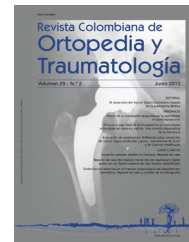




www.elsevier.es/rccot



ORIGINAL

Traducción, adaptación cultural y validación de una escala de función del miembro superior: DASH



Luis Alejandro García González^{a,*}, Francisco Javier Aguilar Sierra^b,
Constanza Moreno Serrano^c y Mónica Enciso^d

^a Cirujano de Mano y Miembro Superior, Unidad de Cirugía de Miembro Superior, Departamento de Ortopedia y Traumatología, Hospital Universitario San Ignacio, Pontificia Universidad Javeriana, Bogotá, Colombia

^b Cirujano de mano, Instituto de Ortopedia Infantil Roosevelt, Bogotá, Colombia

^c Cirujana de Mano, Unidad de Cirugía de la Mano, Departamento de Ortopedia y Traumatología, Fundación Santa Fe de Bogotá, Colombia

^d Cirujana de Hombro, Clínica Colsubsidio, Bogotá, Colombia

Recibido el 15 de julio de 2016; aceptado el 29 de junio de 2017

Disponible en Internet el 14 de septiembre de 2017

PALABRAS CLAVE

DASH;
Escala de valoración
clínica;
Validación;
Miembro superior

Resumen

Introducción: El objetivo del estudio fue realizar la traducción, adaptación cultural y validación de una escala de función del miembro superior, denominada DASH, por sus siglas en inglés (discapacidad del brazo, el hombro y la mano).

Materiales y métodos: Se realizó un estudio metodológico de validación de escalas: traducción (constó de traducciones inglés-español, comité de expertos inicial, dos contratraducciones, un segundo comité de expertos y una traducción definitiva); evaluación de confiabilidad (pruebas de coherencia interna, test-retest y cambio mínimo detectable); respuesta al cambio (mediciones en pacientes antes y después de tratamiento), y estudio de validez (se hicieron pruebas de validez de apariencia, de constructo y de criterio).

Resultados: La prueba traducida fue aplicada a 338 pacientes con enfermedades de mano y hombro, 311 fueron utilizables. La consistencia interna, mediante alfa de Cronbach, se estimó $> 0,96$; el resultado del coeficiente de concordancia de Lin fue 0,86. El cambio mínimo detectable fue 15,88, la capacidad de respuesta al cambio mediante la prueba de la *t* de Student aplicada sobre la diferencia en la puntuación arrojó una $p < 0,001$; adicionalmente se corrió una correlación no paramétrica de Spearman entre la diferencia en la puntuación DASH y la diferencia en la escala visual analógica de cambio en estado de salud (EVA SALUD) y se obtuvo un $r = 0,62$ ($p < 0,001$).

* Autor para correspondencia. Dirección de correspondencia: Carrera 16 A # 82 16. Tel.: +3187083459.

Correo electrónico: lagarcia@javeriana.edu.co (L.A. García González).

KEYWORDS

DASH;
Scales for clinical
evaluation;
Validation;
Upper limb

Discusión: La escala DASH traducida y adaptada al español demostró una buena fiabilidad, estabilidad y capacidad de respuesta al cambio, logró discriminar entre diferentes enfermedades y tiene una validez estimada buena en cuanto a contenido, constructo y criterio.

Nivel de evidencia clínica a: Nivel II.

© 2017 Sociedad Colombiana de Ortopedia y Traumatología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Translation, cultural adaptation and validation of the DASH - upper extremity functional scale

Abstract

Introduction: The objective of the study was to perform translation, cultural adaptation and validation of a function scale of the upper limb, called DASH, for its acronyms in English (disability of the arm, shoulder and hand).

Materials and methods: A methodological study of scale validation was performed: translation (consisting of English-Spanish translations, initial expert committee, two backtranslations, a second committee of experts and a definitive translation); reliability assessment (internal consistency tests, test-retest and minimum detectable change); response to change (measurements in patients before and after treatment), and validity study (validity tests of appearance, construct and criteria were performed).

Results: The translated test was applied to 338 patients with hand and shoulder diseases, 311 were usable. The internal consistency, using Cronbach's alpha, was estimated to be > 0.96 ; The result of the Lin coefficient of agreement was 0.86. The minimum detectable change was 15.88, the ability to respond to change using the Student t test applied on the difference in score yielded a $p < 0.001$; A nonparametric Spearman correlation was recorded between the difference in the DASH score and the difference in the visual analogue scale of change in health status (EVA HEALTH) with $r = 0.62$ ($p < 0.001$).

Discussion: The DASH scale, translated and adapted to Spanish demonstrated good reliability, stability and responsiveness to change, managed to discriminate between different diseases and has an estimated validity in terms of content, construct and criteria.

Level of clinical evidence: Level II.

© 2017 Sociedad Colombiana de Ortopedia y Traumatología. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

La medición de variables es la base de la estandarización de conceptos relacionados con la salud. En el caso de la percepción de las personas sobre su estado de salud, la magnitud de sus síntomas y sus discapacidades existe una dificultad importante para dicha medición. Para intentar superar esta dificultad se ha implementado la teoría psicométrica desde las ciencias psicológicas para desarrollar escalas que permitan medir de forma fiable estas variables conocidas como «blandas». El empleo de escalas para la evaluación del estado de salud ha aumentado notoriamente en todas las especialidades de la medicina durante los últimos años. En Ortopedia y, específicamente, en Cirugía de la Mano, esta tendencia no ha sido diferente. Actualmente se publican estudios clínicos que presentan como parte importante de la evaluación de los individuos o pacientes la utilización de una escala o más. La mayoría de estas son específicas de una región anatómica (codo, muñeca, etc.) o patología en particular (síndrome del túnel del carpo [STC], lesiones de tendones flexores, inestabilidad del hombro, etc.).

Además, la disponibilidad de escalas diferentes para realizar la evaluación de un mismo constructo y los criterios individuales y arbitrarios de elección imposibilitan la comparación de los resultados de diferentes estudios¹⁻¹³.

La escala de discapacidad del hombro, el brazo y la mano fue desarrollada por un equipo conjunto de expertos comisionados por el Institute for Work & Health de Canadá, la American Academy of Orthopaedic Surgeons, la American Association for Hand Surgery, la American Orthopaedic Society for Sports Medicine, el American Shoulder & Elbow Surgeons, la American Society for Surgery of the Hand, la Arthroscopy Association of North America y la American Society of Plastic and Reconstructive Surgeons, tratando unificar en un solo instrumento la evaluación de la gran mayoría o todas las patologías de la extremidad superior del adulto. Para esto se revisaron 13 escalas diferentes referentes al miembro superior, se estudiaron sus ítems y se partió de 821 distractores¹⁴⁻¹⁹.

El DASH es un cuestionario autoadministrado que permite su aplicación no concurrente (por correo) reduciendo así la posibilidad de sesgo de observador. Se han realizado estudios

para estimar la fiabilidad y la validez en diferentes regiones del miembro superior. La coherencia interna y la fiabilidad test-retest superan el valor de 0,95. El DASH se correlaciona altamente (>0,75) con otras medidas de función, discapacidad y dolor. Se ha demostrado su sensibilidad al cambio en pacientes con patologías del manguito rotador, STC, osteoartritis del hombro, etc. El proceso de validación continúa en proceso²⁰⁻²⁷.

Durante los últimos 5 años se ha publicado el resultado de la validación del DASH a los siguientes idiomas: sueco, francés, japonés, chino, coreano, alemán, armenio, griego, español europeo y otros, con resultados satisfactorios en todos los casos en cuanto a su fiabilidad y validez para medir la discapacidad del miembro superior tanto en entidades agudas como en crónicas, aisladas y generalizadas; esto demuestra su asimilación en los continentes europeo y asiático, y es indicio de su universalización progresiva. Igualmente, como se describe más adelante, el DASH aparece como método de medición en estudios clínicos de miembro superior con frecuencia creciente y se ha posicionado como herramienta estándar para el reporte de desenlaces de intervenciones en mano, codo y hombro²⁸⁻³⁸.

Materiales y métodos

Se desarrolló un estudio metodológico de validación de escalas, en el cual se incluyeron como población de referencia a personas adultas con enfermedades o estados posteriores al traumatismo o posteriores al tratamiento de la extremidad superior con los siguientes criterios de inclusión: paciente mayor de 16 años; de cualquier sexo; diagnóstico clínico y/o paraclínico de una de las siguientes patologías: síndrome del túnel carpiano, tenosinovitis estenosante de los flexores de los dedos (dedos en gatillo), tenosinovitis estenosante del primer compartimiento extensor (enfermedad de Quervain), tendinosis o rotura del manguito de los rotadores, hombro congelado, artrosis del hombro o inestabilidad del hombro; aceptación para participar en el estudio, y firma de formato de consentimiento informado.

Los criterios de exclusión fueron: rechazo a participar en el estudio, incapacidad para cumplir con controles o seguimiento y analfabetismo.

Estimación del tamaño de la muestra

Se definió el análisis factorial como la prueba con mayor requerimiento de muestra. Por tanto, esta se compuso de 10 pacientes por cada ítem que debía evaluarse (300 individuos) siguiendo lo propuesto por Norman y Streiner³⁹ y por Tinsley y Tinsley⁴⁰. Además, se incluyó un 10% más para compensar pérdida de seguimiento de individuos y otro 10% por cuestionarios inutilizables. Esto dio como resultado una muestra de 360 pacientes.

Traducción y adaptación cultural

Se hicieron 3 traducciones del inglés al español, siguiendo un patrón predeterminado. De estas, una fue hecha por una persona no médica, otra por un médico no ortopedista y la otra por un cirujano de mano. Se definió una traducción

inicial (T1) por consenso. Las anotaciones y discusiones se registraron, lo mismo que la forma en que se llegó a un acuerdo. La T1 fue retrotraducida al inglés por dos traductores de lengua nativa inglesa (T2) y la versión final se logró después de un comité de expertos.

Análisis de los resultados

La información de los cuestionarios diligenciados fue almacenada en una base de datos en Excel X para Macintosh, se crearon subbases para el manejo de las diferentes pruebas y estas fueron procesadas en Stata 10 para Windows.

Se corrieron las pruebas que se enumeran a continuación.

Demografía

Estadística descriptiva sobre variables demográficas básicas.

Estimación de fiabilidad test-retest

Se aplicó el instrumento a un subgrupo de pacientes de cada grupo una semana después de la aplicación inicial, en quienes se cumplía que no hubiesen iniciado tratamiento y/o declarasen ausencia de cambio en sus síntomas y función de acuerdo con el formato correspondiente. Se midió concordancia entre los dos valores obtenidos para cada individuo mediante un coeficiente de correlación de concordancia de Lin^{41,42}. La muestra para esta prueba fue definida en 30 individuos para obtener un intervalo de confianza del 95% (IC95%) con amplitud menor de 0,2 con 2 pruebas por individuo y extrapolando los cálculos para obtener un CCI de 0,9 de acuerdo con lo propuesto por Giraudeau y Mary⁴³. Para compensar pérdidas por incumplimiento de pacientes, se definió un total de 35 individuos para esta prueba.

Estimación de la precisión de la prueba

Se pretendió estimar el error alrededor de una puntuación obtenida al aplicar el DASH, tanto transversal (hay un rango de valores entre los cuales se encuentra la verdadera puntuación) como longitudinal (mediciones seriadas muestran cambios en la puntuación que pueden representar cambio real en función o variación por error).

Siguiendo el método propuesto por Ravaud et al., y seguido por Beaton et al.⁴⁴, la precisión transversal parte del error estándar de la medición (EEM). Este se calculó a partir del coeficiente de fiabilidad de Cronbach (R) con la siguiente fórmula:

$EEM_t = DE \times \sqrt{1 - R}$, donde DE corresponde a la desviación estándar basal de la muestra y R al coeficiente de fiabilidad. Para determinar el nivel de confianza con qué definir el rango de valores donde se debe encontrar la puntuación verdadera, se multiplicó el EEM por el valor de Z deseado, en este caso, para un intervalo de confianza de 95%, se multiplicó por 1,96.

La precisión longitudinal del DASH se estimó mediante el cálculo del cambio mínimo detectable (CMD), esto es, el valor mínimo aceptable de cambio para considerar que el cambio obtenido en dos tomas diferentes de la prueba

obedece a un cambio en la función o los síntomas y no a la variabilidad por error.

Para estimar la precisión longitudinal, el EEM se calcula a partir del coeficiente de confiabilidad test-retest (r) y la DE basal con la siguiente fórmula:

$$EEM_1 = DE \times \sqrt{1 - r}$$

El cambio mínimo detectable se calcula con el 95% de fiabilidad al multiplicar el EEM por 1,96 (valor de Z para un intervalo de confianza del 95%).

Estimación de la respuesta al cambio

La capacidad de identificar un cambio en el estado de salud es una característica importante de una escala. En este estudio se planeaba estimar esta sensibilidad al cambio comparando las puntuaciones obtenidas en pacientes a quienes se les aplicó la prueba dos veces en dos situaciones distintas, pero que debían sobreponerse por sus características. Primero, se esperaba obtener un efecto clínicamente importante de intervenciones terapéuticas sobre los individuos del estudio^{45,46}, por ejemplo, liberación del túnel del carpo, infiltración de tenosinovitis de mano, cirugía de reparación del manguito de los rotadores, fisioterapia, etc. Para medir este efecto, se solicitó a los participantes del estudio rellenar el cuestionario entre 6 y 12 semanas después de efectuada la intervención o iniciado el tratamiento conservador si bien este no es el máximo efecto esperado, especialmente en problemas del hombro, pues es clínicamente plausible esperar un cambio apreciable. En segundo lugar, durante el control clínico los pacientes rellenaron un formato de escala visual con la pregunta: «en comparación con antes de iniciar el tratamiento, ¿qué tanto ha cambiado su brazo, hombro o mano?», con una escala visual que pretendía identificar la presencia de cambio, su dirección y magnitud.

Los resultados para la primera medición se evaluaron por diferencia en puntuación (el valor absoluto de la resta de la puntuación inicial menos la final), tamaño del efecto (la media del cambio dividida por la DE de las puntuaciones de base) y la respuesta media estandarizada (RME, medida como el resultado de dividir el promedio de cambio sobre la DE del cambio en las puntuaciones)⁴⁷⁻⁵⁰ y se estimó la significancia estadística de las diferencias mediante una prueba de la t pareada^{51,52}. Para la segunda medición, adicional a lo anterior, se comparó la magnitud y dirección del cambio reportado en el formato de escala visual con la magnitud y dirección de cambio en la puntuación del DASH mediante el cálculo del coeficiente de correlación de Spearman; para estas mediciones se estimó un tamaño de muestra de 43 pacientes con un error alfa de 0,05, poder del 90% para encontrar diferencias entre puntuación basal de 40 y posterior al tratamiento de 25 con DE de 15 para los dos grupos.

Estimación de consistencia interna

Se buscó evaluar la homogeneidad de los ítems que componen la escala al comparar la puntuación de cada ítem con la escala entera y consigo mismo. Lo ideal es tener un grado de

moderado a alto de homogeneidad y se midió mediante la prueba de la alfa de Cronbach como recomiendan Streiner y Norman⁵³, aplicada a la totalidad de la muestra y a los diferentes segmentos de esta, es decir, a cada grupo etiológico, previamente y después del tratamiento. Adicionalmente se corrió un análisis de factores confirmatorio también en la totalidad de la muestra para evaluar peso de los factores y corroborar la distribución de los dominios descritos por los autores de la escala.

Estimación de la validez

Se planeó exponer el cuestionario ante un grupo de expertos en el área de Cirugía de la mano, de Hombro y en Medición, solicitando un reporte escrito sobre su opinión acerca de la apariencia del cuestionario, si este era capaz de evaluar discapacidad del miembro superior, si la cantidad, organización y contenido de las preguntas era suficiente y apropiado para evaluar problemas en toda la extremidad y si era lo suficientemente específico para evaluar problemas específicos. Con esta encuesta se esperaba estimar la evidencia de validez basada en apariencia y contenido. La evidencia de validez basada en el constructo o estructura interna se exploró mediante análisis de factores, pero de acuerdo con autores como Goodwin y Nunnally^{54,55}, este método es insuficiente para relacionar la escala con eventos externos a esta; así, se practicó comparación de grupos, de la forma siguiente:

1. Validez de constructo por severidad:

- a) Comparación de puntuaciones de pacientes con problemas de la mano y del hombro. Se esperaba obtener puntuaciones mayores en los segundos y, por tanto, discriminar entre gravedad de enfermedades y niveles de compromiso funcional.
- b) Comparación de puntuaciones de acuerdo con la situación laboral del individuo: empleado en ocupación habitual, empleado con restricciones debidas a su problema de hombro, brazo o mano, o incapacitado por su problema de hombro, brazo o mano.

2. Validez de constructo longitudinal:

- a) Comparación de puntuaciones iniciales con postratamiento (12 semanas). Se superpone con la respuesta al cambio, pero es útil también para determinar discriminación entre mayor y menor compromiso funcional y clínico.
- b) Correlación de puntuaciones con escala visual análoga de dolor. Permite también estimar su capacidad de discriminación por severidad.

Validez de criterio

La estimación de la validez basada en criterio concurrente se abordó mediante la comparación entre las puntuaciones del DASH y la versión en español del cuestionario SF - 12aV2 validada en Colombia y avalada por Qualitymetrics.

Las comparaciones anteriormente mencionadas se exploraron estadísticamente mediante pruebas de correlación no paramétrica de Spearman.

Resultados

Traducción y adaptación cultural

Se produjeron 3 traducciones de inglés a español siguiendo un formato específico, posteriormente se reunió el comité de traducción y, por consenso, se definió una versión definitiva T1.

Se hizo una contratraducción al inglés por parte de 2 profesores del departamento de Lenguas de la Universidad cuya lengua materna es el inglés.

Finalmente se reunió el comité de revisión incluyendo a los autores, a los traductores de español a inglés y a los contratraductores, y se obtuvo la versión definitiva después de verificar la adherencia al original de la versión contratraducida y la equivalencia clínica, semántica y cultural de las escalas.

Fase 2: aplicación del DASH

Entre julio de 2005 y marzo de 2007 se aplicó la versión a estudio del DASH a 338 pacientes, quienes rellenaron 465 cuestionarios. 24 cuestionarios fueron anulados por tener más de 3 ítems sin respuesta. La tasa de respuesta para cada ítem fue de 0,97 por término medio (de 0,83 a 0,99), con una mediana de 0,98. El único ítem con tasa menor de 0,95 fue la pregunta 21 (actividades íntimas de pareja con tasa de 0,83).

La base analizada es de 311 pacientes y la información demográfica se encuentra en la [tabla 1](#).

Estimación de fiabilidad test-retest

El resultado del coeficiente de concordancia de Lin fue 0,86; IC95%: de 0,76 a 0,95 en una muestra de 38 pacientes.

Estimación de la precisión transversal

Esta se estimó mediante la medición del EEM con un intervalo de confianza del 95%, basada en el coeficiente alfa de confiabilidad (Rxx). Para esto se calculó el EEM y se multiplicó por 1,96

$$EEM_{95} = 1,96 \times EEM$$

$$EEM_{95} = 1,96 \times DE_{base} \times \sqrt{1 - R_{xx}}$$

$$EEM_{95} = 1,96 \times 21,65 \times \sqrt{1 - 0,96}$$

$$EEM_{95} = 1,96 \times 4,33$$

$$EEM_{95} = 8,49$$

Este resultado significa que, para una puntuación determinada, en un paciente y momento cualquiera, el valor real

Tabla 1 Resumen de información demográfica de la muestra

Total	311
<i>Media de edad (DE)</i>	50 ($\pm 12,6$)
<i>Mínima</i>	17
<i>Máxima</i>	84
<i>Hombres</i>	91
<i>Mujeres</i>	220
<i>Estatura media</i>	162 cm
<i>Hombres</i>	170
<i>Mujeres</i>	158
<i>Peso</i>	66 kg
<i>Hombres</i>	74,3
<i>Mujeres</i>	63
<i>IMC medio</i>	25
<i>Mínimo</i>	17
<i>Máximo</i>	38
<i>Ocupación</i>	
Manual	136
No manual	143
Sin información	32
Trabajo usual	163
Desempleado	6
Ama de casa	85
Estudiante	10
Retirado	19
Incapacidad	21
Otro	8
<i>Esquema de salud</i>	
POS	171
ARP	9
Prepago	121
SOAT	1
Particular	9
<i>Ubicación de la patología</i>	
Hombro	124
Mano	178
Ambos	9

del DASH se debe encontrar 8,49 puntos por encima o por debajo del valor puntual con una seguridad del 95%.

Estimación de la precisión longitudinal

Para esta, se estima el cambio mínimo detectable, con una fórmula similar a la anterior, pero en lugar de incluir el coeficiente alfa de Cronbach, se utilizó el coeficiente de fiabilidad de Lin (0,86), con la fórmula anteriormente descrita, cambiando los coeficientes, el valor del CMD_{95} fue 15,88. Este valor indica la mínima variación que se puede considerar real y no producto de cambios fortuitos en dos medidas repetidas en un mismo individuo.

Estimación de la respuesta al cambio

Esta estimación se llevó a cabo en 66 individuos de la muestra total, quienes tuvieron una puntuación basal pro-

medio de 52 (DE: 20,74) y una puntuación al control de 31 (DE: 21,21), la diferencia absoluta en puntuación fue de 27 con DE: 16,74; así, la respuesta media estandarizada fue de $20,66/23,97=0,86$ y el tamaño del efecto fue de $20,66/20,74=0,99$. De acuerdo con lo propuesto por Cohen, este es un tamaño de efecto grande ($>0,80$).

Se corrió una prueba de la t pareada comparando las puntuaciones basales con las de control y se encontró un valor de $p < 0,001$ y se confirmó la diferencia real entre las puntuaciones previa y posterior al tratamiento.

Se midió el coeficiente de correlación de Spearman entre la diferencia entre los dos valores de DASH y la diferencia en 2 valores de la EVA de dolor y se obtuvo un $r=0,62$ ($p < 0,001$) y entre DASH y EVA de cambio en estado de salud con $r=0,5668$ ($p < 0,001$), indicando una asociación moderada entre los cambios en la puntuación del DASH y las escalas visuales de dolor y estado de salud.

Estimación de la coherencia interna

El resultado del alfa de Cronbach fue estimado con resultado $> 0,96$ para los 30 ítems de la escala.

Estimación de la validez de constructo

Se corrió un análisis de factores confirmatorio de componentes principales sobre las puntuaciones del DASH de los 322 individuos participantes, encontrando que un componente explicaba el 49,42% de la varianza, con valor propio de 14,83 y se retuvieron 4 componentes que en conjunto explican el 64,17% de la varianza.

En la figura 1 se presenta el gráfico Scree que muestra el cambio en la dirección de la curva después del cuarto componente.

Se hizo una rotación varimax a los 4 componentes retenidos y se encontró la siguiente distribución de cargas para cada componente:

Componente 1: movilidad de hombro y fuerza: explica los siguientes ítems de la escala:

5. Empujar una puerta pesada.
6. Poner un objeto en una repisa ubicada por encima de su cabeza.

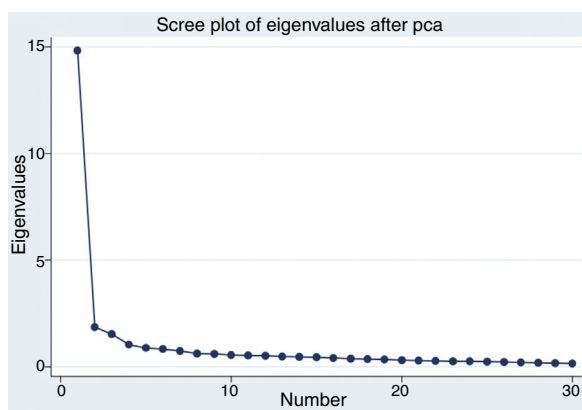


Figura 1 Gráfico Scree mostrando el cambio en la dirección de la curva después del cuarto componente.

7. Realizar oficios caseros pesados (como lavar pisos o paredes).

8. Podar o arreglar un jardín o las plantas de su casa.

9. Tender la cama.

12. Cambiar una bombilla alta (ubicada por encima del nivel de su cabeza).

13. Lavarse o secarse el pelo.

14. Lavarse la espalda.

15. Ponerse un buzo o saco cerrado.

18. Realizar actividades recreativas que requieren esfuerzo utilizando su brazo, hombro o mano (p. ej., jugar tenis, bolos, tejo, etc.).

19. Realizar actividades recreativas que requieren mover libremente su brazo (p. ej., tenis de mesa, natación, volar cometas, etc.).

20. Usar medios de transporte para ir de un lugar a otro.

21. Realizar actividades íntimas de pareja.

22. ¿Qué tanta dificultad ha tenido en la última semana para participar en actividades sociales normales con su familia, amigos o vecinos por el problema en su brazo, hombro o mano?

Componente 2: actividades manuales

1. Abrir un frasco nuevo o apretado.

2. Escribir.

3. Girar una llave.

4. Preparar una comida.

10. Cargar una bolsa o un maletín.

16. Cortar comida con un cuchillo.

17. Realizar actividades de recreación que requieren poco esfuerzo utilizando el brazo, el hombro o la mano (p. ej., jugar cartas, tejer, etc.).

Componente 3: síntomas y autoimagen

23. ¿Durante la última semana se vio limitado en sus actividades diarias como resultado del problema en su brazo, hombro o mano?

24. Dolor de brazo, hombro o mano.

25. Dolor en el brazo, hombro o mano cuando usted realiza alguna actividad específica.

26. Hormigueo o sensación de agujas en su brazo, hombro o mano.

27. Debilidad en su brazo, hombro o mano.

28. Rigidez en su brazo, hombro o mano.

29. ¿Qué tanta dificultad tuvo en la última semana para dormir por el dolor en su brazo, hombro o mano? 30. Me siento incapaz, menos seguro o menos útil debido a mi problema del brazo, hombro o mano

Componente 4: solamente explica el ítem 11 referente a carga de un objeto pesado (mayor de 5 kg)

Esta distribución de cargas denota un predominio de lo físico en la escala, con la posibilidad de diferenciar los ítems relacionados con el hombro y la fuerza de las actividades manuales y de los síntomas.

Validez de constructo por severidad

a) Se comparó el valor medio del DASH en pacientes con problemas en hombro y en mano. El resultado se encuentra en la tabla 2.

Tabla 2 Comparación de valor de DASH por ubicación anatómica

Localización	n	DASH promedio	DE	EVA dolor promedio
Mano	171	48	22,05	50
Hombro	121	49	21,2	64
Prueba de la t		p=0,6		p=0,01

Tabla 3 Resultado de ANOVA comparando STC, manguito de los rotadores y dedos en gatillo

	DASH		valor de p		
	Media	DE	STC-manguito	Manguito-gatillo	STC-gatillo
STC	49	20,73			
Manguito	54	21,60	0,719	0,004*	0,047*
Gatillo	37	22,14			

* Valor de p significativo.

- b) Se comparó el valor de DASH en pacientes con STC unilateral, con lesión del manguito de los rotadores unilateral y con dedos en gatillo mediante análisis de varianza. Los resultados se encuentran en la [tabla 3](#).
- c) Se comparó el valor medio de DASH entre pacientes en trabajo usual (n = 163) y pacientes incapacitados (n = 21) mediante una prueba de la t con resultado de p = 0,047.

Estimación de la validez de criterio

Se estimó mediante la comparación entre los valores de DASH y de SF 12 A v2 en una submuestra aleatoriamente obtenida de 100 pacientes. De estos se recuperaron 81 cuestionarios válidos. El valor promedio de la esfera física (PCS) fue de 39 ($\pm 8,7$) y la esfera mental fue de 47 ($\pm 9,46$). Se corrió una correlación no paramétrica de Spearman entre DASH y PCS = -0,5676 ($p < 0,001$) y la correlación entre DASH y MCS = -0,3787 ($p < 0,001$). Para asegurar la capacidad de generalización de los resultados de la submuestra se compararon los valores de DASH y DE de esta con la totalidad y se obtuvo una p de 0,73 en una prueba de la t pareada.

Discusión

La inclusión de medidas psicométricas para la evaluación del estado de salud ha abierto las posibilidades de medición de discapacidad, dolor y otras expresiones de estados mórbidos que aquejan a las personas; específicamente, las escalas que exploran la vivencia del paciente con su enfermedad o situación médica nos han acercado más a entender el impacto de estas situaciones y de los tratamientos en cada paciente y apreciar su punto de vista. En ortopedia y traumatología ha sido tradicional evaluar los resultados por medidas denominadas fuertes, como la consolidación de una fractura o no, la alineación de una extremidad, la fuerza y los arcos de movimiento de una articulación. Si bien estas medidas son muy importantes y su cercanía a la normalidad nos indican moderadamente un buen resultado, estas

no incluyen la percepción del paciente y lo fragmentan en datos que no implican su satisfacción y capacidad de retornar a sus actividades premórbidas, por ejemplo, después del tratamiento de una fractura del antebrazo, con adecuada consolidación de los huesos, pero con pérdida de 45° de supinación. Un ama de casa puede expresar satisfacción completa, pero un tenista competitivo no podrá efectuar golpes básicos y no podrá regresar a su deporte. Este paciente seguramente no estará satisfecho. Las medidas tradicionales nos dirán que ambos resultados son satisfactorios. Si a estos dos pacientes les aplicamos un instrumento de medición de discapacidad, sus resultados serán muy diferentes. Este ejemplo busca explicar la inmensa utilidad de introducir instrumentos como las escalas de medición de discapacidad en medicina y específicamente en ortopedia. Las escalas que incluyen la subjetividad del paciente se están introduciendo en la práctica clínica ortopédica desde hace más de dos décadas. Algunas de ellas son ya universales y han permitido la comparación de resultados entre poblaciones diferentes mediante su validación. Desafortunadamente, la primera generación de escalas en ortopedia fue desarrollada para áreas anatómicas específicas y aun para procedimientos quirúrgicos específicos, lo que limitó su aplicación.

La iniciativa de los autores del DASH se basó en la necesidad de disponer de una escala que fuese aplicable en todas las situaciones patológicas y traumáticas de toda la extremidad superior, desde la mano hasta el hombro, y que permitiese dar a los tratantes un instrumento universal de medición de discapacidad originada en el miembro superior con suficiente validez y capacidad discriminativa. La vocación universal de esta escala se manifiesta también en el hecho de que ha sido abierta a la disposición de todos los profesionales para su utilización en el idioma original y para su validación en otros idiomas.

Seguir este impulso ha sido la motivación de desarrollar este estudio de traducción, adaptación cultural y validación del DASH. Se logró hacer la traducción siguiendo una metodología específica y ordenada, obteniendo un instrumento similar al original y comprensible para la población estudiada. Esto se manifiesta por la alta proporción de escalas utilizables en este estudio (95% del total de cuestionarios aplicados). Igualmente, solo el ítem 21, referente a actividades íntimas de pareja, tuvo una tasa de respuesta menor al 93%, con el 83% de respuestas. La inclusión de este ítem puede haber generado una pérdida menor de exactitud de los datos, pero esta no se comprobó al comparar las puntuaciones incluyendo y excluyéndolo. La distribución demográfica de los participantes en el estudio refleja adecuadamente la de la población estudiada, es decir, la proporción de mujeres (71%) es la que se ve en las consultas fuente de pacientes. Esto se explica por la mayor prevalencia en mujeres de las entidades incluidas. En el caso del STC, la proporción de personas que consultan es 4 mujeres por cada hombre. Esta proporción es similar en el caso de los dedos en gatillo y para el hombro doloroso está alrededor de 2:1.

La estabilidad de la escala se confirmó satisfactoriamente al encontrar un coeficiente de concordancia de 0,86 en la prueba test-retest. Este valor está por debajo del reportado inicialmente de 0,96, pero supera el límite inferior de 0,75 previamente propuesto por Nunnally.

La precisión de la escala busca estimar la magnitud de error de una puntuación obtenida en un momento determinado. La precisión transversal obtenida a partir del error estándar de la medición informa que, con un grado de confianza específico, en este caso el 95%, el valor real de esta puntuación estará 8,4 puntos por encima o por debajo del valor obtenido. Este resultado es casi idéntico al reportado en la versión original (8,8)⁵⁶.

La respuesta al cambio fue estimada primariamente mediante el cálculo de la diferencia absoluta en puntuación de dos aplicaciones del DASH. El punto está bien por encima del cambio mínimo detectable, lo que indica la validez de este valor. Adicionalmente se calculó el tamaño del efecto y la respuesta media estandarizada (0,99 y 0,86, respectivamente). Estos dos valores relacionan el cambio promedio en las puntuaciones con la desviación estándar de la muestra y pretenden dar una cifra única que refleje la magnitud de este cambio, como se mencionó antes, desde el punto de vista de Cohen. Un valor mayor de 0,80 se puede considerar un efecto grande; en este caso indican que las diferencias encontradas entre las dos aplicaciones son producto de un cambio real y no del azar. Estos valores superan los reportados por Beaton et al. en su reporte de validación original. Se intentó confirmar la tendencia y dirección del cambio en la puntuación del DASH con las escalas visuales de dolor y cambio en estado de salud mediante el cálculo de correlaciones de Spearman, pero se obtuvo una asociación moderada. Esta se puede intentar explicar por la dificultad encontrada para la comprensión por parte de los participantes de estas escalas visuales análogas que introdujo mayor variabilidad en sus resultados⁵⁷.

La coherencia interna hallada es prácticamente igual a la reportada por todos los estudios de validación del DASH. Este elevado valor de 0,96, que sugiere una redundancia de ítem, justificó la modificación del DASH original por sus autores y redujo el número de ítems a 11, denominado quickDASH, ya disponible para aplicación con unas características similares al original.

La validez fue estimada desde varios ángulos tratando de cubrir de la mejor forma esta difícil parte del proceso emprendido. La validez de constructo se abordó mediante el desarrollo de un análisis factorial de componentes principales. El resultado es muy similar al reportado por estudios de validación del DASH en japonés, chino, holandés y coreano, y al análisis exploratorio reportado por los autores de la escala. También se exploró la validez de constructo por gravedad, comparando las puntuaciones de pacientes con problemas en hombro y en mano. Hay que resaltar que una de las hipótesis de este trabajo, el hecho de que las puntuaciones del DASH serían mayores en lesiones más proximales, fue desvirtuada, pues el valor promedio de los pacientes con problemas del hombro fue 49 y de los pacientes afectados en la mano fue 48 ($p=0,6$). Se encontró diferencia al comparar los valores de la escala visual de dolor entre los grupos anatómicos ya que los pacientes con problemas del hombro tuvieron una puntuación promedio de 64 y los afectados en la mano reportaron una puntuación de 50 ($p=0,01$). Esto reflejó una pobre capacidad de discriminación de la escala por gravedad si hemos de confiar en el enunciado de que las patologías del hombro afectan más a la capacidad funcional que las de la mano. Desafortunadamente no contamos con un instrumento fiable paralelo para evaluar la gravedad

de cada caso que nos permitiera una comparación mejor; otros estudios publicados cuentan con otras escalas validadas previamente para comparar, pero no era nuestro caso; intentando solucionar este problema, se hizo una comparación tomando solo los pacientes con diagnóstico definido de STC, dedo en gatillo y lesión del manguito de los rotadores, y comparando sus puntuaciones de DASH, esto con el objeto de limitar mejor el espectro de cada enfermedad. Así se encontró diferencia estadísticamente significativa entre las puntuaciones de pacientes con dedo en gatillo y los otros dos grupos (con STC $p=0,047$ y con lesión del manguito $p=0,004$). En esta situación existe la explicación clínicamente plausible de que el compromiso que produce un dedo en gatillo es menor que el producido por las otras dos entidades y permite suponer que la gravedad del STC y las lesiones del manguito de los rotadores son similares en representación subjetiva para los pacientes. Este debe ser objeto de mayor investigación basada en los pacientes. La validez de constructo también fue investigada comparando las puntuaciones de pacientes incapacitados frente a las de los pacientes laboralmente activos y se confirmó una diferencia significativa con $p=0,047$.

La validez de criterio fue estimada mediante la comparación por correlación entre las puntuaciones del DASH y las de la escala SF 12 A v2 validada para Colombia. Se escogió la versión de 12 ítems para acortar el tiempo de administración del cuestionario y la versión aguda para estar de acuerdo con el marco temporal del DASH que pregunta por síntomas y función durante la semana previa a la administración. Esta estimación mostró una correlación aceptable con la esfera física y pobre con la mental, demostrando que el DASH aplicado mide síntomas y discapacidad de origen físico y que otros componentes de la salud no intervienen en la puntuación.

Las limitaciones que presentó el desarrollo de este estudio y que pueden haber afectado a sus resultados consisten en la dificultad para obtener los cuestionarios de control, fuesen inmediatos (para medir estabilidad) o tardíos (para medir respuesta) en un número importante de participantes. Esta falla disminuyó el número de pacientes para hacer estas mediciones, pero permitió tener los mínimos necesarios para desarrollar los cálculos planeados.

La escala DASH fue traducida siguiendo parámetros estrictos de calidad y se obtuvo una versión en español fiable, comprensible y de características comparables a la versión original. En este estudio se confirmó que existe redundancia de ítem y que la escala puede ser reducida sin perder sus características estadísticas necesarias. La escala demostró poseer una adecuada estabilidad y capacidad de respuesta al cambio en pacientes con las enfermedades incluidas en este estudio. La capacidad de discriminación de esta escala fue confirmada al comparar una enfermedad de curso benigno y generalmente corto como el dedo en gatillo con enfermedades crónicas como el STC y las lesiones del manguito de los rotadores del hombro, la incapacidad de discriminar entre el STC y la lesión del manguito de los rotadores puede responder más a limitaciones en la definición de la gravedad que a las características de la escala, pero debe ser investigado a profundidad en futuros estudios.

Se dispone ahora de un instrumento válido y fiable para medir estado funcional y respuesta al tratamiento en pacientes con enfermedades que afectan a los miembros

superiores. La labor de validación en otras patologías, especialmente en traumatología, debe continuar con la aplicación clínica y generalizada de la escala en consultas especializadas de Cirugía de la Mano y de Hombro y Codo. Adicionalmente se debe probar su aplicabilidad y validez en otras regiones del país.

Responsabilidades éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad de los datos. Los autores declaran que han seguido los protocolos de su centro de trabajo sobre la publicación de datos de pacientes.

Derecho a la privacidad y consentimiento informado. Los autores declaran que en este artículo no aparecen datos de pacientes.

Financiación

Este estudio no recibió financiación por ninguna entidad. Fue sufragado en su totalidad por los autores.

Conflicto de intereses

Ninguno de los autores tiene conflictos de interés para declarar.

Bibliografía

1. Institute for Work & Health. Toronto, ON. Canadá. <http://www.dash.iwh.on.ca/index.htm>
2. Lumpkin JR; Subcommittee on Populations, National Committee on vital and health Statistics (NCVHS): Classifying and reporting functional status. 2001. <http://www.ncvhs.hhs.gov/010716rp.htm>.
3. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability & responsiveness of the disabilities of the arm, shoulder, and hand outcome measure in different regions of the upper extremity. *J Hand Ther.* 2001;14:128–46.
4. Sánchez R, Gómez C. Conceptos básicos sobre validación de escalas. *Rev Col Psiquiatría.* 1998;27:121–30.
5. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res.* 1987;160–4.
6. Richards RR, An K-N, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg.* 1994;3:347–52.
7. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. *Arthritic Care Res.* 1991;4:143–9.
8. Lippitt SB, Harryman DT, Matsen FA. A practical tool for evaluation of function: the Simple Shoulder Test. En: Matsen FA, Fu FH, Hawkins RJ, editores. *The shoulder: a balance of mobility and stability.* Rosemont, Illinois: The American Academy of Orthopaedic Surgeons; 1993. p. 501–18.
9. Beaton DE, Richards RR. Measuring function of the shoulder: a cross-sectional comparison of five questionnaires. *J Bone Joint Surg.* 1996;78-A:882–90.
10. Gerber C. Integrated scoring systems for the functional assessment of the shoulder. En: Matsen FA, Fu FH, Hawkins RJ,

editores. *The shoulder: a balance of mobility and stability.* Rosemont, Illinois: The American Academy of Orthopaedic Surgeons; 1993. p. 531.

11. Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability. The Western Ontario Shoulder Instability Index (WOSI). *Am J Sports Med.* 1998;26:764–72.
12. L'Insalata J, Warren R, Cohen S, Altchek D, Peterson M. A self-administered questionnaire for assessment of symptoms and function of the shoulder. *J Bone Joint Surg.* 1997;79A:738–48.
13. Morrey BF, An KN, Chao EYS. Functional evaluation of the elbow. En: Morrey BF, editor. *The elbow and its disorders.* 2.^a edición Philadelphia: W.B. Saunders; 1993. p. 86–97.
14. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, et al. The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. *J Bone Joint Surg Br.* 2008;90:466–73.
15. Hildebrand K, Patterson S, Regan W, MacDermid J, King G. Functional outcome of semiconstrained total elbow arthroplasty. *J Bone Joint Surg.* 2000;82A:1379–86.
16. Ewald FC, Scheinberg RD, Poss R, Thomas WH, Scott RD, Sledge CB. Capitulocondylar total elbow arthroplasty. Two to five-year follow-up in rheumatoid arthritis. *J Bone Joint Surg.* 1980;62-A:1259–63.
17. Turchin DC, Beaton DE, Richards RR. Validity of observer-based aggregate scoring systems as descriptors of elbow pain, function, and disability. *J Bone Joint Surg.* 1998;80-A:154–62.
18. MacDermid JC. Development of a scale for patient rating of wrist pain and disability. *J Hand Ther.* 1996;9:178–83.
19. MacDermid JC, Turgeon T, Richards RS, Beadle M, Roth JH. Patient rating of wrist pain and disability: a reliable and valid measurement tool. *J Orthop Trauma.* 1998;12:577–86.
20. MacDermid JC, Richards RS, Donner A, Bellamy N, Roth JH. Responsiveness of the Short Form-36, Disability of the Arm, Shoulder and Hand Questionnaire, patient-rated wrist evaluation and physical impairment measurements in evaluating recovery after a distal radius fracture. *J Hand Surg.* 2000;25A:330–40.
21. Amadio PC, Silverstein MD, Ilstrup DM, Schleck CD, Jensen LM. Outcome assessment for carpal tunnel surgery: the relative responsiveness of generic, arthritis-specific, disease-specific, and physical examination measures. *J Hand Surg.* 1996;21A:338–46.
22. Levine DW, Simmons BP, Koris MJ, Daltroy LH, Hohl GG, Fossel AH, et al. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. *J Bone Joint Surg.* 1993;75A:1585–92.
23. Moriatis-Wolf J, Green A. Influence of comorbidity on self-assessment instrument scores of patients with idiopathic adhesive capsulitis. *J Bone Joint Surg Am.* 2002;84:1167–73.
24. Hudak PL, Amadio PC, Bombardier C, the Upper Extremity Collaborative Group (UECG). Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. *Am J Ind Med.* 1996;29:602–8.
25. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the disabilities of the arm, shoulder and hand outcome measure in different regions of the upper extremity. *J Hand Ther.* 2001;14:128–46.
26. Atroshi I, Gummesson C, Andersson B, Dahlgren E, Johansson A. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire. Reliability and validity of the Swedish version evaluated in 176 patients. *Acta Orthop Scand.* 2000;71:613–8.
27. Veehof MM, Slegers EJ, van Veldhoven NH, Schuurman AH, van Meeteren NL. Psychometric qualities of the Dutch language version of the disabilities of the arm, shoulder, and hand questionnaire (DASH-DLV). *J Hand Ther.* 2002;15:347–54.
28. Imaeda T, Toh S, Nakao Y, Nishida J, Hirata H, Ijichi M. Validation of the Japanese Society for Surgery of the Hand version of the Disability of the Arm, Shoulder, and Hand questionnaire. *J Orthop Sci.* 2005;10:353–9.
29. Dubert T, Voche P, Dumontier C, Dihn A. Le questionnaire DASH. Adaptation française d'un outil d'évaluation international. *Chirurgie de la Main.* 2001;20:294–302.
30. Lee EWC, Chung MMH, Li APS, Lo SK. Construct validity of the chinese version of the Disabilities of the arm, shoulder and hand Questionnaire (dash-hkpwh). *J Hand Surg. (British and European Volume.)* 2005;30B:29–34.
31. Joo-Yup Lee, Jae-Young Lim, Joo Han Oh, Young-Mi Ko. Cross-cultural adaptation and clinical evaluation of a Korean version of the disabilities of arm, shoulder, and hand outcome questionnaire (K-DASH). *J Shoulder Elbow Surg.* 2008;17:570–4.
32. Germann G, Harth A, Wind G, Demir E. Standardisation and validation of the German version 2.0 of the Disability of Arm, Shoulder, Hand (DASH) questionnaire. *Unfallchirurg.* 2003;106:13–9.
33. Hunsaker FG, Cioffi DA, Amadio PC, Wright JG, Caughlin B. The American Academy of Orthopaedic Surgeons Outcomes Instruments – Normative Values from the General Population. *J Bone Joint Surg.* 2002;84-A:208–15.
34. Gay RE, Amadio PC, Johnson JC. Comparative responsiveness of the Disabilities of the Arm, Shoulder, and Hand, the Carpal Tunnel Questionnaire and the SF-36 to clinical change after carpal tunnel release. *J Hand Surg. (American).* 2003;28A:250–4.
35. McKee MD, Kim J, Kebaish K, Stephen DJG, Kreder HJ, Schemitsch EH. Functional outcome after open supracondylar fractures of the humerus. *J Bone Joint Surg.* 2000;82-B:646–51.
36. Ring D, Perey BH, Jupiter JB. The functional outcome of operative treatment of ununited fractures of the humeral diaphysis in older patients. *J Bone Joint Surg.* 1999;81-A:177–90.
37. Su BW, Solomons M, Barrow A, Senoge ME, Gilberti M, Lubbers L, et al. Device for zone-II flexor tendon repair. A multicenter, randomized, blinded, clinical trial. *J Bone Joint Surg Am.* 2005;87:923–35.
38. Ruch DS, Ginn TA, Yang CC, Smith BP, Rushing J, Hanel DP. Use of a distraction plate for distal radial fractures with metaphyseal and diaphyseal comminution. *J Bone Joint Surg Am.* 2005;87:945–54.
39. Norman GR, Streiner DL. Componentes principales y análisis de factores. En: Norman, Streiner, editores. *Bioestadística.* Barcelona: Mosby/Doyma libros; 1996. p. 129–42.
40. Tinsley HEA, Tinsley DJ. Uses of factor analysis in counseling psychology research. *J Counseling Psychology.* 1987;34:414–24.
41. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Statist Med.* 2002;21:1131–5.
42. Lin L, Hedayat AS, Wu W. A unified approach for assessing agreement for continuous and categorical data. *J Biopharmaceutical Statistics.* 2007;17:629–52.
43. Giraudeau B, Mary JY. Planning a reproducibility study: How many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient. *Statist Med.* 2001;20:3205–14.