



# La versión en español del QuickDASH: Fiabilidad, error de medición, validez de constructo y sensibilidad para la evaluación de resultados en el Síndrome del Túnel Carpiano

## *The Spanish Version of the QuickDASH. Reliability, Measurement Error, Construct Validity and Responsiveness for Outcomes Assessment in Carpal Tunnel Syndrome*

Roberto S. Rosales<sup>1</sup>  Hugo Sambad-Rodriguez<sup>1</sup> Carmen Menaya-Fernandez<sup>1</sup>  
Yolanda Martin-Hidalgo<sup>2</sup> Luis Reboso-Morales<sup>2</sup>

<sup>1</sup>Unidad de Mano y Microcirugía, GECOT, La Laguna, Santa Cruz de Tenerife, España

<sup>2</sup>Departamento de Ortopedia, University Hospital of La Candelaria, Santa Cruz de Tenerife, España

Address for correspondence Roberto S. Rosales, MD, MSCR, PhD, Unidad de Mano y Microcirugía, GECOT, La Laguna, Santa Cruz de Tenerife, España (e-mail: titorosales@telefonica.net).

Rev Iberam Cir Mano 2024;52(2):e114–e123.

### Resumen

**Objetivo** El propósito de este artículo fue evaluar la confiabilidad, el error de medición, la validez de construcción y la respuesta de la versión española del QuickDASH para la evaluación de resultados en el síndrome del túnel carpiano (STC). **Métodos** Una población total de 40 pacientes con diagnóstico de STC completó las versiones en español del QuickDASH, CTS-AL y EQ-5D una semana antes de la cirugía, el día previo a la cirugía y tres meses después de la intervención. La consistencia interna fue analizada mediante el coeficiente  $\alpha$  de Cronbach. Para evaluar la concordancia absoluta se utilizaron el coeficiente de correlación intraclase ( $ICC_{2,1}$ ), el coeficiente de correlación de Lin (CCC) y el límite de acuerdo de Bland y Altman (LoA), mientras que el método no paramétrico de Passing-Bablok (P-B) se empleó para evaluar el sesgo sistemático constante y/o proporcional entre las mediciones. La precisión transversal se analizó con el error estándar de medición (SEM). La precisión longitudinal para el coeficiente de fiabilidad test-retest se analizó con el error estándar de la diferencia de medición ( $SEM_{diff}$ ) y el cambio mínimo detectable al 90 % ( $MDC_{90}$ ) y al 95 % ( $MDC_{95}$ ) de nivel de confianza. Para evaluar la validez de construcción, se formuló la hipótesis de que el QuickDASH tendría una fuerte correlación positiva con el CTS-AL y una

### Palabras Clave

- fiabilidad
- validez
- validez de construcción
- QuickDASH
- Síndrome del Túnel Carpiano

recibido  
02 de mes de julio de 2024  
aceptado  
11 de octubre de 2024

DOI <https://doi.org/10.1055/s-0044-1793950>.  
ISSN 1698-8396.

© 2024. SECMA Foundation. All rights reserved.  
This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)  
Thieme Revinter Publicações Ltda., Rua do Matoso 170, Rio de Janeiro, RJ, CEP 20270-135, Brazil

correlación negativa moderada con el índice EQ-5D, utilizando el coeficiente de correlación de Pearson ( $r$ ) con un nivel de significancia de 0.05.

**Resultados** El coeficiente de Cronbach fue de 0.912. ICC<sub>2,1</sub> y CCC mostraron una concordancia absoluta alta (0.868 y 0.738 respectivamente). La regresión P-B no mostró diferencias constantes y proporcionales significativas entre las dos administraciones del QuickDASH. El QuickDASH mostró una respuesta (ES = 2.1; SRM = 1.97) inferior al CTS-AL (ES = 3.53; SRM = 3.50) y superior al índice EQ-5D (ES = 0.78; SRM = 0.83).

**Conclusión** La versión española (España) del instrumento QuickDASH presentó buen nivel de confiabilidad, validez de construcción y respuesta para la evaluación de resultados en STC.

## Abstract

**Objective** This paper's purpose was to assess the reliability, measurement error, construct validity, and responsiveness of the Spanish version of the QuickDASH for outcomes assessment in Carpal Tunnel Syndrome (CTS).

**Methods** A total population of 40 patients was diagnosed with CTS. The Spanish versions of the QuickDASH, CTS-AL, and EQ-5D were completed by the patients 1 week, the day before surgery, and 3 months after the surgery. Internal consistency was analyzed by Cronbach  $\alpha$  coefficient. The intraclass correlation (ICC<sub>2,1</sub>), Lin's coefficient (CCC), and Bland and Altman Limit of Agreement (LoA) were used for the absolute agreement assessment, and the non-parametric Passing-Bablok (P-B) were used to assess constant and proportional systematic bias between the measurements. Cross-sectional precision was analyzed with the Standard Error of the Measurement (SEM). Longitudinal precision for the test-retest reliability coefficient was analyzed with the Standard Error of the Measurement difference (SEMdiff) and the Minimal Detectable Change at 90% (MDC90) and 95% (MDC95) confidence levels. For assessing construct validity, we hypothesized that the QuickDASH would have a strong positive correlation with the CTS-AL and a moderate negative correlation with the EQ-5D Index, using a Pearson correlation coefficient ( $r$ ) with a level of significance of 0.05.

**Results** Cronbach coefficient was 0.912. ICC<sub>2,1</sub> and CCC showed a high absolute agreement (0.868 and 0.738 respectively). The P-B regression showed no significant constant and proportional differences between the two administrations of QuickDASH. The QuickDASH showed a responsiveness (ES = 2.1; SRM = 1.97) lower than the CTS-AL (ES = 3.53; SRM = 3.50), and higher than the EQ-5D Index (ES = 0.78; SRM = 0.83).

**Conclusion** The Spanish (Spain) version of the QuickDASH instrument presented a good level of reliability, construct validity, and responsiveness for outcomes assessment in CTS.

## Keywords

- ▶ reliability
- ▶ validity
- ▶ construction validity
- ▶ QuickDASH
- ▶ Carpal Tunnel Syndrome

## Introducción

El uso de instrumentos de evaluación de resultados de salud, llamados resultados informados por los pacientes (PRO), ha crecido en la investigación clínica en cirugía de mano. Uno de los instrumentos PRO más utilizados ha sido el DASH (Disability of the Arm, Hand, and Shoulder). En 2005, Beaton et al.<sup>1</sup> desarrollaron una versión abreviada al comparar tres métodos de reducción de ítems (método de retención de conceptos, correlación ítem-total equidiscriminativa y teoría de respuesta al ítem o modelo de Rasch) con el propósito de reducir la carga para los encuestados, manteniendo las propiedades psicométricas

del instrumento PRO. La versión abreviada, QuickDASH, fue una versión reducida de la medida de resultados DASH, y se constituyó por 11 ítems en lugar de 30 para medir la función física y los síntomas en personas con uno o más trastornos musculoesqueléticos de la extremidad superior. Considerando que ya existía una versión en español del DASH con buen nivel de fiabilidad, validez y sensibilidad,<sup>2-4</sup> se lanzó y aprobó una nueva versión en español de la versión abreviada, QuickDASH, por el Institute for Work and Health (IWH) de Ontario, Canadá ([https://dash.iwh.on.ca/sites/dash/public/translations/QuickDASH\\_Spanish\\_Spain\\_2018.pdf](https://dash.iwh.on.ca/sites/dash/public/translations/QuickDASH_Spanish_Spain_2018.pdf)). Sin embargo, no se

ha reportado nueva evidencia sobre las propiedades psicométricas de la versión en español del instrumento QuickDASH.

El propósito de este artículo fue evaluar la confiabilidad, el error de medición, la validez de constructo y la capacidad de respuesta de la versión española del QuickDASH para la evaluación de resultados en el síndrome del túnel carpiano (STC).

## Métodos

### Población de estudio

La población del estudio consistió en 40 pacientes consecutivos (27 mujeres) con una edad media de 54.5 años (DE 11.2) (→ **Tabla 1**) diagnosticados de STC según criterios clínicos y neurofisiológicos, reclutados de la lista de espera para cirugía de STC en el Sistema Nacional de Salud en Tenerife, España. Todos los procedimientos realizados en este estudio que involucraron a participantes humanos estuvieron de acuerdo con las normas éticas del comité nacional de investigación institucional del Hospital Universitario de La Candelaria, Facultad de Medicina, Universidad de La Laguna, y con la Declaración de Helsinki de 1964 y sus enmiendas posteriores o estándares éticos comparables. El comité de ética revisó y aprobó este estudio (Protocolo PI-11/16). Se obtuvo el consentimiento informado por escrito de todos los participantes individuales incluidos en el estudio.

**Tabla 1** Datos demográficos

	Total (N = 40)
<b>Edad</b>	
Media (DE)	54.48 (11.18)
Mediana (Q1, Q3)	50.5 (47.5, 59.5)
<b>CTS-AL_pre</b>	
Media (DE)	3.76 (0.68)
Mediana (Q1, Q3)	3.9 (3.4, 4.2)
<b>QuickDASH1</b>	
Media (DE)	63.88 (19.46)
Mediana (Q1, Q3)	65.9 (53.4, 77.3)
<b>Eq. 5D_pre</b>	
Media (DE)	0.56 (0.31)
Mediana (Q1, Q3)	0.7 (0.2, 0.8)
<b>Género</b>	
Masculino	13 (32.5%)
Femenino	27 (67.5%)
<b>Mano afectada</b>	
Derecha	26 (65.0%)
Izquierda	14 (35.0%)

Abreviaturas: Q1, primer cuartil; Q3, tercer cuartil; DE, desviación estándar.

### Criterios de elegibilidad

Los criterios de inclusión fueron: (1) entumecimiento u hormigueo con o sin dolor en al menos 2 de los dedos de la distribución del nervio mediano,<sup>5,6</sup> (2) aumento de los síntomas con las pruebas de provocación del túnel carpiano (prueba de Phalen y/o prueba de Phalen inversa),<sup>6</sup> (3) síntomas durante más de dos meses,<sup>5</sup> (4) fracaso del tratamiento conservador,<sup>5</sup> y una prueba de conducción nerviosa que muestra neuropatía mediana en la muñeca (latencia motora distal > 4,5 milisegundos, latencia sensorial de los dedos de la muñeca > 3,5 milisegundos, o velocidad de conducción sensorial en el segmento del túnel carpiano < 40 m/segundo).<sup>7,8</sup> Los criterios de exclusión fueron: signos clínicos o electrofisiológicos de compresión del nervio proximal, diabetes o cualquier enfermedad metabólica, artritis reumatoide u otras enfermedades inflamatorias generales.<sup>5,6,9</sup>

#### Diseño clínico

Se realizó un estudio observacional con un diseño transversal para el Error Estándar de Medición (EEm) y análisis de validez de constructo y un diseño de cohorte clásico para evaluación de confiabilidad y capacidad de respuesta test-retest que se adhirió a las guías STROBE como diseño clínico utilizado en esta investigación.<sup>10</sup>

### Instrumentos y Medidas de Resultados

Las versiones estándar en español de los cuestionarios QuickDASH ([www.iwh.on.ca](http://www.iwh.on.ca)), Síndrome del Túnel Carpiano 6 ítems de Atroshi-Leryn (CTS-AL)<sup>11</sup> y Euroqol 5D Index (EQ-5D)<sup>12</sup> fueron completadas por los pacientes 1 semana antes de la cirugía.

QuickDASH es la versión más corta de los instrumentos DASH PRO desarrollados para medir la discapacidad de las extremidades superiores. Está constituido por 11 ítems y se puntúa de 0 (mejor = menor discapacidad) a 100 (peor = mayor discapacidad). Se deben completar al menos 10 de los 11 elementos para que se calcule una puntuación. Los valores asignados para todas las respuestas completadas simplemente se suman y se promedian, lo que produce una puntuación sobre cinco. Luego, este valor se transforma en una puntuación sobre 100 restando uno y multiplicando por 25. Esta transformación se realiza para que la puntuación sea más fácil de comparar con otras medidas escaladas en una escala de 0 a 100. En este trabajo se utilizaron las herramientas del servicio de cálculo desarrollado por el Instituto del Trabajo y la Salud para obtener los puntajes QuickDASH. [http://www.orthopaedicscore.com/scorepages/disabilities\\_of\\_arm\\_shoulder\\_hand\\_score\\_quickdash.html](http://www.orthopaedicscore.com/scorepages/disabilities_of_arm_shoulder_hand_score_quickdash.html).<sup>13</sup>

El CTS-AL mide la gravedad de los síntomas relacionados con la STC. Está constituido por 6 ítems. Cinco de los 6 ítems en el CTS-AL tienen un texto similar a los ítems correspondientes en la escala de gravedad de síntomas de 11 ítems (cuestionario estándar CTS-SS).<sup>14</sup> El ítem restante (resultado de la fusión de 2 ítems de la escala de gravedad de síntomas) contiene texto de ambos ítems. Sin embargo, el CTS-AL presenta un diseño completamente diferente y mejorado. La puntuación es similar a la escala de gravedad de síntomas de 11 ítems; para cada paciente, las respuestas de los ítems se puntúan

de 1 (mejor) a 5 (peor) y luego se promedian para los 6 ítems, obteniendo así una puntuación CTS-AL (se permite únicamente 1 respuesta faltante en los ítems).

El índice genérico de salud EuroQol (EQ-5D) incluye un cuestionario de cinco partes y una escala analógica visual de autovaloración. En este estudio, el EQ-5D se utilizó como un índice de calidad de vida relacionada con la salud, con cinco dimensiones: movilidad (MO), cuidado personal (CP), actividades cotidianas (AC), dolor/malestar (DM) y ansiedad/depresión (AD). El rango posible para cada una de las variables de dimensión es de 1 a 3, donde 1 = sin problemas, 2 = problemas moderados y 3 = problemas extremos. Una vez que el conjunto de datos analíticos tuvo las variables de dimensión adecuadamente nombradas, el índice EQ-5D se generó utilizando el algoritmo de puntuación EQ-5D TM para Excel <http://www.ahrq.gov/professionals/clinicians-providers/resources/rice/EQ5Dscore.html>. Las puntuaciones del índice EQ-5D variaron de 0 (muerte) a 1 (mejor salud). Algunos pacientes con enfermedades graves de larga duración pueden tener estados de salud con valores de utilidad inferiores a cero, es decir, desde una perspectiva social, se consideran en estados "peores que la muerte"<sup>15</sup>. La puntuación del índice no se calculará si faltan respuestas en una o más de las dimensiones. En este estudio, no se observaron ítems faltantes en ninguno de los tres instrumentos PRO.

Para evaluar la fiabilidad test-retest, se realizó una segunda autoadministración del QuickDASH 1 semana después de la primera administración y se respondió un cuestionario simple que constaba de preguntas sobre los cambios en el estado de salud de las extremidades superiores durante la semana anterior. Finalmente, QuickDASH se autoadministró a la población de muestra 3 meses después de la liberación abierta del túnel carpiano para analizar la capacidad de respuesta.

La propiedad de medición "validez de constructo" incluyó tres aspectos: validez estructural, prueba de hipótesis y validez transcultural. Se utilizó prueba de hipótesis para analizar la validez de constructo del QuickDASH. La validez de la prueba de hipótesis de constructo es el grado en que las puntuaciones de un instrumento PRO son consistentes con las hipótesis basadas en el supuesto de que el instrumento PRO mide válidamente el constructo que se va a medir.<sup>16-18</sup> Para evaluar la validez de constructo, se planteó la hipótesis a priori de que QuickDASH tendría una correlación positiva de moderada a fuerte con el CTS-AL y una correlación negativa de baja a moderada con el índice EQ-5D.

La capacidad de respuesta es la capacidad de un instrumento PRO para detectar cambios a lo largo del tiempo en el constructo a medir y está relacionada con la validez longitudinal o las puntuaciones de cambio a lo largo del tiempo. Para ello se estableció a priori la hipótesis: el QuickDASH, como instrumento específico para extremidades superiores, tendría una capacidad de respuesta menor que el CTS-AL (instrumento PRO específico para enfermedades), y mayor que un instrumento genérico como el EQ-5D Index.

## Análisis de datos

### Fiabilidad

La consistencia interna, o el grado en que el QuickDASH midió un solo concepto, se evaluó con el coeficiente  $\alpha$  de Cronbach ( $\alpha > 0,7$  indicó una buena consistencia interna). La reproducibilidad o fiabilidad test-retest se analizó mediante el coeficiente de correlación intraclase (modelo de efectos aleatorios bidireccional y definición de concordancia absoluta ( $ICC_{2,1}$ ),<sup>18</sup> Coeficiente de correlación de concordancia de Lin (CCC), Límites de concordancia (LoA) de Bland y Altman y el modelo de regresión ortogonal no paramétrico de Passing-Bablok (P-B) entre las dos administraciones antes de la cirugía, tomando 1 semana como tiempo de lavado.

Los límites de acuerdo de Bland-Altman (LoA)<sup>19</sup> calculan la diferencia entre ambas mediciones para cada sujeto ( $di = Y - X$ ) y se comparan con la media  $(X_i + Y_i) / 2$  de ambas mediciones para cada individuo. Si se asume una distribución normal de las diferencias, se espera que el 95 % de las diferencias se sitúe entre los límites del intervalo, lo cual indica un buen nivel de concordancia. La normalidad de la distribución de las diferencias fue analizada mediante la prueba de Shapiro-Wilk; además, también se analizaron la curtosis y la asimetría, con un nivel de significancia de 0.05.

La regresión de Passing-Bablok (P-B) se utilizó en caso de no necesitar asumir una distribución normal de las diferencias. El análisis P-B es una estimación no paramétrica de la línea de regresión ortogonal entre dos métodos o mediciones. La ecuación lineal es  $Y = A + BX + \epsilon$ , donde A representa la diferencia constante entre las dos mediciones, B la diferencia proporcional, y  $\epsilon$  es una variable aleatoria con media igual a cero, que representa el error aleatorio no sistemático entre ambas mediciones. Cuando  $A = 0$  y  $B = 1$ , significa que ambas mediciones presentan el mismo error, lo cual indica una excelente concordancia y que ambas son comparables e intercambiables. Con base en la regresión no paramétrica P-B, se pudo determinar si existían diferencias sistemáticas significativas constantes y/o proporcionales entre ambas mediciones.<sup>18</sup>

### Error de medición

La precisión transversal se analizó con el error estándar de medición ( $SEM = DE$  multiplicado por la raíz cuadrada del coeficiente de  $1 - Cronbach$ ). La precisión longitudinal para el coeficiente de confiabilidad test-retest se analizó con el Error Estándar de la diferencia de Medición ( $SE_{diff} = SD$  multiplicado por la raíz cuadrada de  $1 - ICC$  multiplicada por la raíz cuadrada de 2) y el Cambio Mínimo Detectable al 90% de nivel de confianza ( $MDC_{90} = diferencia SEM$  multiplicada por 1,65) y nivel de confianza del 95% ( $MDC_{95} = diferencia SEM$  multiplicada por 1,96).<sup>20</sup>

### Hipótesis de validez de constructo

Las hipótesis de validez de constructo se analizaron con el coeficiente de correlación de Pearson ( $r$ ), utilizando un nivel

de significancia estadística de 0.05. Valores entre 0,8 y 1,0 indican una relación muy fuerte, entre 0,6 y 0,8 una relación fuerte, entre 0,4 y 0,6 una relación moderada, entre 0,2 y 0,4 una relación débil y menos de 0,2 una relación muy débil o ninguna relación.<sup>21</sup>

### Sensibilidad

La sensibilidad de cada instrumento PRO se evaluó calculando el tamaño del efecto (ES = cambio promedio de las puntuaciones / desviación estándar (DE) de las puntuaciones iniciales) y la media de respuesta estandarizada (SRM = cambio promedio de las puntuaciones / DE del cambio). Una SRM o un ES alto indica una alta sensibilidad al cambio, y un ES o SRM >0.8 representa una mejora clínica significativa.<sup>18</sup>

### Tamaño de la muestra

Para un ICC esperado de 0.85 en la fiabilidad test-retest, con un ancho del intervalo de confianza (IC) del 95% de 0.20 y 2 medidas, el tamaño de muestra esperado fue de 31. Para la validez de constructo, el cálculo del tamaño de muestra a priori para el análisis de correlación mostró que, basándose en la hipótesis nula propuesta ( $H_0$  = la correlación es igual a cero), con una prueba bilateral, un nivel de significancia de 0.05, una potencia del 80% y un  $r$  mínimo esperado de 0.4, se necesitaría un tamaño de muestra de 37 pacientes (Stata 16.1, StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, EE. UU.).

## Resultados

### Fiabilidad

El análisis de consistencia interna mostró un  $\alpha$  de Cronbach de 0.912 con una covarianza interítem promedio de 0.522. El

análisis de concordancia absoluta test-retest demostró un  $ICC_{2,1}$  de 0.868 (IC del 95%: 0.750 a 0.930) ( $p < 0.001$ ) y un CCC de 0.738 (IC del 95%: 0.597 a 0.877) ( $p < 0.001$ ). El análisis de Bland-Altman mostró que la diferencia absoluta estimada de sesgo entre ambas administraciones del QuickDASH fue de 5.47 unidades (DE 13.39), con un LoA de -20.767 a 31.707, siendo el 95% de los casos incluidos en el intervalo (2 casos sobre el límite, lo que representa un 5%, y ningún caso por debajo del límite) (►Fig. 1). Sin embargo, no se confirmó la suposición de normalidad de la distribución de las diferencias (prueba de Shapiro-Wilk  $p = 0.005$ , prueba de asimetría  $p = 0.0019$ , prueba de curtosis  $p = 0.0029$ ).

La regresión de Passing-Bablok mostró una diferencia mediana entre ambas administraciones de las puntuaciones del QuickDASH en la fiabilidad test-retest de 2.3 unidades, con diferencias constantes y proporcionales no significativas (►Tabla 2) (►Fig. 2).

### Error de Medición

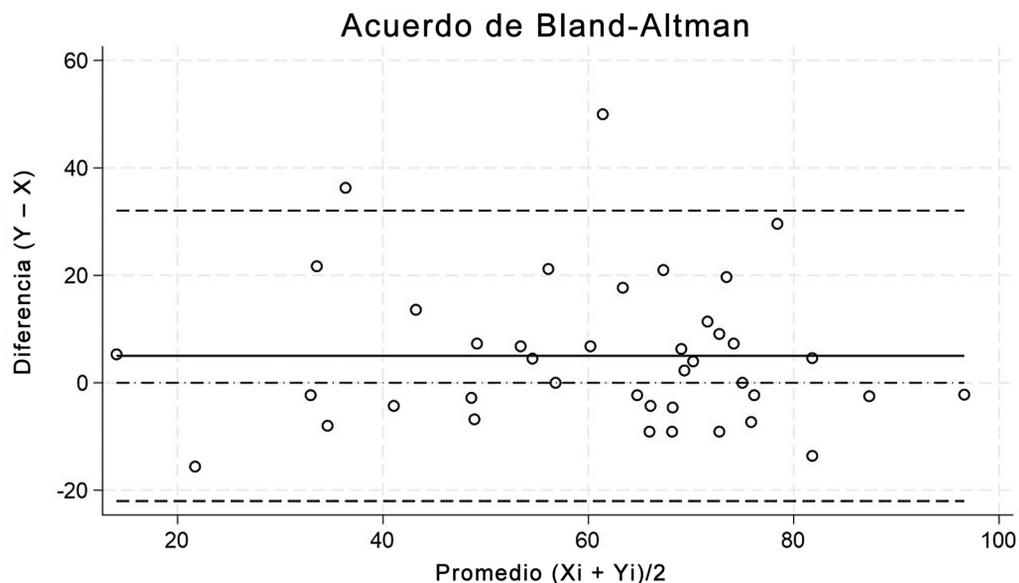
El análisis de precisión transversal demostró un SEM de 5,785. La precisión longitudinal mostró una  $SE_{diff}$  de 10,0001 con un  $MDC_{95} = 19,6$  y un  $MDC_{90} = 16,5$  (►Tabla 3).

### Validez de constructo

El QuickDASH presentó una fuerte correlación positiva con el CTS-AL ( $r = 0,635$ ) y una correlación negativa moderada con el índice EQ-5D ( $r = -0,492$ ), siendo relevante ( $p < 0,001$ ) (Material complementario).

### Sensibilidad

El QuickDASH mostró una capacidad de respuesta (ES = 2,1; SRM = 1,97) inferior al CTS-AL (ES = 3,53; SRM = 3,50) y superior al índice EQ-5D (ES = 0,78; SRM = 0,83) (►Tabla 4).



**Fig. 1** Límites de Acuerdo de Bland-Altman. Los límites de acuerdo de Bland-Altman (LoA) calculan la diferencia entre ambas medidas para cada sujeto ( $d_i = Y - X$ ) y se comparan con el promedio ( $(X_i + Y_i)/2$ ) de ambas medidas para cada individuo. Se muestra que el 95% de los casos estaban incluidos en el intervalo (solo 2 casos sobre el límite, lo que representa un 5%, y ningún caso por debajo del límite).

Bland-Altman Agreement = Acuerdo de Bland-Altman

Average  $(X_i + Y_i)/2$  = Promedio  $(X_i + Y_i)/2$

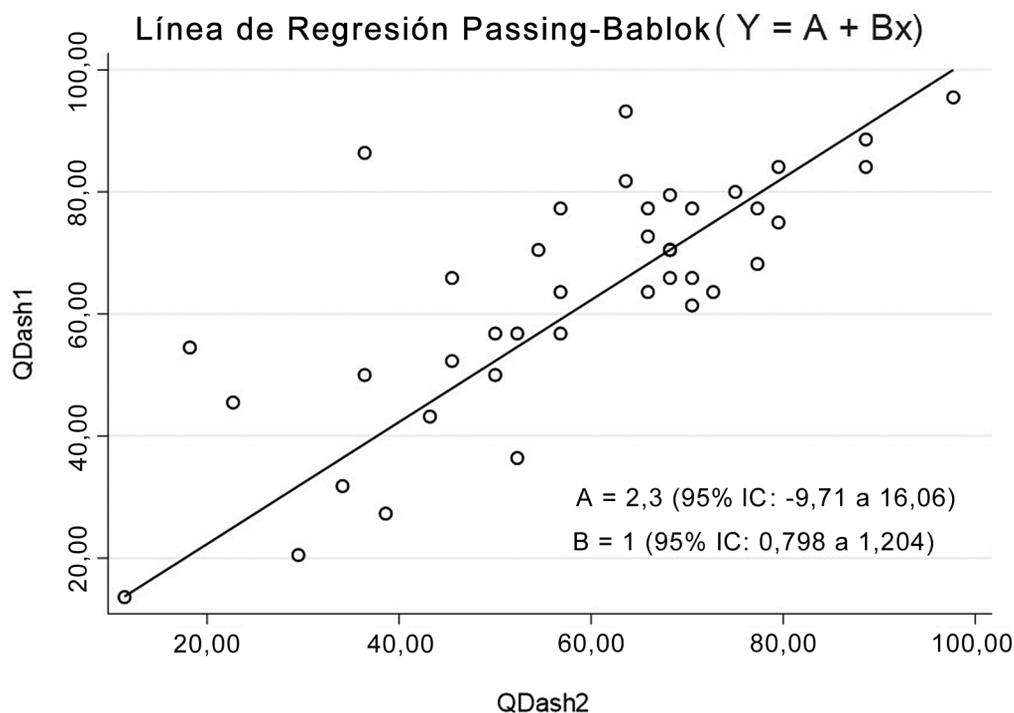
Difference  $(Y - X)$  = Diferencia  $(Y - X)$

**Tabla 2** Acuerdo Test-Retest de Passing-Bablok: estadísticas descriptivas

Variable	Válida	Perdida	Obs	Mediana	Media	Mínimo	Máximo	DE
Y: QDash1	40	0	40	65.9	63.88	13.6	95.5	19.45837
X: QDash2	40	0	40	63.6	58.41	11.4	97.7	19.73363
Y-X	40	0	40	2.3	5.47	-15.9	50	13.38622
100*(Y-X)/X	40	0	40	4.58%	16.0%	-30.5%	199.5%	43.0%

Abreviaturas: QDash1, primera administración de QuickDASH; QDash2, segunda administración de QuickDASH con 1 semana de diferencia; DE, desviación estándar.

100\*(Y-X)/X = porcentaje de las diferencias tomando como referencia el QDash2.



**Fig. 2** Línea de Regresión Passing-Bablok. A es la diferencia constante entre las dos medidas, B es la diferencia proporcional. Si A = 0 y B = 1, significa que ambas medidas presentaron el mismo error y mostraron una excelente concordancia. Observar que las hipótesis nulas (A = 0 y B = 1) estaban incluidas en el Intervalo de Confianza del 95% y que el P-B mostró diferencias constantes y proporcionales no significativas. Passing-Bablok Regression Line = Línea de Regresión Passing-Bablok

CI = IC -9.71 to 16.06 = -9,71 a 16,06

0.798 to 1.204 = 0,798 a 1,204

QDASH1 = QDASH1

QDASH2 = QDASH2

**Tabla 3** Error de medición del QuickDASH

A) Precisión transversal				
Cronbach α	DE	SEM	transversal Precisión. IC del 95%	Precisión transversal. 90% IC
0.912	19.458	5.785	-/+ 11.339	-/+ 9.546
B) Precisión longitudinal				
ICC <sub>2,1</sub>	DE	SE <sub>diff</sub>	MDC <sub>95</sub>	MDC <sub>90</sub>
0.868	19.4584	10.000	19.600	16.500

Abreviaturas: IC, intervalo de confianza; CCI: coeficiente de correlación intraclase; MDC, cambio más pequeño detectable (SDC) definido por COSMIN; MDC<sub>90</sub>: cambio mínimo detectable con un nivel de confianza del 90%; MDC<sub>95</sub>: Cambio mínimo detectable con un nivel de confianza del 95%; DE: desviación estándar; SE<sub>diff</sub>, Error Estándar de la diferencia de Medición; SEM, Error Estándar de la Medición.

**Tabla 4** Capacidad de respuesta y validez de constructo longitudinal del QuickDASH

Instrumento PRO	Pre-op	Post-op	Cambio pre-post	Sensibilidad	
	Media (DE)	Media (DE)	Media (DE)	ES	SRM
CTS_AL	3.76 (0.68)	1.38 (0.44)	2.38 (0.68)	3.53	3.50
QuickDASH	63.88 (19.46)	23.01 (13.19)	40.87 (20.75)	2.10	1.97
Eq. 5D Index	0.56 (0.31)	0.80 (0.19)	0.24 (0.29)	0.78	0.83

Abreviaturas: CTS\_AL, Síndrome del Túnel Carpiano 6 ítems de la escala Atroshi-Leryn; Eq. 5D, escala EuroQol; ES, tamaño del efecto; Post-op, Postoperatorio; Pre-op, Preoperatorio; Instrumento PRO, Instrumento de resultados informados por los pacientes; QuickDASH, versión más corta del instrumento DASH PRO; DE: desviación estándar; SRM, media de respuesta estandarizada.

## Discusión

Los resultados han demostrado que el instrumento PRO QuickDASH presenta un buen nivel de consistencia interna y un alto nivel de fiabilidad test-retest, con un excelente nivel de coeficientes de concordancia absoluta ( $ICC_{2,1}$ , CCC) sin diferencias sistemáticas constantes o proporcionales significativas entre las puntuaciones de las dos administraciones, realizadas con una semana de diferencia. Los resultados de los análisis de validez y sensibilidad coincidieron con las hipótesis de constructo formuladas a priori, lo que prueba una buena validez de constructo transversal y longitudinal.

Se han desarrollado diferentes herramientas para evaluar la calidad de un instrumento PRO. El trabajo del equipo original del Medical Outcomes Trust, que desarrolló los instrumentos genéricos clásicos, el SF-36 y el SF-12, sirvió de base para dos directrices: "Evaluating the Measurement of Patient-Reported Outcomes" (EMPRO),<sup>22</sup> y "Consensus-based standards for the selection of health measurement instruments" (COSMIN).<sup>17,23</sup> Con base en el COSMIN, se pueden distinguir tres dominios de calidad principales para evaluar un instrumento PRO: fiabilidad, validez y sensibilidad. Cada dominio puede incluir diferentes propiedades de medición.<sup>18</sup> La fiabilidad incluía tres propiedades de medición: consistencia interna, fiabilidad de prueba y repetición y error de medición. El coeficiente  $\alpha$  de Cronbach es una estimación de la consistencia interna y depende del número de ítems de una escala y de su magnitud de intercorrelación y valores  $>0,7$  indican una buena consistencia interna. En esta investigación observamos un  $\alpha$  de Cronbach de 0,912 muy cercano al reportado por Gabel,<sup>24</sup> Alnahdi,<sup>25</sup> y Claro da Silva<sup>26</sup>; y ligeramente inferior al valor de 0,94 reportados por Hammond et al.<sup>27</sup> (versión británica), Beaton et al.<sup>1</sup> en el desarrollo del QuickDASH y al reportado por Schønnemann et al.<sup>28</sup> ( $\alpha = 0,96$  versión danesa) [2016]. La concordancia absoluta en el análisis de confiabilidad test-retest demostró un excelente nivel de concordancia con un  $ICC_{2,1}$  de 0,85 y un CCC de 0,74 muy similar al QuickDASH brasileño ( $ICC_{2,1} = 0,81$ )<sup>29</sup> (► **Tabla 5**). El error de medición es otra propiedad psicométrica incluida en el dominio de confiabilidad por el COSMIN. SEM o precisión transversal nos brinda información sobre la variación alrededor de una puntuación observada. La verdadera puntuación se encuentra dentro de este rango. En esta investigación, el

SEM de 5,78 fue ligeramente superior en comparación con el informado por Beaton et al en la descripción de la versión más larga DASH (SEM = 4,4).<sup>1</sup> Para la precisión longitudinal de la medición calculamos el  $SE_{diff}$ , MDC<sub>90</sub> y MDC<sub>95</sub>. El MDC es la puntuación de cambio mínima requerida antes de que se pueda considerar con confianza que un individuo ha cambiado por algo más que la variabilidad diaria.<sup>30</sup> En esta investigación, nuestro resultado en el nivel de MDC<sub>90</sub> = 16,5 fue muy similar al rango de 11,0–17,2 informado en una revisión sistemática de la calidad de las propiedades de medición del QuickDASH realizada por Kennedy et al.<sup>31</sup> Sin embargo, nuestro MDC<sub>95</sub> de 19,6 fue inferior al publicado por Mintken et al.<sup>32</sup> Cuando comparamos el MDC, debemos tener en cuenta que este índice de fiabilidad absoluta depende de la población del estudio, el intervalo de lavado, el momento del seguimiento en el que se realizó el análisis test-retest y el DE de los datos.<sup>3</sup> Un aspecto importante en la discusión sobre la fiabilidad test-retest es que la puntuación media de diferencia entre las dos administraciones del QuickDASH realizadas con una semana de diferencia (5,47) (► **Tabla 2**) fue inferior al MDC<sub>95</sub>, lo que prueba que ese cambio debe considerarse como un cambio más allá de la variabilidad diaria.

El COSMIN estableció que las estadísticas apropiadas para evaluar el error de medición son el SEM, los límites de acuerdo (LoA) y el cambio detectable más pequeño (SDC) o MDC. Los cambios dentro de los límites de acuerdo o menores que el MDC<sub>95</sub> son probablemente atribuibles al error de medición, mientras que los cambios fuera de los límites de acuerdo o mayores que el SDC deben considerarse como cambios reales. En esta investigación, incluimos los límites de acuerdo como una estadística apropiada para la concordancia test-retest porque Bland and Altman<sup>33</sup> desarrollaron los límites de acuerdo como un análisis alternativo al ICC; sin embargo, la suposición de normalidad de las diferencias entre las dos medidas fue violada (Shapiro-Wilk  $p = 0,005$ ) en nuestra muestra. En consecuencia, los resultados de los límites de acuerdo no deberían ser considerados.

Rosales y Atroshi<sup>18</sup> introdujeron en el campo de la cirugía de mano el uso del análisis no paramétrico de Passing-Bablok (P-B) para determinar errores sistemáticos constantes y/o proporcionales entre dos medidas, con la ventaja de que no era necesaria la suposición de normalidad. Nuestros resultados demostraron que no había diferencias

**Tabla 5** Revisar traducciones de QuickDASH desde 2011 hasta 2023

Dash traducido	Autor del estudio, año	Población del estudio	Pruebas reportadas en el estudio			Validez	Sensibilidad
			Fiabilidad: Consistencia Interna (Alfa de Cronbach)	Fiabilidad: Test-retest (ICC a menos que se indique lo contrario)	ICC (2,1) = 0.91		
árabe	Alnahdi 2021 <sup>25</sup>	Participantes con trastornos musculoesqueléticos de las extremidades superiores	QuickDASH 0.90	ICC (2,1) = 0.91	Construtto GAF r = -0,53 PRNP r = 0,52 RANGO 36: Funcionamiento físico r = -0,77 Bienestar emocional r = -0,47 Dolor r = -0,75	Cambio observado en la tercera entrevista. SRM = 0,94	
brasileño	Claro da Silva 2020 <sup>26</sup>	Participantes con trastornos de las extremidades superiores.	DASH 0.93 QuickDASH 0.88	ICC = 0.81	Construtto DASH r = 0.91 SF-12 Componente físico r = -0,55 Componente mental r = -0,49		
Chino	Cao 2019 <sup>33</sup>	Pacientes que padecen trastornos crónicos de las extremidades superiores	QuickDASH 0.81	ICC = 0.90	Construtto Dash r = 0,820 EVA r = 0,463 Subescalas SF-36 Función física r = -0,630 Rol físico r = -0,471 Dolor corporal r = -0,563 Salud general r = -0,414 Vitalidad r = -0,053 Función social r = -0,178 Rol emocional r = -0,010 Salud mental r = -0,165	-	
Danés	Schønnemann 2016 <sup>28</sup>	Pacientes con fractura de radio distal.	QuickDASH 0.96	ICC = 0.94	-	-	
Inglés (británico)	Hammond 2018 <sup>27</sup>	Pacientes con artritis reumatoide	DASH 0.98 DASH Work 0.94	ICC (2,1) = 0.95	Concurrente QuickDASH r = 0.61-0.91 DASH r = 0.61-0.99 DASH Trabajo r = 0.53-0.80 DASH Spam r = 0.52-0.78		
Italiano	Franchignoni 2011 <sup>34</sup>	Pacientes con trastornos de las extremidades superiores	QuickDASH 0.87	-	-	-	

(Continued)

Tabla 5 (Continued)

Dash traducido	Autor del estudio, año	Población del estudio	Pruebas reportadas en el estudio			Validez	Sensibilidad
			Fiabilidad: Consistencia Interna (Alfa de Cronbach)	Fiabilidad: Test-retest (ICC a menos que se indique lo contrario)	Fiabilidad: Test-retest (ICC a menos que se indique lo contrario)		
coreano	Hong 2018 <sup>29</sup>	Pacientes con síndrome del túnel carpiano	QuickDASH 0,89 ADL 0,89 Actividad Social 0,70 Síntoma 0,72	ICC (2,1) = prueba para cada ítem osciló entre 0,64 y 0,98 y 0,83 para todos los QuickDASH	Construc-to DASH r = 0,975 = 0,975 Dolor EVA r = 0,365 Escala Blend de SNC r = -0,074 Fuerza de agarre r = -0,309 Fuerza de pinza r = -0,327	Cambio observado seis meses después de la operación. SRM = 1,00	

Abreviaturas: ADL, Actividades de la vida diaria; GAF, Evaluación Global de Función; CCI: coeficiente de correlación intraclass; NRPS: escala numérica de valoración del dolor; SRM: media de respuesta estandarizada; EVA, escala analógica visual.

sistemáticas constantes ( $A = 2.3$ ; IC del 95%: -9.71 a 16.06) ni proporcionales ( $B = 1$ ; IC del 95%: 0.80 a 1.20) entre las dos administraciones del QuickDASH, realizadas con una semana de diferencia.

El análisis de validez de constructo confirmó la hipótesis a priori como prueba de la buena validez del QuickDASH para medir el constructo propuesto mediante un diseño transversal (correlaciones entre QuickDASH, CTS-AL y EQ-5D) y mediante un diseño longitudinal o análisis de sensibilidad (comparación de ES y SRM del QuickDASH, CTS-AL y EQ-5D desde la línea base hasta 3 meses después de la cirugía).

Un punto fuerte de este estudio es que, basado en la lista de verificación del COSMIN,<sup>17,23</sup> la fiabilidad, validez y sensibilidad de este trabajo cumplieron con la mayoría de los requisitos de diseño como estándares para una buena calidad metodológica de un instrumento PRO: descripción de cómo se manejaron los ítems faltantes, porcentaje de valores faltantes, cálculo del tamaño de muestra, al menos dos mediciones independientes con un intervalo de lavado apropiado y estable de 1 semana en el que las condiciones eran similares para ambas mediciones en el análisis de fiabilidad, las pruebas de hipótesis sobre correlaciones y sensibilidad se formularon a priori antes de la recolección de datos, las direcciones esperadas de las correlaciones y tamaños de efecto se incluyeron en la hipótesis, se incluyó la magnitud absoluta o relativa esperada de los tamaños de efecto en la hipótesis, hubo una descripción adecuada de los instrumentos comparadores y sus propiedades, y finalmente, un diseño clínico y método estadístico adecuados para las hipótesis a ser probadas.

Una limitación de este estudio es la falta de análisis de interpretabilidad, que puede proporcionar información sobre la MICD (diferencia clínica mínima importante) o el grado en que se puede asignar un significado clínico cualitativo a las puntuaciones cuantitativas de un instrumento o al cambio en las puntuaciones.<sup>17,18,23</sup> Se recomiendan estudios adicionales sobre la MICD para completar el análisis de las propiedades de medición del instrumento PRO QuickDASH en español.

En conclusión, este estudio ha demostrado que la versión española (de España) del instrumento QuickDASH tiene buena fiabilidad, validez de constructo y sensibilidad para la evaluación de resultados en la STC.

#### Conflicto de intereses

Los autores declaran que no existe ningún conflicto de intereses.

#### Referencias

- 1 Beaton DE, Wright JG, Katz JN Upper Extremity Collaborative Group. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am* 2005;87(05): 1038–1046
- 2 Rosales RS, Delgado EB, Díez de la Lastra-Bosch I. Evaluation of the Spanish version of the DASH and carpal tunnel syndrome health-related quality-of-life instruments: cross-cultural adaptation process and reliability. *J Hand Surg Am* 2002;27(02):334–343

- 3 Rosales RS, Diez de la Lastra I, McCabe S, Ortega Martinez JJ, Hidalgo YM. The relative responsiveness and construct validity of the Spanish version of the DASH instrument for outcomes assessment in open carpal tunnel release. *J Hand Surg Eur Vol* 2009;34(01):72–75
- 4 Kennedy CA, Beaton DE, Solway S, McConnell S, Bombardier C. The DASH and QuickDASH outcome measure user's manual. 3rd ed. Toronto, Ontario: Institute for Work & Health; 2011
- 5 Atroshi I, Gummesson C, Johnsson R, Sprinchorn A. Symptoms, disability, and quality of life in patients with carpal tunnel syndrome. *J Hand Surg Am* 1999;24(02):398–404
- 6 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 Am* 2003;28(02):250–254
- 7 Atroshi I, Larsson GU, Ornstein E, Hofer M, Johnsson R, Ranstam J. Outcomes of endoscopic surgery compared with open surgery for carpal tunnel syndrome among employed patients: randomised controlled trial. *BMJ* 2006;332(7556):1473
- 8 Kimura J. *Electrodiagnosis in diseases of nerve and muscle: principles and practice*. 2nd ed. Philadelphia: FA Davis; 1989
- 9 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 Am* 1996;21(03):338–346
- 10 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JPSTROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61(04):344–349. Doi: 10.1016/j.jclinepi.2007.11.008
- 11 Rosales RS, Atroshi I. Spanish versions of the 6-item carpal tunnel syndrome symptoms scale (CTS-6) and palmar pain scale. *J Hand Surg Eur Vol* 2013;38(05):550–551
- 12 Badia X, Roset M, Montserrat S, Herdman M, Segura A. [The Spanish version of EuroQol: a description and its applications. European Quality of Life scale]. *Med Clin (Barc)* 1999;112(Suppl 1):79–85
- 13 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(06):602–608
- 14 Levine DW, Simmons BP, Koris MJ, et al. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. *J Bone Joint Surg Am* 1993;75(11):1585–1592
- 15 Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *Br J Rheumatol* 1997;36(05):551–559
- 16 Rosales RS, Gutiérrez R, Hidalgo Y, Reboso L, Atroshi I. Cómo escribir un artículo original en Cirugía de la Mano. *Rev Iberoam Cir Mano* 2016;44(01):47–58
- 17 Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res* 2018;27(05):1171–1179
- 18 Rosales RS, Atroshi I. The methodological requirements for clinical examination and patient-reported outcomes, and how to test them. *J Hand Surg Eur Vol* 2020;45(01):12–18
- 19 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307–310
- 20 Kennedy CA, Beaton DE, Solway S, McConnell S, Bombardier C. The DASH and QuickDASH outcome measure user's manual. 3rd ed. Toronto, Ontario: Institute for Work & Health; 2011
- 21 Chung MK. Correlation Coefficient. In: Salkin NJ, Ed. *Encyclopedia of Measurement and Statistics*. London, Sage Publications, 2007: 189–201
- 22 Valderas JM, Kotzeva A, Espallargues M, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res* 2008;17(02): 179–193
- 23 Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21(04):651–657
- 24 Gabel CP, Michener LA, Melloh M, Burkett B. Modification of the upper limb functional index to a three-point response improves clinimetric properties. *J Hand Ther* 2010;23(01):41–52
- 25 Alnahdi AH. Validity and reliability of the Arabic quick disabilities of the arm, Shoulder and Hand (QuickDASH-Arabic). *Musculoskelet Sci Pract* 2021;53:102372
- 26 da Silva NC, Chaves TC, Dos Santos JB, et al. Reliability, validity and responsiveness of Brazilian version of QuickDASH. *Musculoskelet Sci Pract* 2020;48:102163
- 27 Hammond A, Prior Y, Tyson S. Linguistic validation, validity and reliability of the British English versions of the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire and QuickDASH in people with rheumatoid arthritis. *BMC Musculoskelet Disord* 2018;19(01):118
- 28 Schønnemann JO, Eggers J. Validation of the Danish version of the Quick-Disabilities of Arm, Shoulder and Hand Questionnaire. *Dan Med J* 2016;63(12):A5306
- 29 Hong SW, Gong HS, Park JW, Roh YH, Baek GH. Validity, Reliability and Responsiveness of the Korean Version of Quick Disabilities of the Arm, Shoulder, and Hand Questionnaire in Patients with Carpal Tunnel Syndrome. *J Korean Med Sci* 2018; 33(40):e249
- 30 Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland-Morris questionnaire. *Phys Ther* 1996;76(04):359–365, discussion 366–368
- 31 Kennedy CA, Beaton DE, Smith P, et al. Measurement properties of the QuickDASH (disabilities of the arm, shoulder and hand) outcome measure and cross-cultural adaptations of the QuickDASH: a systematic review. *Qual Life Res* 2013;22(09):2509–2547
- 32 Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and Numeric Pain Rating Scale in patients with shoulder pain. *J Shoulder Elbow Surg* 2009;18(06): 920–926
- 33 Cao S, Zhou R, Zhou H, et al. Reliability and validity of Simplified Chinese version of Quick Disabilities of the Arm, Shoulder, and Hand (QuickDASH) questionnaire: cross-cultural adaptation and validation. *Clin Rheumatol* 2019;38(11):3281–3287
- 34 Franchignoni F, Ferriero G, Giordano A, Sartorio F, Vercelli S, Brigatti E. Psychometric properties of QuickDASH - a classical test theory and Rasch analysis study. *Man Ther* 2011;16(02): 177–182