**CUESTIONARIOS**

Argimón Pallás JM, Jiménez Villa J. Métodos de investigación clínica y epidemiológica. Barcelona. Ed Elsevier, 3ª ed 2004. Cap 21: Validación de cuestionarios. p. 196-206.

**Tabla: Características a considerar en la validación de un cuestionario**

Métodos de investigación clínica y epidemiológica, pág. 196

**1º Viabilidad:** sencillo, viable y aceptado.

**2º Fiabilidad:**

**Intraobservador = test-retest o repetibilidad** (Se mide mediante la Kappa de Cohen si la variable es cuantitativa categórica, y mediante el Coeficiente de Correlación si es cuantitativa continua)

**Interobservador** (Se mide mediante la Kappa de Cohen si la variable es cuantitativa categórica, y mediante el Coeficiente de Correlación si es cuantitativa continua)

**Consistencia interna** (Se mide mediante el Coeficiente alfa de Cronbach si es cuantitativa continua, o el coeficiente de Kuder-Richardson KR-20 si es dicotómica)

**3º Sensibilidad al cambio** (Se mide mediante la diferencia de medias, y/o el tamaño del efecto, también llamado diferencia estandarizada de medias o *d* de Cohen).

**4º Validez:**

**Validez lógica**

**Validez de contenido**

**Validez de criterio** **(validez para predecir un criterio externo)**

**Validez de constructo**

****

***I. VIABILIDAD.***

Los mejores instrumentos no sirven si su aplicación resulta compleja y costosa. Características como el tiempo empleado en la cumplimentación del cuestionario, la sencillez y la amenidad del formato, y el interés, la brevedad y la claridad de las preguntas, así como la facilidad de puntuación, el registro y la codificación, y la interpretación de los resultados, son aspectos para la viabilidad.

***II. FIABILIDAD.***

**1º Fiabilidad intra-observador = test-retest o repetibilidad** (Se mide mediante el Coeficiente de Correlación si la variable es cuantitativa continua, o mediante el Coeficiente de Concordancia mediante la Kappa de Cohen si la variable es cuantitativa categórica).

Cuando un cuestionario se administra a la misma población en dos ocasiones diferentes en el tiempo, se obtienen resultados idénticos o similares.

Las limitaciones de la evaluación de la repetibilidad de un cuestionario son las mismas que las de cualquier otro instrumento de medida. Si el tiempo trascurrido entre ambas aplicaciones del cuestionario es muy largo, el fenómeno que se mide puede haber presentado variaciones, mientras que, si es demasiado corto, puede existir un recuerdo de las respuestas dadas en la primera ocasión. En ambos casos se obtendrá una medida distorsionada de la repetibilidad. Además, algunos participantes pueden no aceptar que se les administre el cuestionario en dos ocasiones, sobre todo si es extenso.

**2º Fiabilidad inter-observador** (Se mide mediante el Coeficiente de Correlación si la variable es cuantitativa continua, o mediante el Coeficiente de Concordancia mediante la Kappa de Cohen si la variable es cuantitativa categórica).

La evaluación de la fiabilidad inter-observador consiste en estimar el grado de concordancia entre dos o más evaluadores (observadores) independientes. La demostración de una alta fiabilidad inter-observador implica que la fiabilidad intra-observador también es alta. No obstante, si la fiabilidad inter-observador es baja, no se puede asegurar si se debe a la existencia de diferencias entre los observadores o se debe a un solo observador.

**3º Consistencia interna entre las preguntas de una dimensión** (Se mide mediante el Coeficiente alfa de Cronbach).

**Se refiere a si dos o más ítems (preguntas) que miden un mismo atributo (dimensión) presentan homogeneidad entre ellos.**

Los cuestionarios se desarrollan para medir separadamente diferentes componentes o dimensiones (factores) de un problema. Un cuestionario de salud suele estar dividido en preguntas que tratan de medir salud física y mental, o un cuestionario de satisfacción en apartados que identifican, por ejemplo, los componentes de competencia profesional, las cualidades personales del profesional sanitario y la accesibilidad de los servicios. En todas estas situaciones es de esperar que exista una buena homogeneidad entre las distintas preguntas que miden una misma dimensión (carácter, categoría o atributo). Si en un cuestionario de satisfacción, los usuarios contestan que su médico se preocupa bastante de ellos como persona, es de esperar que, al preguntarles si su médico está dispuesto a escucharles, contesten afirmativamente. De lo contrario se pensará que los distintos ítems que componen la satisfacción con el médico que les atiende son poco consistentes entre sí y que el cuestionario es poco fiable.

Cuando un cuestionario está compuesto por diferentes **variables sintéticas[[1]](#footnote-1)**, cada una de las cuales pretende medir una dimensión diferente del fenómeno, debe evaluarse la consistencia interna de los ítems (preguntas) que constituyen cada una de las variables sintéticas.

La técnica estadística para su análisis es Coeficiente alfa de Cronbach, que expresa la consistencia interna entre los varios ítems de una dimensión. Sus valores están comprendidos entre 0 y 1, y su interpretación es similar al Coeficiente de Correlación.

Pueden calcularse diferentes valores del alfa de Cronbach con las respuestas de los 5 ítems que constituyen una dimensión y ver si el valor mejora, empeora o se mantiene igual eliminando uno de los ítems, pues es la forma en la que se intenta su posible eliminación. De todas formas, antes de decidir eliminar un ítem (pregunta), debe evaluarse si esto puede afectar a la validez (ver luego) del cuestionario, ya que puede ser preferible mantener la pregunta aún a costa de una consistencia interna igual o incluso ligeramente menor.

Como norma general, se sugiere que el valor del Coeficiente alfa de Cronbach ha de ser igual o superior a 0,7 para considerar que un instrumento tiene buena consistencia interna.

**4º Fuentes de error de la fiabilidad.**

La fiabilidad de una medida puede afectarse por algunos de los factores siguientes:

**1. Cambios a través del tiempo en las características estudiadas (PRESENTISMO).** Al repetir el cuestionario se debe tener en cuenta qué medidas son susceptibles de variar con el tiempo. De hecho, muchas actitudes, creencias o estilos de vida pueden hacerlo, como la frecuencia de cepillado de los dientes o la práctica de ejercicio físico. Si ha existido un cambio, una repetibilidad baja no implica necesariamente una escasa fiabilidad del cuestionario.

**2. Cambios debidos a las condiciones de administración del cuestionario.** Algunos factores personales del entrevistado, como el estado emocional, el cansancio, el estado de salud o las condiciones del entorno (calor, frío, etc.), pueden influir en el modo de contestar a las preguntas y alterar la fiabilidad.

**3. Variaciones debidas al propio cuestionario.** En ocasiones el formato del cuestionario, la formulación de las preguntas o las instrucciones para cumplimentarlo son poco comprensibles y pueden ser interpretadas de forma distinta por el encuestado al repetir la prueba. Por consiguiente, hay que insistir que las instrucciones sean cortas y precisas y que las preguntas estén definidas de forma operativa con un lenguaje claro y sin ambigüedades.

**4. Cambios atribuidos a los encuestadores.** Los encuestadores deben atenerse a la estructura y secuencia del cuestionario y a cómo han estado formuladas las preguntas. La transcripción de las respuestas a las preguntas abiertas debe ser literal. Al poner en marcha un estudio es de suma importancia entrenar a los entrevistadores para evitar estos errores y conseguir al mismo tiempo que existe uniformidad ente ellos.

**5. Errores en el manejo de los datos.** Estos errores se pueden producir al codificar, grabar y /o transformar las variables para su análisis.

NOTA ADICIONAL: Que las preguntas de una encuesta sean fiables no demuestra que sean necesariamente válidas. Por ejemplo, los encuestados pueden sobreestimar uniformemente al número de horas que trabajan o la cantidad de alcohol que consumen. Por lo tanto, aunque los resultados sean reproducibles (fiables), pueden ser uniformemente erróneos. Así pues, un instrumento para encuesta puede ser reproducible (fiable), pero no válido. Sin embargo, si no es fiable no será válido.

***III. SENSIBILIDAD AL CAMBIO.*** (Se mide mediante la diferencia de medias, y/o el tamaño del efecto, también llamado diferencia estandarizada de medias o *d* de Cohen).

El concepto de **sensibilidad (*sensitivity*)** se refiere a la capacidad de un instrumento para detectar los casos verdaderos.

**El concepto sensibilidad al cambio (*responsiveness*) es la capacidad que tiene un instrumento para detectar numéricamente los cambios clínicos importantes en el atributo verdadero que se mide.** La sensibilidad al cambio es un aspecto crucial en los ensayos clínicos, la valoración de programas y los análisis de coste-utilidad; en otras palabras, cuando el instrumento se utiliza como variable de respuesta. Cuando interesa que un instrumento sea sensible al cambio, hay que prestar atención a la escala numérica con la que se miden las respuestas, es decir, cuántos más puntos haya en las categorías de respuesta, más sensible será el instrumento a los cambios significativos.

La sensibilidad al cambio se mide con la “diferencia de medias” y/o con el “tamaño del efecto (effect size)”, que relaciona la media de las diferencias entre las puntuaciones antes y después de una intervención (numerador), con la desviación estándar de la puntuación antes de la intervención (denominador).

El conocimiento de la sensibilidad al cambio de un instrumento es importante para la estimación del cálculo del tamaño muestral necesario para demostrar los cambios debidos al tratamiento en un ensayo clínico. Cuando mayor sea la sensibilidad al cambio del instrumento, menor será el tamaño de muestra necesario.

**20011231-EvPrD, Valid EVA dolor, 13 cm es mín cambio clín signific. Gallagher**

Gallagher EJ, Liebman M, Bijur PE. Prospective validation of clinically important changes in pain severity measured on a visual analog scale. Ann Emerg Med. 2001 Dec;38(6):633-8.

Gallagher y col llevaron a cabo un estudio prospectivo con 96 pacientes en dos servicios de urgencias urbanos en los que evaluaron a los pacientes aquejados de dolor, mediante una Escala Visual Analógica (EVA) de Dolor de 100 mm, a los 30, 60, 90 y 120 minutos, pidiéndoles que contestaran simultáneamente si su dolor era: “mucho menor”, “un poco menor”, “aproximadamente el mismo”, “un poco mayor” y “mucho mayor” respecto al test anterior.

A priori definieron “mínimo cambio clínicamente significativo” como la diferencia (en milímetros) entre la puntuación EVA actual y la inmediatamente anterior, cuando los pacientes informaban “un poco más” o “un poco menos”.

Hubo 141 pares de mediciones consecutivas en que los pacientes informaron como “un poco más” o “un poco menos” dolor. Esta diferencia clínicamente significativa se asoció con una diferencia medias de 13 mm (IC 95%, 10 a 16) en la EVA de dolor.

***IV. VALIDEZ.***

Grado en que el test logra medir lo que se propone medir. La validez se refiere a la capacidad de un test para la medición numérica del rasgo variable inobservable, constructo teórico, variable latente o atributo psicológico, para el que se ha diseñado. Tiene diferentes aspectos o componentes, que deben ser evaluados en lo posible. Éstos no son independientes entre sí, pues mantienen una relación de influencia recíproca y conforman las diversas partes de un todo, que es lo que conocemos de forma genérica como “validez”.

**1º Validez lógica.**

La validez lógica o aparente (*face validity*) es el grado en que parece que un cuestionario, una parte de él o un ítem (pregunta), mide lo que quiere medir. Lo ideal es que todos los términos estén definidos en un glosario mediante definiciones operacionales preferentemente, o descriptivas.

Hay que evitar en lo posible la ambigüedad y los eufemismos por conveniencia y deseabilidad social. Por ejemplo: *“¿En las dos últimas semanas ha experimentado algún cambio en su sentido de trascendencia?”* Si el concepto “trascendencia” no está definido, su ambigüedad impide la validez lógica de la respuesta.

La decisión sobre si las preguntas deben tener o no validez lógica ha de tomarse antes de iniciar su redacción. Si las preguntas carecen de validez lógica es muy probable que los encuestados rechacen contestar. De todos modos, en alguna ocasión puede ser de interés formular preguntas con una validez lógica inferior a la menor posible. Por ejemplo, cuando se trata de temas muy sensibles, conflictivos o que no están bien vistos socialmente, si se realizan preguntas directas (con mucha validez lógica), es muy probable que el encuestado no responda o falsee la respuesta, por lo que puede ser preferible realizar preguntas que aborden el tema de una forma más indirecta, aunque con una menor validez aparente.

**2º Validez de contenido.**

El objetivo que persigue llevar a cabo un estudio de validación del contenido es analizar hasta qué punto los elementos que componen el test son una muestra relevante y representativa del constructo sobre el que se van a realizar las inferencias, lo que puede estimarse a través de la opinión de expertos.

La relevancia implica la necesidad de una clara y exhaustiva especificación de todas las posibles conductas observables del constructo a medir (especificación del dominio de conductas). La representatividad hace referencia a la necesidad de que todas esas conductas estén representadas en el test (representatividad del dominio).

**Ejemplo de baja validez de contenido**

Veamos un ejemplo sobre la búsqueda del candidato más apropiado para un empleo.

Una escuela quiere contratar **a un nuevo profesor de ciencias** y un panel de directores comienza a analizar a los distintos candidatos. Elaboran una lista acotada y luego diseñan una prueba, con la que acaban escogiendo al candidato con mejor puntuación. Lamentablemente, éste demuestra ser un **maestro de ciencias naturales mediocre**.

Después de ver la prueba, la junta de educación comienza a ver en qué se equivocó. La gran mayoría de las preguntas fueron sobre matemáticas, entonces, obviamente, la escuela encontró al mejor profesor de matemáticas.

Sin embargo, estaban buscando que el profesor de ciencias enseñara biología, química y psicología. La validez de contenido de la prueba fue pobre y no representó plenamente **el constructo de "bondad de profesor de ciencias (naturales)”**

Rectificando, **la escuela rediseñó la prueba** y la sometió a un panel de expertos en educación. Después de pedirles a los candidatos que completaran el nuevo examen, la escuela encontró a otra persona que demostró ser **una profesora de ciencias naturales excelente y completa**. Esta prueba tuvo una validez de contenido mucho más elevada y representó totalmente cada elemento del constructo.

**3º Validez de criterio (validez para predecir un criterio externo)**

Se mide la respuesta al cuestionario frente a un criterio externo observable. Cuando la variable es dicotómica, un ejemplo es la validación (medición del Valor Predictivo Positivo y Negativo) del test de sangre oculta en heces frente a la mortalidad por Cáncer ColoRectal en 10 años. Cuando la variable es cuantitativa continua, puede utilizarse el Coeficiente de correlación de Pearson (cuestionario de bondad de vendedor frente al número de ventas en una semana), y con más precisión el Coeficiente de Correlación Intraclase.

En ocasiones se puede disponer de algún método alternativo de medida del fenómeno estudiado cuya validez ha sido demostrada, que se toma como referencia para determinar la validez de la encuesta. Siempre que se disponga de un método de referencia adecuado, deberá evaluarse la validez de criterio del cuestionario.

Un ejemplo: se puede **validar el consumo de tabaco declarado**, comparándolo con los valores obtenidos de nicotina en sangre o de monóxido de carbono en el aire espirado. En otros casos el investigador tendría que “fiarse” de medidas menos objetivas como la historia clínica o los resultados obtenidos mediante otro cuestionario.

Otro ejemplo: las manifestaciones de los individuos sobre el uso de su medicación que se han validado comparándolas con los registros farmacéuticos de recetas.

La validez de criterio puede evaluarse de dos formas: la validez concurrente y la validez predictiva. Para valorar la **validez concurrente** se relaciona la nueva medida con la de referencia, siendo ambas administradas simultáneamente. Cuando el criterio de referencia no esté disponible hasta un tiempo después (por ejemplo, el desarrollo de una enfermedad), se valora hasta qué punto la nueva medida es capaz de predecirlo correctamente y se habla de **validez predictiva**.

**4º Validez de constructo** [Un constructo alude a un concepto teórico psicológico inobservable].

A veces el constructo no tiene criterio externo objetivable (con lo que es imposible estimar la validez de criterio externo). Otras veces existe, pero no está al alcance del investigador. En estos casos, el procedimiento más empleado es evaluar la validez de constructo que necesita, al menos, medir la validez discriminante y la validez convergente.

1) La **validez discriminante** se refiere a la capacidad para distinguir entre subgrupos de pacientes o individuos con y sin el constructo (o atributo de interés), o también discriminar con distintos niveles del constructo (atributo de interés). Por ejemplo, es de esperar que la **calidad de vida relacionada con la salud** será peor en los pacientes asmáticos con gran afectación funcional que en aquellos con formas más leves.

El método más sencillo para evaluar la validez discriminante es el de los grupos extremos, que consiste en administrar el cuestionario a dos grupos de individuos: uno con la característica o conducta de interés, y otro que carece de ella, y luego medir el coeficiente de correlación entre las respuestas de ambos grupos. Es de esperar que la correlación sea nula o baja, porque si la correlación fuera alta, entonces no está discriminando el rasgo latente o constructo. Este enfoque presenta dos problemas. El primero es la propia definición de grupos extremos, ya que no siempre existe un criterio adecuado para conocer quién tiene y quién no tiene la característica de interés. En este caso se puede dividir la muestra en función de la puntuación obtenida con el propio instrumento, seleccionando por ejemplo el tercil con mejores puntuaciones y el tercil con peores puntuaciones. El segundo problema es similar al que se presenta al evaluar una prueba diagnóstica: puede ser relativamente fácil discriminar entre dos grupos muy extremos, pero ésta no es la utilidad que se pretende dar al instrumento en la práctica habitual. Por tanto, comprobar que un cuestionario es útil para diferenciar entre dos grupos extremos no es suficiente para demostrar su validez de constructo.

Ejemplo:

**19961015-EvPrDia, Validez y fiabilidad Cuestionario función familiar Apgar. Bellón**

Bellón Saameño JA, Delgado Sánchez A, Luna del Castillo JD, Lardelli Claret P. Validez y fiabilidad del cuestionario de función familiar Apgar-familiar. Aten Primaria. 1996 Oct 15;18(6):289-96.

Este estudio, que tenía por objeto evaluar la validez de constructo y fiabilidad de un cuestionario de función familiar en una muestra de 656 pacientes, **se utilizó el método de los grupos extremos para evaluar la validez de constructo “función familiar”**. Se partió de la hipótesis de que los casados tienen una mejor **función familiar** que los divorciados, por lo que deberían obtener puntuaciones superiores en el cuestionario.

Al aplicarlo hubo 127 [viudo, separado o divorciado], con una odds de Apgar disfuncional / funcional de 43/84 = 0,52, frente a 529 [casado o soltero], con un odds de Apgar disfuncional / funcional de 63/466 = 0,12; con lo cual OR = 3,79 (IC 95%, 2,41-5,95).

**2)** La segunda condición para evaluar la validez del constructo medir la **validez convergente**. Consiste en comprobar que las respuestas del cuestionario cuyo constructo se quiere validar, tienen un alto coeficiente de correlación con las respuestas de un cuestionario con un constructo de similar significado, generalmente ya acreditado. Si se obtiene una correlación baja, la validez del constructo es nula (o baja).

Ejemplo:

**19961015-EvPrDia, Validez y fiabilidad Cuest apoyo social Duke-UNC-11. Bellón**

Bellón Saameño JA, Delgado Sánchez A, Luna del Castillo JD, Lardelli Claret P. Validez y fiabilidad del cuestionario de apoyo social funcional Duke-UNC-11. Aten Primaria. 1996 Sep 15;18(4):153-6, 158-63.

Este estudio evaluó la validez de constructo y la fiabilidad del cuestionario de **apoyo social** funcional Duke-UNC-11 en una muestra de 656 pacientes. **Para evaluar la validez de constructo “apoyo social”**, los autores eligieron determinadas características que, según la literatura, están relacionadas con el **apoyo social** y que son la edad, el estado civil, el estado de salud, la utilización de servicios, la salud mental, la función familiar y el número de convivientes. Para ello, compararon las distribuciones de estas variables entre los sujetos con un **apoyo social normal o bajo**, según el cuestionario, y determinaron los coeficientes de correlación entre sus puntuaciones en el cuestionario evaluado y el resto de las escalas cuantitativas.



**5º Fuentes de error de la validez en el diseño de los cuestionarios.**

**1. Orden de las preguntas.** Es conveniente situar las más conflictivas al final del cuestionario, ya que si se ponen al principio existe la posibilidad de que la persona encuestada rechace seguir contestando o no lo haga con la sinceridad deseada.

**2. La redacción de las preguntas o el efecto marco.** En 1981, a dos grupos de estudiantes de la Universidad de Stanford y British Columbia, aleatoriamente distribuidos, se les planteó la siguiente situación: Imagínese que USA está preparándose para un brote de una inusual enfermedad asiática, con la que se espera que mueran 600 personas. Se proponen dos programas alternativos para combatir la enfermedad.

Al grupo 1 se le comunica así:

A) Si se adopta el programa A, salvará a 200 personas;

B) Si se adopta el programa B, hay un tercio de probabilidades de que 600 personas se salven y dos tercios de que ninguna persona se salve.

En esta versión del problema (aversión al riesgo como ganancia), con una muestra N= 152, el 72% respondió a favor del programa A y el 28% a favor del B.

Al grupo 2 se le comunica así:

A’) Si se adopta el programa A’, 400 personas morirán;

B’) Si se adopta el programa B’, hay un tercio de probabilidades de que nadie muera y dos tercios de probabilidades de que mueran las 600 personas.

En esta versión del problema (aversión al riesgo como pérdida), con una muestra N= 155, el 22% respondió a favor del programa A’ y el 78% a favor del B’.

Aunque no hay sustancial diferencia entre ambas versiones, cada una evoca diferentes asociaciones emocionales y evaluaciones. El marco de riesgo se refiere a que es frecuente elegir “no correr riesgo” cuando la situación se presenta como ganancia segura y es frecuente elegir “correr riesgos” cuando la situación se presenta como pérdida segura.

**3. Errores de categorización de las respuestas.** En los cuestionarios con opciones de respuestas predeterminadas hay que tener en cuenta todas las posibilidades de respuesta, de modo que el encuestado encuentre siempre su opción adecuada.

En otras ocasiones, son los factores personales del entrevistado los que introducen los sesgos en las respuestas. Hay individuos que, al ser encuestados, tienden a contestar de forma distinta a como lo harían normalmente.

**4. Sesgo de conveniencia social.** Algunas personas tienden a dar la respuesta más aceptable socialmente, o la que piensan que contestará la mayoría. Este sesgo depende de muchos factores individuales, como la edad, sexo, clase social y el contexto en el que se realiza la pregunta. Ejemplos son las preguntas sobre el alcohol o el aborto. Un mecanismo para intentar evitarlo es dar en primer lugar las opciones menos convenientes, de modo que sea más fácil seleccionarlas. En muchas ocasiones, el sesgo de conveniencia social se produce sin que el entrevistado tenga intención de engañar, mientras que en otras no dice lo que piensa de forma deliberada.

**5. Tendencia sistemática a dar siempre la misma respuesta.** Como contestar “sí”, “verdadero”, o “de acuerdo” a las cuestiones planteadas. En el caso más extremo estas respuestas se dan independientemente del contenido de la pregunta. Como ejemplo, se podría dar el caso de un individuo que conteste afirmativamente a la pregunta de si toma la medicación todos los días a la hora indicada, y también lo haga cuando se le preguntara si a menudo se le olvida tomar la medicación. En el extremo opuesto están los que contestan sistemáticamente “no” o “no de acuerdo”.

**6. Características del encuestador.** La respuesta de una pregunta a la persona de si cumple con la medicación prescrita puede ser muy distinta si la realizan su médico o lo hace un entrevistador no implicado directamente en su cuidado médico.

**7. Elección de la categoría o puntuación intermedia.** En las escalas de puntuación de Likert existe la posibilidad, además, de que aparezcan otros sesgos. Uno de ellos se refiere a que algunos individuos siempre escogen la categoría o puntuación intermedia. El efecto de ese sesgo es la reducción, en la práctica, de las posibles respuestas. Así si se usa la escala de Likert (con cinco opciones), las extremas recogen la posibilidad de “siempre” y ”nunca”, un sesgo de aversión hacia los extremos implica que la escala queda reducida a tres categorías, ya que el individuo no contesta ni “nunca” ni “siempre”, con la consiguiente pérdida de fiabilidad. Existen dos posibles soluciones a este problema. La primera es evitar los términos absolutos, y en su lugar utilizar “casi nunca” y “casi siempre”. La segunda consiste en aumentar las categorías posibles en cada respuesta, de modo que, si en realidad se desea una pregunta con cinco categorías de respuesta, finalmente tenga siete, siendo las categorías extremas las ocupadas por los términos absolutos.

***V. ADAPTACIÓN TRANSCULTURAL DE LOS FORMULARIOS Y SU POSTERIOR VALIDACIÓN.***



Practicar el cuestionario en una población representativa para la VALIDACIÓN

1. Es mejor la denominación **variable sintética** que escala, pues las escalas son ordinales, numéricas, etc. Las escalas se utilizan para construir (constructo) variables sintéticas que pretendan medir o estimar la cantidad de un rasgo o carácter abstracto, tales como la capacidad de liderazgo, el dolor que experimenta un politraumatizado, la gravedad de un proceso asmático, la discapacidad funcional del anciano, la calidad de vida de un trasplantado renal o el grado de sentimiento entre acuerdo y desacuerdo sobre una pregunta. [↑](#footnote-ref-1)