

Construcción y validación de instrumentos de medición en el ámbito de la salud. Revisión de Literatura

Construction and validation of measurement instruments in the field of health. Literature Review

Recibido: 20/05/2022

Aceptado: 14/06/2022

Zenaida Rojas-Apaza

[orcid 0000-0001-5443-171X](#)

Estudiante de Maestría en Estomatología de la Facultad de Odontología, Universidad Nacional Mayor de San Marcos. Lima, Perú

Gilmer Torres-Ramos

[orcid 0000-0002-2590-6736](#)

Docente de la Facultad de Odontología, Universidad Nacional Mayor de San Marcos. Lima, Perú

Enna Lucila Garavito-Chang

[orcid 0000-0003-2925-8630](#)

Docente de la Facultad de Odontología, Universidad Norbert Wiener. Lima, Perú

Sara Castañeda-Sarmiento

[orcid 0000-0001-9240-6232](#)

Docente de la Facultad de Odontología, Universidad Nacional Mayor de San Marcos. Lima, Perú

Roxana Patricia López-Ramos

[orcid 0000-0003-4374-9575](#)

Docente de la Facultad de Odontología, Universidad Nacional Mayor de San Marcos. Lima, Perú

RESUMEN

El proceso de construcción y validación de instrumentos de medición en el ámbito de la salud es un tema complejo que comprende diversas etapas, la información existente de este proceso metodológico es dispersa y poco clara. El objetivo de este artículo es sintetizar y proponer recomendaciones basadas en la literatura científica sobre un adecuado proceso metodológico para la construcción y validación de instrumentos de medición en el ámbito de la salud. Se realizó una exhaustiva revisión de la literatura en bases de datos como Medline, Scielo, Science direct y Google Scholar entre 1994 y el 2020, relacionados a construcción y validación de instrumentos de medición. Se obtuvieron 62 artículos, después de la lectura y exclusión, se seleccionó 34 artículos considerados adecuados para cumplir con los objetivos planteados. En base a la información encontrada se describen las etapas para un adecuado proceso de construcción y validación, se presentan las propiedades psicométricas como validez y confiabilidad, además de las diversas pruebas estadísticas utilizadas para garantizar un instrumento válido y confiable.

Palabras claves: Validez; Confiabilidad; Instrumentos de medición.

Citar como Rojas Z, Torres G, Garavito E, Castañeda S, López R. Construcción y validación de instrumentos de medición en el ámbito de la salud. Revisión de literatura. *Odontol Pediatr* 2022;21 (1); 46 - 52.

Abstract

The process of construction and validation of measurement instruments in the field of health is a complex issue that includes various stages, the existing information on this methodological process is scattered and unclear. The objective of this article is to synthesize and propose recommendations based on the scientific literature on an adequate methodological process for the construction and validation of measurement instruments in the field of health. An exhaustive review of the literature was carried out in databases such as Medline, Scielo, Science direct and Google Scholar between 1994 and 2020, related to the construction and validation of measurement instruments. 62 articles were obtained, after reading and exclusion, 34 articles considered adequate to meet the stated objectives were selected. Based on the information found, the stages for an adequate construction and validation process are described, the psychometric properties such as validity and reliability are presented, in addition to the various statistical tests used to guarantee a valid and reliable instrument.

Keywords: Validity; Reliability; Measurement instrument.

INTRODUCCIÓN

Es muy frecuente que, en estudios de investigación en el ámbito de la salud, se haga uso de diversas herramientas e instrumentos de medición, dentro de ellos se encuentran los cuestionarios, que forman una parte importante para el desenvolvimiento de la práctica clínica, la evaluación del estado de salud y el aporte a la investigación, el resultado de estos instrumentos tiene gran influencia respecto a las decisiones, tratamientos, programas de salud pública y políticas institucionales¹. Razón por la cual los instrumentos deben de cumplir dos propiedades psicométricas importantes como la validez, el cual representa la veracidad de los resultados y la confiabilidad que hace referencia a la estabilidad de los mismos; lo cual nos garantiza un resultado adecuado y veraz^{2,3}. Es importante tener presente que no todo instrumento es válido y confiable. La validación de instrumentos de

investigación es un proceso importante, complejo y muchas veces confuso, algunos investigadores hacen uso instrumentos que no han sido validados adecuadamente, pues el desconocimiento de un correcto proceso de validación los lleva a elegir erróneamente el instrumento a utilizar². Es necesario que el investigador compruebe que el instrumento se encuentre correctamente validado, solo así medirá lo que pretende medir de manera constante, a lo largo del tiempo, dentro de un entorno y población específica⁴.

Si bien es cierto existen muchos instrumentos de medición validados adecuadamente, disponibles para el uso del investigador, pero muchas veces se necesita construir nuevos instrumentos de medición para recoger datos sobre fenómenos no explorados con anterioridad⁴. La creación de un nuevo instrumento

de medición es complejo, utiliza varios recursos y requiere del aporte de capacidades y conocimientos de diferentes áreas, es por ello que el investigador tiene que estar capacitado para el desarrollo de la herramienta¹.

El presente estudio tiene como objetivo: Sintetizar y proponer recomendaciones basadas en la literatura científica sobre un adecuado proceso metodológico para la construcción y validación de instrumentos de medición en el ámbito de la salud

Es muy frecuente que, en estudios de investigación en el ámbito de la salud, se haga uso de diversas herramientas e instrumentos de medición, dentro de ellos se encuentran los cuestionarios, que forman una parte importante para el desenvolvimiento de la práctica clínica, la evaluación del estado de salud y el aporte a la investigación, el resultado de estos instrumentos tiene gran influencia respecto a las decisiones, tratamientos, programas de salud pública y políticas institucionales¹. Razón por la cual los instrumentos deben de cumplir dos propiedades psicométricas importantes como la validez, el cual representa la veracidad de los resultados y la confiabilidad que hace referencia a la estabilidad de los mismos; lo cual nos garantiza un resultado adecuado y veraz^{2,3}. Es importante tener presente que no todo instrumento es válido y confiable. La validación de instrumentos de investigación es un proceso importante, complejo y muchas veces confuso, algunos investigadores hacen uso instrumentos que no han sido validados adecuadamente, pues el desconocimiento de un correcto proceso de validación los lleva a elegir erróneamente el instrumento a utilizar². Es necesario que el investigador compruebe que el instrumento se encuentre correctamente validado, solo así medirá lo que pretende medir de manera constante, a lo largo del tiempo, dentro de un entorno y población específica⁴.

Si bien es cierto existen muchos instrumentos de medición validados adecuadamente, disponibles para el uso del investigador, pero muchas veces se necesita construir nuevos instrumentos de medición para recoger datos sobre fenómenos no explorados con anterioridad⁴. La creación de un nuevo instrumento de medición es complejo, utiliza varios recursos y requiere del aporte de capacidades y conocimientos de diferentes áreas, es por ello que el investigador tiene que estar capacitado para el desarrollo de la herramienta¹.

El presente estudio tiene como objetivo: Sintetizar y proponer recomendaciones basadas en la literatura científica sobre un adecuado proceso metodológico para la construcción y validación de instrumentos de medición en el ámbito de la salud.

METODOLOGÍA

Se realizó una búsqueda bibliográfica exhaustiva en diversas bases de datos como Medline, Scielo, Science direct y Google Scholar entre los años 1994 y el 2020; utilizando los terminos: “construction”, “validation”, “ validity”, “reliability”, “instruments measurement”, “health questionnaire”, “scale”, “survey”; uniendolos con el operador booleano AND. Los criterios de inclusión fueron: artículos en idioma español o inglés, publicados entre los años 1994 y 2020, que incluyeran temas relacionados sobre construcción, validación de instrumentos de medición y sus aspectos metodológicos. Los criterios de exclusión fueron artículos de opinión.

La búsqueda bibliográfica se llevó a cabo en diciembre del 2020, inicialmente se obtuvieron 62 artículos, del los cuales se seleccionaron 34 considerados adecuados para su posterior análisis y el cumplimiento de los objetivos planteados.

RESULTADOS

De la búsqueda inicial realizada en las diferentes bases de datos anteriormente mencionadas se identificó un total de 62 artículos, de los que se seleccionaron 40 por tener relación con el tema de investigación. Se procedió a leer los resúmenes, buscando concordancia con el tema de investigación y descartando aquellos no relacionados con la metodología de Validación de Instrumentos en el Ámbito de la Salud. Finalmente se escogió un total de 34 artículos, cuyo objetivo guarda relación con el presente estudio.

Características psicométricas de un instrumento de medición

Las características psicométricas de un instrumento de medición justifican la fiabilidad de los datos que se obtienen con el mismo. Todo instrumento antes de ser utilizado en el ámbito de la investigación debe estar calibrado para producir resultados con el menor error posible. El uso de estas estrategias se ha vuelto más intensa en las últimas décadas, motivando las buenas prácticas para la obtención del instrumento adecuado. Dentro de las principales características psicométricas de un instrumento de medición tenemos a la validez, confiabilidad, sensibilidad y factibilidad (tabla 1)⁵

Validez

La validez se refiere al grado en que un instrumento mide exactamente lo que pretende medir^{5,2,6,7,3}. Implica qué tan bien la información que será recopilada con el instrumento abarca el área real de investigación⁸. La validez es una pieza fundamental en el diseño de un instrumento de investigación y puede estimarse de diferentes maneras, dentro de

las principales se puede mencionar a la validez de contenido, validez de criterio y validez de constructo. Cada una de manera independiente aporta evidencia a la validación global del instrumento, siendo esta de carácter acumulativo^{7,2}.

Tipos de validez: Validez de contenido, validez de criterio y validez de constructo.

Validez de contenido

Es una valoración cualitativa del instrumento, donde se evalúa si el instrumento diseñado abarca todas las dimensiones del fenómeno que se pretende medir, bajo este argumento se considera que un instrumento es válido en su contenido si abarcará todos los aspectos relacionados con la variable motivo de estudio^{7,9}. De manera global, la validez de contenido evalúa el nuevo instrumento para asegurar que este incluya todos los elementos que son relevantes y elimine aquellos que podrían significar indeseables⁸. Una forma recomendable de establecer este tipo de validez es mediante un “Juicio de expertos”.

Validez de Criterio

La validez de criterio hace referencia al grado en que el instrumento elaborado emite una puntuación similar a los de otros instrumentos que evalúan las mismas dimensiones, por lo tanto, representa una comparación con uno considerado como Gold Standard^{10,11}.

Validez de constructo

La validez de un constructo evalúa el grado en que el instrumento refleja la teoría del fenómeno para el cual fue diseñado. Busca determinar la relación teórica de los ítems del nuevo instrumento con los conceptos

Tabla 1. Características psicométricas de un instrumento de medición⁷

Propiedad	Método	Descripción	Estadístico
V. Contenido	Juicio de expertos	Explora el instrumento según criterios de claridad, pertinencia y relevancia de los ítems a través de la opinión de un panel de expertos	Índice de Lawshe
	V. Concurrente	Mide el grado de correlación entre un instrumento y otra magnitud que mida el mismo criterio	CCI
V. Criterio	V. Predictiva	Mide el grado de correlación entre la medida de un concepto y una medida posterior del mismo concepto. Mide como un instrumento predice una evaluación	CCI
	Grupos extremos	Compara las puntuaciones de diferentes grupos de individuos	t de student Mann Whitney
V. Constructo	V. Convergente	Compara el grado de correlación con otras variables relacionadas con el constructo	CCI
	V. Divergente	Compara el grado de correlación con otras variables no relacionadas con el constructo	CCI
	Análisis factorial	Identifica factores o constructos de un instrumento y la relación de cada ítem con cada factor o constructo	Coeficientes λ o cargas factoriales $\geq 0,3$
	Consistencia interna	Estima el grado en que los ítems se correlacionan entre ellos	Alfa de Cronbach-KR-20
Confiabilidad	Test-retest	Estimar la estabilidad temporal del instrumento	CCI-C Kappa
	Formas paralelas	Compara concordancia entre dos o más instrumentos contruidos para medir el mismo constructo, en diferentes momentos	CCI-C Kappa
	Mitades partidas	Compara las correlaciones entre dos mitades de un instrumento	CCI-C Kappa
Sensibilidad	Confiabilidad Interobservador	Estima el acuerdo entre dos o más evaluadores que evalúan a los mismos individuos con el mismo instrumento	CCI-C Kappa
	Capacidad de un instrumento para detectar cambios a través del tiempo		Prueba de hipótesis
Factibilidad	Capacidad de aplicabilidad en el campo que se quiere utilizar		-

Nota: Tomado de "Características psicométricas de un instrumento", por Carvajal A, Centeno C, Watson R, Martínez M, Sanz Rubiales Á. ¿Cómo validar un instrumento de medida de la salud? An Sist Sanit Navar. 2011;34(1): Pag 67

teóricos, de tal manera que proporciona evidencia para una correcta interpretación de los resultados, pueden ser de dos tipos; validez convergente y discriminante.⁵ El análisis factorial es una de las técnicas estadísticas utilizada para determinar los o dominios del instrumento en desarrollo¹⁰.

Confiabilidad

La confiabilidad es el grado de consistencia con que los ítems de un instrumento miden el fenómeno sin error de medición. Determinará si el instrumento produce resultados consistentes en tiempo y espacio, evaluando aspectos de coherencia, equivalencia, estabilidad y homogeneidad. El cumplimiento de la confiabilidad le brindará al instrumento el criterio esencial de calidad^{2,5}. Algunas opciones para estimar la confiabilidad son: análisis de consistencia interna, test-retest y confiabilidad interobservador.

Sensibilidad

La sensibilidad mide la capacidad que tiene un instrumento para detectar cambios importantes en sujetos evaluados a lo largo del tiempo “posterior a la intervención”, es considerado como un tipo especial de validez de constructo. Se mide con el estadístico “tamaño del efecto” para evaluar la sensibilidad al cambio relacionando la media de las diferencias entre las puntuaciones antes y después de la intervención con la desviación estándar de la puntuación antes de la intervención. Es importante conocer en los instrumentos relacionados con la salud la diferencia mínima relevante para conocer la eficacia de los tratamientos. Esta característica también es importante para estimar el tamaño muestral de un estudio cuanto mayor sea la sensibilidad del instrumento menor será el tamaño de la muestra necesario^{7,12}.

Factibilidad

La factibilidad mide si el instrumento es fácil de utilizarlo en el campo para el cual fue diseñado. Se evaluará el tiempo que se requiere para completarlo, sencillez, formato atractivo, claridad de los ítems, entre otros aspectos para la codificación y la interpretación de los resultados. Es recomendable valorar esta característica en diferentes poblaciones, esto llevará a conocer si el instrumento es adecuado para ser aplicado en sujetos en diferentes situaciones¹².

Metodología Para Validar Un Instrumento De Investigación

Según la revisión de la literatura, se logró identificar tres fases para desarrollar y validar un instrumento de investigación (Fig. 1):

Fase 1: Elaboración y diseño del instrumento

Fase 2: Verificar la validez del instrumento

Fase 3: Verificar la confiabilidad del instrumento

Fase 1: Elaboración y diseño del instrumento

(1) Conceptualización y definición de los objetivos del instrumento

En el ámbito de la salud se utilizan diversos instrumentos como escalas o cuestionarios para medir conocimientos, actitudes, prácticas, intenciones o comportamientos¹³. Como primer paso para la elaboración de estos instrumentos es la conceptualización, pues debemos definir de manera clara, precisa y razonablemente detallada que se desea medir: “definir el constructo”¹⁴.

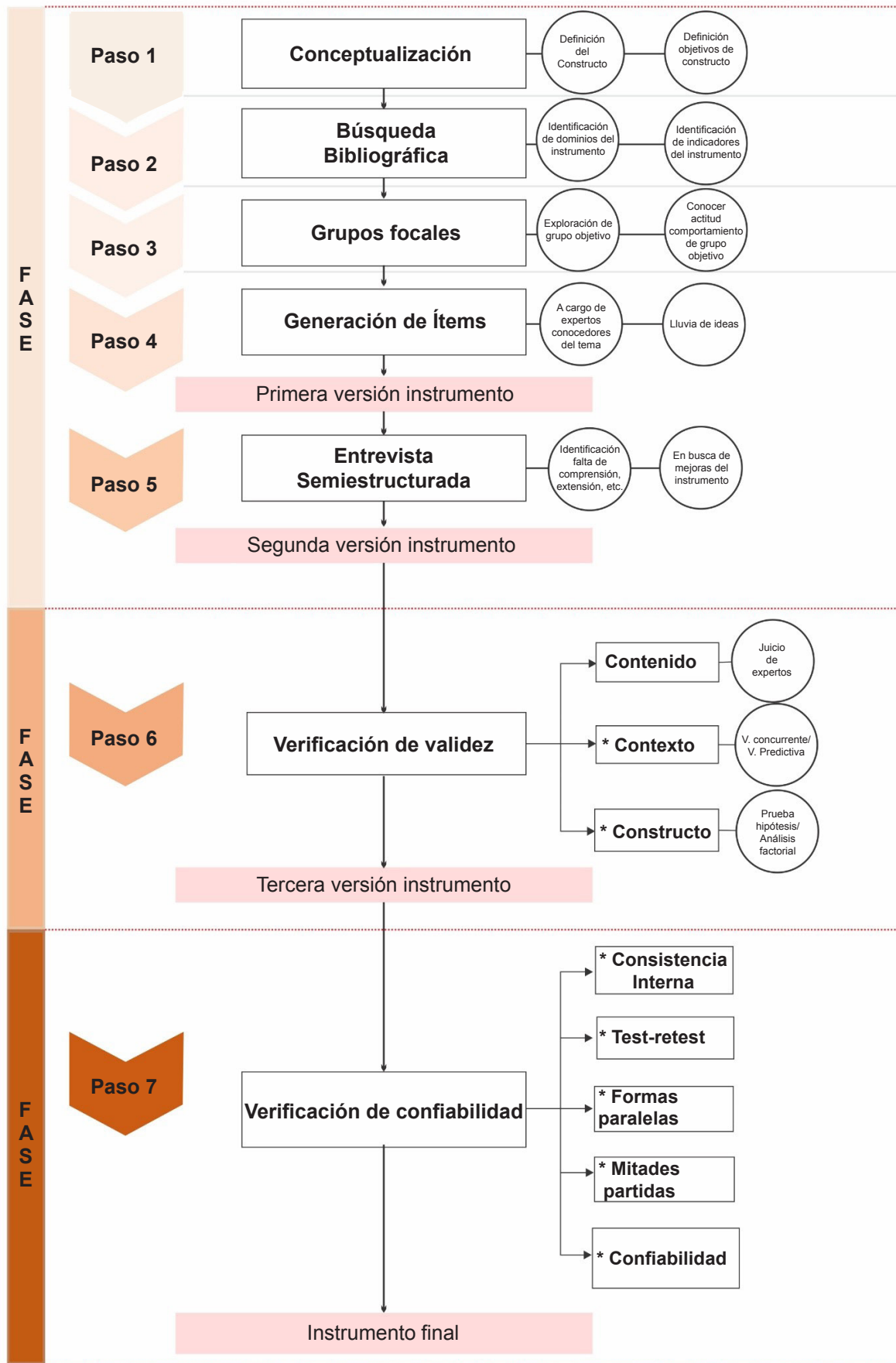


Tabla 2. Valores de Alfa de Crombach. (Adaptado de Aithal)

Valor de alfa de Crombach	Magnitud de confiabilidad	Interpretación
0,01 a 0,20	Muy Baja	Un problema grave en el diseño del cuestionario. Debe descartarse o revisar
0,21 a 0,40	Baja	Baja consistencia interna, escasa interrelación entre elementos. Debe descartarse o revisar
0,41 a 0,60	Moderada	Moderada consistencia y fiabilidad internas en un cuestionario, debe revisarse
0,61 a 0,80	Alta	Alta consistencia y fiabilidad internas en un cuestionario, el investigador debe revisar algunos elementos para mejora
0,81 a 1,00	Muy Alta	Perfecta consistencia interna en un cuestionario dado

Por otro lado es fundamental establecer los objetivos de la medición del constructo, considerando el porqué de la medición, en que personas o público objetivo; de esa manera se podrá definir claramente la población de interés, las características del formato y la forma de administración del instrumento^{6,1}.

(2) Búsqueda bibliográfica, identificación de elementos clave y dominios del instrumento

Es importante realizar una búsqueda bibliográfica exhaustiva, para definir el constructo y de la forma más clara y completa posible, de tal manera que sea factible la identificación de los dominios del nuevo instrumento¹⁵. Una vez definido los dominios se proseguirá a la identificación de indicadores que deriven de dichos dominios, los cuales deben estar correctamente respaldados por evidencia científica producto de la búsqueda bibliográfica^{16,17,18}. también se debe identificar la función que desempeñara el instrumento que se pretende construir, pues se requiere considerar una metodología específica para cada fin; por ejemplo, en el área de diagnóstico su finalidad es detección y por ello es importante su sensibilidad, para determinar estadios preclínicos o clínicos tempranos o para clasificación y/o categorización de los pacientes, tomando relevancia su capacidad de discriminación. En el área de

pronóstico, la finalidad del instrumento es predecir futuros eventos, siendo necesaria su capacidad predictora. En el área de tratamiento, tendrá la finalidad de detección de cambios en el estado de salud a través del tiempo^{6,19}. Además, la revisión exhaustiva de la literatura permitirá al investigador confirmar la inexistencia de instrumentos desarrollados anteriormente con el mismo propósito, de ser así se debe justificar el desarrollo del nuevo instrumento, sustentar una posible mejora teórica y estructural sobre los instrumentos ya existentes y de esa manera contribuir sustancialmente proporcionando una validez significativa sobre otros instrumentos^{11,20,10,15}.

(3) Grupo focal: Exploración de público objetivo

Se considera una etapa de exploración cualitativa, pues el método utilizado consiste en entrevistas a profundidad, consideradas observaciones no estructuradas y grupos focales las cuales permitirán al investigador conocer la actitud, comportamiento y percepción del público objetivo, se considera un paso muy importante para conocer la condición subjetiva de la población a estudiar, esto contribuirá a delimitar aspectos teóricos, y estructura general del instrumento^{21,22}.

(4) Desarrollo de los ítems

Una vez definido los dominios del instrumento, se debe iniciar la redacción de los ítems, estos deben estar a cargo de un grupo de expertos conocedores del tema a investigar, así mismo se sugiere utilizar una técnica de lluvia de ideas^{18,1}.

Se debe optar por dos formas para desarrollar los ítems adecuados: la primera se genera mediante un método deductivo esta se da posterior a la descripción del dominio relevante e identificación de elementos, mediante la revisión de la literatura y evaluación instrumentos existentes; la segunda es el método inductivo la cual implica el desarrollo de los ítems tomando en cuenta la exploración cualitativa de la población objetivo. Es importante considerar como una buena práctica la combinación del método deductivo e inductivo para el desarrollo de los ítems.

Pues una adecuada revisión de la literatura proporcionará la base teórica, por otro lado, la exploración cualitativa nos permitirá una interpretación de las necesidades de la población¹¹. Los ítems elaborados deben de incorporarse a un constructo denominado instrumento inicial, se recomienda que esta primera versión tenga como mínimo el doble de ítems que la versión final, deben de ser redactados en base al público objetivo por ejemplo si son diseñados para pacientes: en lenguaje sencillo, evitar términos técnicos y desconocidos y si su finalidad es personal de salud podrá redactarse con terminología médica.

Finalmente, el investigador debe asegurarse que los ítems no emitan mensajes ofensivos o altamente sesgados en términos de identidad social, como: género, etnia, religión, raza, situación económica u orientación sexual^{11,23,24}.

(5) Prueba piloto preliminar: entrevista semiestructurada

Es recomendable realizar una prueba piloto preliminar a un grupo pequeño de la muestra 30 a 50 individuos, mediante una entrevista semiestructurada el investigador tendrá la oportunidad de identificar la falta de comprensión, extensión adecuada o dificultades en las opciones de respuesta de los enunciados, además recoger sugerencias de la población objetivo para mejoras en el instrumento final^{25,26,27}. La entrevista semiestructurada posee mayor flexibilidad a diferencia de las estructuradas, parten de preguntas planeadas, las mismas que pueden ajustarse a los entrevistados. La ventaja es que puede adaptarse a los sujetos con grades posibilidades para motivar al interlocutor, aclarar términos, identificar ambigüedades y reducir formalismos²⁸.

Fase 2: Verificar la validez del instrumento

Esta fase está orientada a verificar si el instrumento elaborado mide lo que pretendemos medir. Tiene un enfoque cualitativo-cuantitativo y representa el primer paso para determinar asociación entre conceptos abstractos con indicadores observables y posibles de medir¹. Es considerado como un proceso acumulativo y/o sumativo; dentro de los procedimientos utilizados para evaluar la validez de un instrumento tenemos tres técnicas frecuentemente utilizadas: validación de contenido, validación de criterio y validación de constructo.

(6) Validez de contenido: Juicio de expertos:

Denominado también "Juicio de Expertos", viene a ser una opinión informada de personas con experiencia y trayectoria, conocedores del tema motivo de

estudio, se considera que dichas personas pueden aportar información, evidencia, y juicios de valor. Según Skjong y Wentworht (2000) para la elección de los jueces expertos es recomendable tener en cuenta ciertos criterios como: (1) Experiencia en realización de juicios de valor y toma de decisiones basada la experticia, (2) reconocimiento y prestigio en la comunidad, (3) disponibilidad para apoyar, y (4) imparcialidad, confianza en sí mismo y adaptabilidad. Otros expertos también recomiendan considerar el entrenamiento, experiencia, entre otros aspectos.²⁹ Según la recomendación de diversos expertos el número de jueces incorporados al panel evaluador debe ser en número de seis a diez³⁰.

Este proceso debe iniciar con la emisión de una carta de invitación por parte del investigador dirigida a cada juez experto donde se describirá de manera estructurada todo el procedimiento para indicar la validez del contenido¹.

Los jueces conocedores del tema, son quienes evaluarán a través de una pauta diseñada por el investigador, si el contenido del instrumento (ítems) reúne las características adecuadas de claridad, pertinencia y relevancia; sin embargo es el investigador quien decidirá hasta dónde extenderá el contenido del instrumento^{6,8}. El resultado de las puntuaciones asignadas por los jueces expertos para cada ítem servirá para calcular el índice de validez de contenido (CVR) para cada ítem empleado por Lawshe (1975), producto de este cálculo se eliminarán aquellos que muestren una puntuación crítica⁸.

(7) Validez de criterio: Comparación con un Gold Standard

Si instrumento elaborado por el investigador mide lo que se pretende medir, sus resultados obtenidos

deben tener similitud con resultados del "estándar de oro" o del criterio externo². Este tipo de validez puede ser de dos tipos: validez concurrente y predictiva.

Validez concurrente: Se considera que el instrumento posee validez concurrente cuando la medición resultante del instrumento elaborado es similar al estándar de oro y por lo tanto puede reemplazar al instrumento existente. Una consideración importante en que ambos instrumentos se aplican en un mismo periodo de tiempo².

Validez predictiva: cuando la medición resultante predice algún evento futuro y la recopilación de datos o aplicación de ambos instrumentos ocurre en distintos momentos².

(8) Validez de constructo:

Este tipo de la validez busca ratificar la teoría que fundamenta la utilización de un determinado instrumento: por ello, la validez de constructo tiene que responder a la interrogante de ¿la medición obtenida con el nuevo instrumento refleja la naturaleza del "constructo" que se pretende medir? Por ejemplo, si la finalidad es medir el nivel de conocimiento referente a una determinada patología "prevención diabetes" (como constructo), el investigador debe tener en cuenta todas las dimensiones teóricas que involucran el tema de estudio, y las relaciones que se configuran entre éstas⁵. Para establecer este tipo de validez, se realizan algunas predicciones basadas en la construcción denominada prueba de hipótesis, por otro lado el análisis factorial será otra alternativa para determinar este tipo de validez².

1) Prueba de hipótesis: Incluye tres estrategias que nos ayudarán a confirmar la validez de constructo mediante: (a) Técnica de grupos conocidos; se aplica

el instrumento en diferentes grupos de individuos, para comparar los resultados obtenidos, lo esperado es que los resultados sean diferentes, debido que el instrumento válido en constructo debe detectar las diferencias en cuanto a los grupos de individuos^{2,9}. (b) Validez convergente: Aplicado cuando no existe el estándar de oro, las puntuaciones del nuevo instrumento serán comparadas con las de otro instrumento que evalúa un constructo similar, lo esperado es que ambas medidas presenten correlaciones significativas, con lo cual se comprueba que son conceptualmente equivalentes³¹. (c) Validez divergente: La comparación será con un instrumento que evalúa un constructo totalmente diferente, el resultado esperado es correlaciones bajas, lo cual indica asociación no significativa entre las variables, confirmando que ambos instrumentos miden constructos diferentes³².

2) Análisis factorial: El análisis factorial es una de las pruebas estadísticas más apropiada para corroborar este tipo de validez, mediante ella podemos identificar la cantidad de factores o constructos que existen en el instrumento, y la medida en que cada ítem está relacionado con cada factor o constructo²¹

Existen dos tipos diferentes; análisis factorial exploratorio y análisis factorial confirmatorio, este tipo de análisis servirá para definir los factores correspondiente a las variables que estén fuertemente relacionadas entre sí, los investigadores recomiendan usar únicamente el análisis factorial exploratorio^{2,33}.

En primera instancia debe usarse el análisis factorial exploratorio, esto ayudara a determinar la estructura interna de ítems y factores (dominios), seguidamente mediante el análisis factorial confirmatorio se dará validez a tal estructura factorial soportada en un marco teórico de referencia³¹.

Fase 3: Verificar la confiabilidad del instrumento

El proceso de validación de un instrumento tiene como finalidad reducir el error en el proceso de medición. Es importante tener presente que el desarrollador del instrumento tiene la responsabilidad de identificar las fuentes de error, los cuales serían perjudiciales para una correcta interpretación de la puntuación y resultados. Un instrumento será confiable si la medida obtenida es estable a través del tiempo, además si dos o más evaluadores obtienen puntuaciones similares en un mismo periodo de aplicación; lo cual se interpreta como una precisión en la medida obtenida⁹.

La confiabilidad de un instrumento se puede estimar y evaluar utilizando su consistencia interna, prueba -retest, confiabilidad entre evaluadores, confiabilidad de forma paralela y confiabilidad de mitad dividida. Un instrumento es altamente confiable si produce el mismo resultado cuando se repite nuevamente en las mismas condiciones. En base a lo señalado es necesario considerar lo siguiente²⁵:

(9) Consistencia interna

La consistencia interna permite estimar el grado en que los ítems de un instrumento se correlacionan entre ellos, si todos miden el mismo constructo y de ser así se considera que muestran una alta correlación, es decir, la escala posee un alto grado de homogeneidad^{2,5}.

Coefficiente Alfa de Cronbach: Es el método más utilizado por los investigadores para medir la consistencia interna, a través de el, se evalúa el nivel de covarianza entre los ítems de un determinado instrumento. El valor del Alfa de Cronbach (Tabla 2) oscila entre 0 y 1, siendo considerado los

valores más cercanos a 1 como aceptables, según algunos expertos, un valor mínimo aceptable es de 0,70 para indicar la coherencia interna adecuada de un instrumento, por el contrario valores por debajo de 0,7 representan una consistencia interna deficiente indicando una mala interrelación entre los elementos²⁵. Los valores del coeficiente alfa de Cronbach están altamente influenciados por cantidad de ítems del instrumento; una pequeña cantidad de ítems por dominio en un instrumento puede reducir los valores de alfa, lo cual afecta la consistencia interna. Este método es aplicado en instrumentos con opción múltiple de respuesta².

Coeficiente Kuder y Richardson (KR-20):

Es otra alternativa para estimar la consistencia de un instrumento, donde cada ítem es analizado individualmente, esta técnica se basa en la existencia de una correlación lineal entre las respuestas a los ítems. Esta prueba es comúnmente utilizada en instrumentos con opciones de respuesta dicotómicas, en cuanto a la interpretación de resultados, los valores obtenidos cercanos a 1.00 se consideran ideales⁵.

(10) Test-retest

El test-retest nos permite estimar la estabilidad temporal del instrumento, mide la similitud de los resultados en dos momentos diferentes, la principal limitación de este método, es causada por los efectos de la memoria. En este método, en la posterior administración del cuestionario, los encuestados pueden responder recordando la respuesta anterior lo cual se considera una confiabilidad artificial. Una forma de reducir este tipo de efecto de memoria es aumentar el tiempo entre la prueba y la nueva prueba, el lapso de tiempo de 10 a 14 días se considera adecuado para repetir la prueba²⁵. La

confiabilidad puede evaluarse usando el coeficiente de correlación intraclase (CCI), el cual es una de las pruebas más adecuadas para estimar la estabilidad de las variables, pues toma en cuenta los errores de medición. Otros coeficientes, como el de Pearson o Spearman, no son adecuados para este tipo de prueba de confiabilidad, porque no consideran tales errores. En cuanto a la muestra, debe considerarse al menos 50 sujetos y para la interpretación de los resultados satisfactorios los valores mínimos de 0,70.1,20².

Formas paralelas

Este método evalúa la concordancia entre dos o más instrumentos contruidos para medir el mismo constructo, los cuales deben ser aplicados en diferentes momentos, en un intervalo de tiempo corto y en los mismos individuos. Se desarrolla en forma paralela un instrumento que es equivalente al original, pero con ítems de formato diferente. Las correlaciones de las respuestas obtenidas en ambas formas permiten estimar la confiabilidad del instrumento, permitiendo asumiendo que las dos formas son intercambiables. Por lo tanto, cuanto mayor es el grado de correlación entre las dos formas paralelas, más equivalentes son las medidas obtenidas. La principal limitación es construir dos instrumentos para medir las mismas características^{5,34}.

Mitades partidas (Split-Halves)

Este método consiste en dividir el instrumento en dos mitades para determinar el coeficiente de confiabilidad entre las puntuaciones de ambas partes. Se sustenta en la lógica de que, como todos los ítems están diseñados en la misma escala y miden el mismo constructo, las puntuaciones deben estar correlacionadas. La correlación de los resultados

obtenidos de ambos instrumentos se calcula través de la fórmula de Spearman Brown³⁴.

Confiabilidad interobservador:

Este método permite estimar el acuerdo que hay entre dos o más evaluadores evalúan a los mismos individuos con el mismo instrumento. Para variables cualitativas se usa la índice kappa, que es la es la proporción o el porcentaje observado de mediciones concordantes. Sin embargo, dado que una parte de esta concordancia puede ser debida al azar, es necesario utilizar otros métodos que tengan en cuenta este hecho. El índice Kappa resume la concordancia entre dos medidas de una variable (p. ej., las obtenidas por dos observadores), cuando está en una escala cualitativa, eliminando la fracción

de la concordancia debida al azar, es decir, la que se obtendría si las dos medidas no estuvieran relacionadas¹².

El coeficiente de correlación intraclass (R) es un índice que sintetiza el grado de concordancia entre dos variables cuantitativas. Su definición matemática deriva del modelo del análisis de la varianza. La variabilidad total de las mediciones (σ^2_x) se puede descomponer, por un lado, en la debida a los verdaderos valores (σ^2_v), es decir, la de las diferencias entre sujetos, y por otro, a la variabilidad residual de los errores de medida (σ^2_e), es decir, la de las diferencias entre las medidas para cada sujeto. Cuanto menor sea esta variabilidad residual en relación con la total, mayor será la concordancia entre las mediciones¹².

CONCLUSIÓN

Este estudio representa un resumen y organización de la información que se encuentra ampliamente dispersa en la literatura respecto a la metodología sobre construcción y validación de instrumentos de medición en el ámbito de la salud, al mismo tiempo que hace una aproximación a los recursos estadísticos que más se utilizan para lograr este propósito. Es por ello, que este estudio resulta de gran utilidad para investigadores de las áreas de salud, que busquen construir, y validar los instrumentos de su interés, de modo que los resultados de sus investigaciones sean reproducibles, válidos, sensibles y útiles para la población estudiada y la comunidad científica.

Se recomienda seguir una metodología adecuada cumpliendo de manera ordenada las diferentes fases de validación de instrumentos en el ámbito de la salud, si bien es cierto el proceso de validación es un proceso sumativo; consideramos que es necesario realizar la mayor cantidad de pruebas posibles en cuanto a validez y confiabilidad respectivamente, para concluir con un instrumento óptimo y adecuado para ser aplicado.

En este artículo se detalla un proceso adecuado para la construcción y validación de instrumentos en el ámbito de la salud, el cual contribuirá para que el odontopediatra, odontólogos generales y demás profesionales de la salud puedan realizar un adecuado proceso en la elaboración de sus instrumentos de medición (cuestionarios, encuestas, etc.)

REFERENCIAS

1. Coluci MZO, Alexandre NMC, Milani D. Construção de instrumentos de medida na área da saúde. *Cienc e Saude Coletiva*. 2015;20(3):925–36.
2. Souza AC de, Alexandre NMC, Guirardello E de B. Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. *Epidemiol. e Serv. saude Rev do Sist Unico Saude do Bras*. 2017;26(3):649–59.
3. Mohajan HK. Two Criteria for Good Measurements in Research: Validity and Reliability. *Ann. Spuru Haret Univ Econ Ser*. 2017;17(4):82
4. El-Den S, Schneider C, Mirzaei A, Carter S. How to measure a latent construct: Psychometric principles for the development and validation of measurement instruments. *Int. J Pharm Pract*. 2020;28(4):326–36.
5. Echevarría-Guanilo ME, Gonçalves N, Juceli Romanoski P. Psychometric properties of measurement instruments: Conceptual basis and evaluation methods- Part II. *Texto e Context Enferm*. 2019;28:1–14.
6. Cruz-Avelar A, Sinaí Cruz-Peralta E. Metodología para la construcción de instrumentos de medición en salud. *Artículo de revisión* 2017;26:100–5.
7. Carvajal A, Centeno C, Watson R, Martínez M, Sanz Rubiales Á. ¿Cómo validar un instrumento de medida de la salud? *An Sist Sanit Navar*. 2011;34(1):63–72.
8. Taherdoost H. Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *SSRN Electron. J*. 2018;(September).
9. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. *Am J Heal Pharm*. 2008;2276–84.
10. Thomas DB, Oenning NSX, Goulart BNG de. Essential aspects in the design of data collection instruments in primary health research. *Rev CEFAC*. 2018;20(5):657–64.
11. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quifonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Heal*. 2018;6(June):1–18.
12. Argimon J, Jiménez J. (2019). *Métodos de investigación clínica y epidemiológica*. (5ta edición). Madrid: Elsevier Castellano.
13. Rattray J, Jones MC. Essential elements of questionnaire design and development. *J Clin Nurs*. 2007;16(2):234–43.
14. Clark LA, Watson D. Constructing validity: New developments in creating objective measuring instruments. *Psychol Assess*. 2019 Dec;31(12):1412-1427.
15. Mendonca K, Guerra R. Development and Validation of an Instrument for Measuring Patient Satisfaction With Physical Therapy. *Rev Bras Fisioter*. 2007;11(5):369–76.
16. Davoudi-Kiakalayeh A, Mohammadi R, Pourfathollah AA, Siery Z, Davoudi-Kiakalayeh S. Alloimmunization in thalassemia patients: New insight for healthcare. *Int J Prev Med*. 2017;8(November).
17. García Delgado P, Gastelurrutia Garralda MÁ, Baena Parejo MI, Fisac Lozano F, Martínez Martínez F. Validación de un cuestionario para medir el conocimiento de los pacientes sobre sus medicamentos. *Aten Primaria* 2009;41(12):661–8.
18. Lamprea M JA, Gómez-Restrepo C. Validity in Scale-testing. *Rev Colomb Psiquiatr*. 2007;36(2):340–8.
19. Anil S, Anand PS. Early childhood caries: Prevalence, risk factors, and prevention. *Front Pediatr*. 2017;5(July):1–7.
20. Alarcón M AM, Muñoz N S. Medición en salud: Algunas consideraciones metodológicas. *Rev Med Chil*. 2008;136(1):125–30.
21. Chen C, Feng X, Li YT, Zhang Q, Jin YS. Development and validation of a self-management ability questionnaire for patients with chronic periodontitis. *Int J Nurs. Sci*. 2019;6(3):259–65. Available at: <https://doi.org/10.1016/j.ijnss.2019.06.006>.
22. Sachdev PK, Freeland-Graves J, Babaei M. Development and validation of the Dental Nutrition Knowledge Competency Scale for low-income women. *Public Health Nutr*. 2020;23(4):691–700
23. Bichi AA, Talib R, Embong R, et al. Development and validation of behavioural and health sciences research instrument. *Indian J. Public Heal Res Dev*. 2019;10(6):1437–43.
24. Aithal A, Aithal PS. Development and Validation of Survey Questionnaire & Experimental Data – A Systematical Review-based Statistical Approach. *Int. J Manag Technol Soc Sci*. 2020;(103996):233–51.
25. Santosh P, Tarver J, Gibbons F, Vitoratou S, Simonoff E. Protocol for the development and validation of a questionnaire to Assess concerning behaviours and mental health In individuals with autism spectrum disorders: The Assessment of concerning behaviour

(ACB) scale BMJ Open 2016;6(3).

26. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth* 2017;11(5):S80–9.
27. Fuentes Pérez C. Satisfacción ambiental de la vivienda. Interacción y entorno. *Rev Iberoam las Ciencias Soc y Humanísticas RICS* 2015;4(8):8–8.
28. Escobar-Pérez J, Cuervo-Martínez Á. Validez De Contenido Y Juicio De Expertos: Una Aproximación a Su Utilización. *Av. en Medición* 2008;6(January 2008):27–36.
29. Yusoff MSB. ABC of Content Validation and Content Validity Index Calculation. *Educ Med J.* 2019;11(2):49–54.
30. Luján-Tangarife JA, Cardona-Arias JA. Construcción y validación de escalas de medición en salud: Revisión de propiedades psicométricas. *Arch Med.* 2015;11(3):1–10.
31. Polit DF. Assessing measurement in health: Beyond reliability and validity. *Int. J. Nurs Stud.* 2015;52(11):1746–53.
32. Pérez-Gil JA, Chacón Moscoso S, Moreno Rodríguez R. Validez de constructo: El uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema* 2000;12(SUPPL. 2):442–6.
33. Lagunes R. Recomendaciones sobre los procedimientos de construcción y validación de instrumentos y escalas de medición en la psicología de la salud. *Psicol y Salud* 2017;27(1):5–18.