

TRABAJO ORIGINAL

Confiabilidad y mínimo cambio detectable de la medición de fuerza de rotadores de hombro con handgrip.

Reliability and minimal detectable change of shoulder rotators strength measurement with handgrip dynamometer.

Confiabilidade e mínima diferença detectável da medida da força do rotador do ombro de preensão handgrip.

Paula Carolina Dib ¹, Gimena Paola Cardoso ², Mauro Federico Andreu ^{3*}, Silvina García Schustereder ⁴, Marisol Alejandra Galloli ⁴

Recibido: 16 de agosto de 2022. Aceptado: 21 de octubre de 2022.

¿Qué se sabe?

Actualmente, el dinamómetro isocinético es el *gold standard* para evaluar la fuerza muscular. El mismo presenta una alta confiabilidad inter e intra observador, pero presenta un alto costo económico y es difícil de transportar por su gran tamaño. El dinamómetro *handgrip* (HG) es un sistema hidráulico que expresa los valores en kilogramos fuerza, pero no se encuentra validado actualmente para la medición de la fuerza de rotadores de hombro.

¿Qué aporta este trabajo?

Se observó una excelente confiabilidad entre evaluadores para la rotación externa y regular a buena para los rotadores internos. Estos resultados avalan la utilidad del HG para medir la fuerza de los rotadores de hombro. Este método de evaluación podría introducirse en la práctica debido a su mayor confiabilidad respecto de las pruebas manuales subjetivos.

Resumen

Objetivos: Determinar la confiabilidad interobservador de la evaluación de la fuerza muscular de rotadores de hombro mediante el dinamómetro handgrip (HG).

Materiales y métodos: Se incluyeron 21 participantes voluntarios sanos. Se analizó la confiabilidad entre la primera y segunda medición realizada por 2 diferentes operadores para la rotación externa e interna, tanto del miembro dominante como del no dominante. Se utilizó el Coeficiente de Correlación Intraclase (CCI) con sus respectivos intervalos de confianza al 95%. Para determinar el umbral de cambio real en la fuerza de HG cuando diferentes evaluadores registran las mediciones, se calculó el mínimo cambio detectable (MCD)

Resultados: Se observó una excelente confiabilidad entre evaluador para la rotación externa dominante y no dominante (CCI= 0.80 [IC95% 0.58 - 0.92] y CCI 0.80 [IC95% = 0.56 - 0.91], respectivamente). La confiabilidad interevaluador para la rotación interna dominante resultó regular (CCI: 0.56 [IC95% 0.18 - 0.80] y para la no dominante buena (CCI:0.61 [IC95% 0.25 - 0.82])).

Conclusión: El estudio mostró que la confiabilidad interobservador de la herramienta HG para medir fuerza de hombro es excelente para los rotadores externos y de regular a buena para los rotadores internos.

Palabras clave: reproducibilidad de los resultados, dinamómetro de fuerza muscular, hombro, manguito de los rotadores, fuerza muscular, confiabilidad.

* Correspondencia: Mauro Federico Andreu, mfindreu@gmail.com

1 Sanatorio Anchorena San Martín, Buenos Aires, Argentina.

2 Hospital DF Santojanni, Ciudad Autónoma de Buenos Aires, Argentina.

3 Universidad Nacional de la Matanza, Buenos Aires, Argentina.

4 Hospital Dr. Pedro Fiorito, Buenos Aires, Argentina.



Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0).

Abstract

Objective: To determine the interrater reliability of shoulder rotator strength measurement using a hand-grip dynamometer.

Material and methods: Twenty-one healthy volunteers were included in the study. Reliability between the first and second measurements, performed by two different operators for external and internal rotation assessment of the dominant and non-dominant limb, was analyzed. Intraclass correlation coefficients (ICCs) and their corresponding 95% confidence intervals were used. In order to determine the threshold for true change in hand-grip strength when different raters record measurements, the minimal detectable change was calculated.

Results: Excellent interrater reliability was found for dominant and non-dominant external rotation (ICC = 0.80 [95% CI: 0.58-0.92] and ICC = 0.80 [95% CI: 0.56-0.91], respectively). Interrater reliability proved to be fair (ICC = 0.56 [95% CI: 0.18-0.80] to good (ICC = 0.61 [95% CI: 0.25-0.82]) for dominant and non-dominant internal rotation, respectively.

Conclusion: This study shows that interrater reliability of hand-grip dynamometer measurements is excellent for external rotators and fair to good for internal rotators.

Keywords: reproducibility of results, muscle strength dynamometer, shoulder, rotator cuff, muscle strength, reliability.

Resumo

Objetivo: Determinar a confiabilidade interobservador da medida da força do rotador do ombro utilizando o dinamômetro de prensão handgrip.

Material e métodos: Vinte e um participantes voluntários assintomáticos foram incluídos. Foi analisada a confiabilidade entre a primeira e a segunda medidas feitas por 2 operadores diferentes para rotação externa e interna, tanto do membro dominante quanto do não dominante. Foi utilizado o Coeficiente de Correlação Intraclasse (CCI) com seus respectivos intervalos de confiança de 95%. Para determinar o limiar de mudança real na força de handgrip quando as medições são registradas por diferentes avaliadores, a mínima diferença detectável (MDD) foi calculada.

Resultados: Excelente confiabilidade interobservador foi observada para rotação externa dominante e não dominante (CCI = 0,80 [IC 95% 0,58 - 0,92] e CCI 0,80 [IC 95% = 0,56 - 0,91], respectivamente). A confiabilidade entre avaliadores foi regular (ICC: 0,56 [IC 95% 0,18 - 0,80] e boa (ICC: 0,61 [IC 95% 0,25 - 0,82]) para rotação interna dominante e não dominante, respectivamente.

Conclusão: Este estudo mostrou que a confiabilidade interobservador de handgrip foi excelente para os rotadores externos e regular a boa para os rotadores internos.

Palavras-chave: reprodutibilidade dos testes, dinamômetro de força Muscular, ombro; manguito rotador, força muscular, confiabilidade.

Fuentes de financiamiento:

Las y los autores declaran no tener ninguna afiliación financiera ni participación en ninguna organización comercial que tenga un interés financiero directo en cualquier asunto incluido en este manuscrito.

Conflicto de intereses:

Las y los autores declaran no tener ningún conflicto de intereses.

Introducción

El dolor de hombro es una afección común dentro de la práctica kinésica diaria. El mismo es un problema médico y socioeconómico importante en la población occidental, que puede conducir a una incapacidad para llevar a cabo las actividades de la vida diaria, laborales y recreativas (1). En la literatura internacional representa el tercer motivo de consulta médica y se estima que entre el 18% y 26% de la población adulta tendrá dolor en dicha articulación en algún momento de la vida (2-4).

Si bien existen controversias con respecto a la precisión diagnóstica en estos pacientes, una característica común es la disminución de la fuerza o una alteración de la activación de los músculos del manguito rotador (5). Las lesiones en deportistas *overhead* se han relacionado con el desequilibrio muscular entre rotadores externos (Re) e internos (Ri) (6,7).

Debido a la importancia de la fuerza muscular, su medición periódica dentro de la evaluación kinésica es fundamental, ya sea para cuantificar la eficacia de un tratamiento, definir programas de prevención de lesiones, o guiar decisiones para la vuelta al deporte o a las actividades de la vida diaria luego de una lesión o cirugía (6,8).

Actualmente, el dinamómetro isocinético es el *gold standard* para evaluar de manera objetiva la fuerza muscular. El mismo presenta una alta confiabilidad inter e intra observador para demostrar los déficits musculares del manguito rotador, pero presenta un alto costo económico, es difícil de transportar por su gran tamaño y no se encuentra disponible en todos los centros de rehabilitación (6,9). La evaluación manual de la fuerza, una alternativa utilizada comúnmente en la práctica clínica diaria, presenta importantes limitaciones respecto a su confiabilidad (10, 11).

El dinamómetro *handgrip* (HG) es un sistema hidráulico que expresa los valores en kilogramos fuerza (Kg/F) (12).

Esta herramienta no se encuentra validada actualmente para la medición de la fuerza de rotadores de hombro. Entendiendo la importancia de valorar la fuerza muscular con una herramienta económica y accesible, y debido a la limitada confiabilidad que presenta la evaluación manual, el objetivo de este trabajo es evaluar la confiabilidad interobservador de la herramienta HG para cuantificar la fuerza isométrica de los Re y Ri de hombro.

Materiales y método

Este trabajo fue aprobado por el Comité de Ética e Investigación del Hospital y se reporta según los lineamientos recomendados por la literatura para el informe de estudios de confiabilidad (13).

Sujetos

El cálculo del tamaño muestral estimado fue de 21 sujetos para una confiabilidad predeterminada (p_0) de 0,30 y una confiabilidad esperada (p_1) de al menos 0,70, asumiendo un error alfa

de 0,05 y un error beta de 0,20 (14,15). Según la literatura un tamaño muestral de 20 es suficiente en estudios de confiabilidad (16, 17).

Mediante un muestreo por conveniencia se incluyeron 21 participantes voluntarios mayores de 18 años que concurren al Servicio de Kinesiólogía del Hospital D. F. Santojanni en el período comprendido entre octubre y diciembre de 2017. Todos los sujetos firmaron el consentimiento informado.

Fueron excluidos aquellos individuos que, mediante un cuestionario básico de cuatro preguntas, refirieron dolor, cirugías o lesiones previas de hombro, dolor o cirugías cervicales en los últimos 6 meses o que realizaban deportes *overhead* (18,19). Se eliminaron aquellos individuos que presentaron dolor durante las mediciones.

Se registraron los siguientes datos demográficos: edad, sexo, miembro superior dominante (determinado por la mano con la cual escribe) (20), actividad física y frecuencia de la actividad física según cantidad de horas semanales.

Evalúadores

Cuatro kinesiólogos de tres años de experiencia rotaron en grupos de dos para realizar las mediciones. Los mismos estaban familiarizados con la herramienta de medición. No obstante, previo al inicio del protocolo, realizaron un entrenamiento en la administración de la herramienta.

Procedimientos

El HG utilizado para la medición fue de la marca BASELINE® Standard 12-0240 (Hydraulic hand dynamometer), cuyos valores fueron expresados en kilogramos fuerza (Kg/F).

Los individuos realizaron 3 repeticiones de 5 segundos de esfuerzo voluntario, en la cual debían incrementar la fuerza gradualmente hasta llegar a su máximo con un descanso de 10 segundos entre cada intento. La contracción isométrica se realizó contra la pared para la Re y contra la fuerza aplicada por el investigador que sostiene el HG, en caso de Ri. (Figura 1) La posición de medición fue en decúbito supino con el miembro superior al costado del cuerpo con 90° de flexión de codo. La herramienta se posicionó a 2 cm por encima de la apófisis estiloides del radio sobre la región dorsal (fuerza de Re) o ventral (fuerza de Ri) según la medición correspondiente, y se adaptó el apoyo con goma espuma para mayor comodidad del individuo (9).

Para el análisis se consideró la repetición con el mayor valor alcanzado. Con el objetivo de minimizar el efecto de aprendizaje se aleatorizó el orden de inicio del grupo muscular evaluado (Re o Ri) (5). Para tal fin se utilizó la página de internet www.randomization.com. El procedimiento de medición comenzó siempre por el miembro superior dominante (D) y luego el no dominante (ND).

El orden de inicio de los evaluadores asignados también fue aleatorizado. Un primer evaluador realizó las primeras mediciones y luego de un intervalo de una hora, un segundo evaluador procedió con las segundas mediciones. Con el objetivo de evitar la lectura de las primeras mediciones por parte del segundo evaluador, los valores de fuerza se registraron en planillas independientes.



Figura 1. A. Posición de medición de fuerza de RE de hombro.
B. Posición de medición de fuerza de RI.

Análisis estadístico

Las variables continuas que asumieron una distribución normal se reportaron como media y desvío estándar (DE). De lo contrario, se utilizó la mediana y rango intercuartílico (RIQ). Las variables categóricas se reportaron como número de presentación y porcentaje.

Previamente a los análisis de confiabilidad, se deben evaluar determinados supuestos (21,22). El supuesto de normalidad de los datos fue evaluado mediante el test de Shapiro-Wilk. El supuesto de homocedasticidad se determinó según el procedimiento utilizado por Damron et al. (22,23). Primero, los datos se inspeccionaron visualmente trazando las diferencias entre dos mediciones y la media de ellas en un gráfico de Bland-Altman. Luego se evaluó formalmente mediante la correlación (R^2) entre las diferencias absolutas y los valores medios. Los valores de R^2 entre 0 y 0.1 se consideraron homocedásticos (24,25). Cuando se halló una distribución asimétrica o presencia de heteroscedasticidad, se intentó la corrección mediante la transformación logarítmica de los datos (24). Para comparar los valores entre la primera y segunda medición se utilizó la prueba t de Student para muestras pareadas o la prueba de rangos de Wilcoxon, según correspondiera. Se consideraron significativos aquellos valores $p < 0.05$. Para el análisis se utilizó el software IBM SPSS Macintosh, 24.0.

Confiabilidad entre evaluadores

Se realizó el cálculo de la confiabilidad relativa entre la primera y segunda medición realizada por 2 diferentes operadores utilizando el Coeficiente de Correlación Intraclase (CCI) con sus respectivos intervalos de confianza al 95%. Se utilizó un modelo de efectos aleatorios de una vía ($CCI_{1,3}$) (26,27).

Para todos los análisis los valores de los CCI se consideraron pobres (< 0.40), regulares (0.41 a 0.59), buenos (0.6 a 0.74), o excelentes (> 0.75) (19,26).

El acuerdo entre la primera y segunda medición se representó gráficamente mediante gráficos de Bland-Altman. Los límites de acuerdo del 95% se calcularon como la diferencia de medias $\pm 1,96$ DE de las diferencias (25,29).

Error estándar de medición y mínimo cambio detectable

El error de medición es el error que no es atribuible a verdaderos cambios en el constructo evaluado. Se determinó calculando el Error Estándar de Medición (EEMed) y el Mínimo Cambio Detectable (MCD). El EEMed representa el desvío estándar de las medidas repetidas en un mismo paciente y es calculado como la raíz cuadrada de la varianza del error del CCI (23, 30,31). Para determinar el EEMed se utilizó el CCI determinado entre la primera y segunda medición.

Para determinar el umbral de cambio real en la fuerza de HG cuando diferentes evaluadores registran las mediciones, se utilizó el MCD (32). El MCD representa el mínimo cambio observado en una herramienta que es real y no debido a un error de medición.

El MCD se calculó con la fórmula $1.96 \times \sqrt{2} \times \text{EEMed}$ con sus respectivos intervalos de confianza. Estos valores fueron expresados en Kg/F (23, 30,31).

Según lo recomendado, cuando el MCD se obtuvo a partir de una variable transformada al logaritmo, se realizó el antilogaritmo del MCD, dando como resultado un factor de multiplicación/división que puede interpretarse de nuevo a la escala original (24,29).

Resultados

Se incluyeron un total de 21 individuos sanos de los cuales uno fue eliminado por presentar dolor durante la medición. Las características de la muestra se reportan en la tabla 1.

Tabla 1. Características de la población	
Características	n=20
Sexo femenino, n (%)	14 (70)
Edad, M (RIQ)	28 (26 – 34,5)
Dominancia derecha, n (%)	17 (85)
Actividad física, n (%)	17 (85)
Actividad física \geq 3 h/s, n (%)	12 (60)
Referencias: h/s, horas por semana.	

En la tabla 2 se reportan los valores de confiabilidad entre evaluadores, calculados a partir de los datos transformados. Se observó una excelente confiabilidad entre evaluadores para la ReD y ReND (CCI: 0.80 [IC95% 0.58 - 0.92] y CCI: 0.80 [IC95% 0.56 - 0.91], respectivamente). La confiabilidad entre evaluadores para la RiD resultó regular (CCI: 0.56 [IC95% 0.18 - 0.80] y para

la RiND fue buena (CCI:0.61 [IC95% 0.25 - 0.82]). o resalta la necesidad de fortalecer acciones para mejorar el cumplimiento del Código en puntos de venta. Es importante señalar el alcance limitado de este estudio que podría replicarse en otras áreas del país para conocer si el CICSLM es acatado de distinta manera en otras regiones del país y garantizar su cumplimiento.

Tabla 2. Confiabilidad interevaluador			
Hand Grip (Kg Fuerza)	Medición 1	Medición 2	CCI (IC 95%) *
RE Miembro Dominante	11 (9,2 - 12,7)	11,5 (9,2 - 15,2)	0.80 (0.58 - 0.92)
RE Miembro No Dominante	12 (9.2 - 14)	10,5 (9,2 - 15)	0.80 (0.56 - 0.91)
RI Miembro Dominante	12 (10,25 - 15,7)	15,5 (13,2 - 19,5)	0.56 (0.18 - 0.80)
RI Miembro No Dominante	14 (10,2 - 17,5)	14 (12 - 19,2)	0.61 (0.25 - 0.82)
Referencias: RE: Rotación Externa; RI: Rotación Interna. Valores expresados como mediana y percentiles 25 y 75. CCI: Coeficiente de Correlación Intraclass Modelo 1,3; IC 95%: Intervalo de Confianza al 95%.			
(*) Valores calculados con un modelo de efectos aleatorios del Lognatural.			

Al comparar las medianas de la primera y segunda medición se observó diferencia estadísticamente significativa en la RiD (12 (10,25 - 15,7) vs 15,5 (13,2 - 19,5); p=0,004).

Previo a la confección de los correspondientes gráficos de Bland-Altman se evaluó el supuesto de normalidad de las diferencias entre la primera y segunda medición y el supuesto de heterocedasticidad. El test de Shapiro Wilk resultó significativo para la ReD. La inspección visual mostró signos de heterocedas-

ticidad en la variable ReD que fue confirmado por el coeficiente $R^2 = -0.45$. Los coeficientes R^2 para la ReD, ReND y RiND fueron 0,022, 0,006 y 0,049, respectivamente. La transformación de los datos no corrigió la heterocedasticidad ni la asimetría de la distribución de la ReD por lo que se optó por representar gráficamente dicha variable en su escala original. En la figura 2 se reportan los correspondientes gráficos de Bland Altman para los cuatro parámetros medidos.

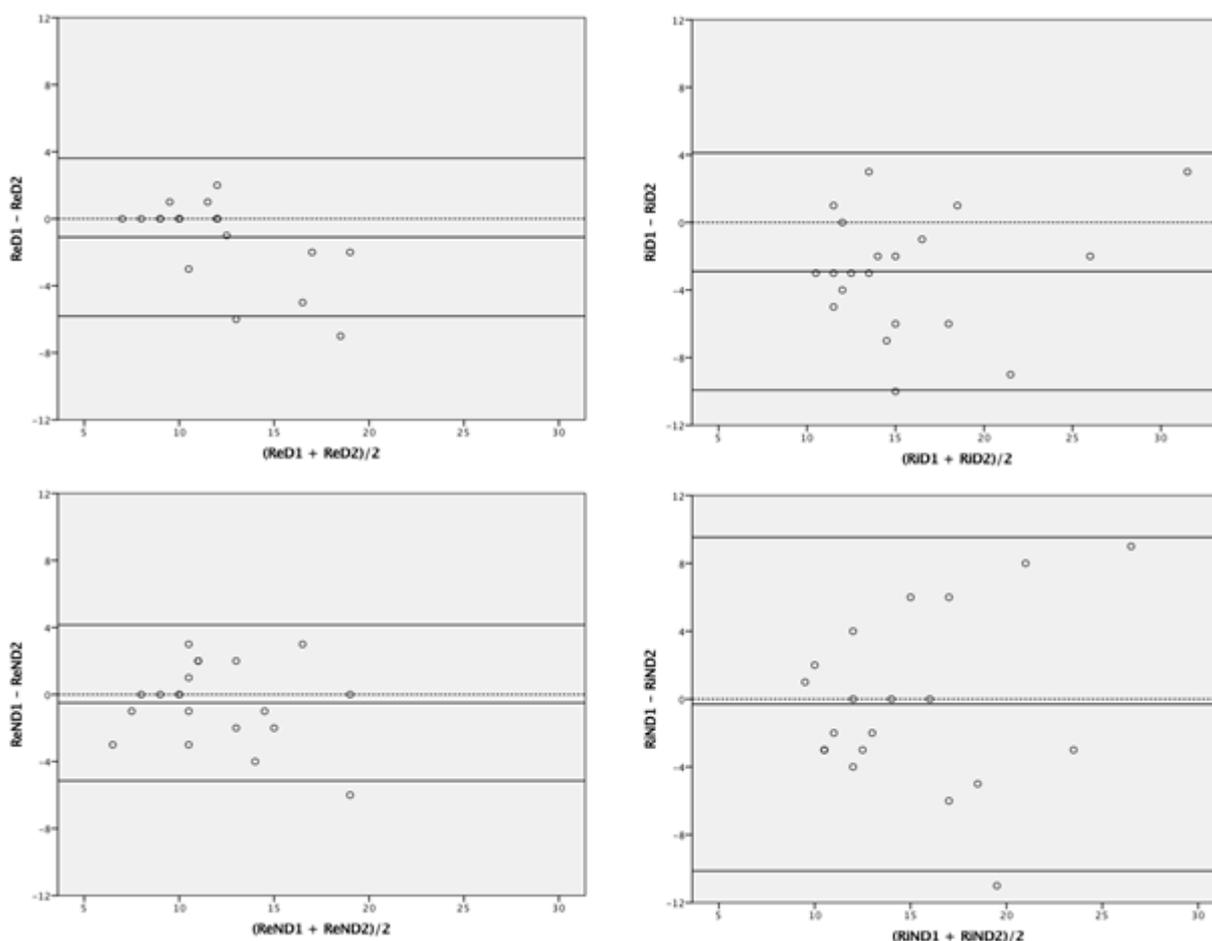


Figura 2. Gráfico de Bland-Altman. La línea horizontal discontinua es la diferencia promedio entre mediciones (kg/fuerza), y las líneas continuas representan los límites de acuerdo del 95%.

 Abreviación: ReD: rotación externa dominante, RiD: rotación interna dominante, ReND: rotación externa no dominante, RiND: rotación interna no dominante.

En la tabla 3 se reportan los correspondientes EEmed y MCD. Se observaron valores similares tanto para la rotación externa como

para la rotación interna. La interpretación de los MCD calculados mediante la transformación logarítmica se explica en la discusión.

Tabla 3. Error estándar de medición y cambio mínimo detectable			
Medición	EEmed (Log¹⁰)	MCD (Log¹⁰)	MCD
HG RE Miembro No Dominante	0,123	0,34	x/+ 1.41
HG RI Miembro Dominante	0,128	0,36	x/+ 1.43
HG RI Miembro No Dominante	0,130	0,36	x/+ 1.43
HG RE Miembro Dominante	0,306	0,85	x/+ 2.34
HG RE Miembro No Dominante	0,26	0,73	x/+ 2.08
HG RI Miembro Dominante	0,359	0,99	x/+ 2.70
HG RI Miembro No Dominante	0,289	0,80	x/+ 2.23

Referencias: Eemed: Error Estándar de medición; MCD: Mínimo Cambio Detectable. RE: Rotación Externa; RI: Rotación Interna; (*) Valores calculados a partir del antilog natural. Todos los valores son expresados en Kg Fuerza.

Discusión

El principal resultado de nuestro estudio fue la aceptable confiabilidad entre evaluadores en la medición de la fuerza isométrica de los rotadores de hombro.

Diversos estudios encontraron una correlación de buena a excelente en la evaluación de fuerza isométrica de rotadores internos y externos de hombro. (9,33,34). En el estudio de Cadogan et al., la confiabilidad entre evaluadores de la Re fue moderada (35). Sin embargo, la comparación con estos resultados es limitada, ya que la herramienta utilizada en nuestro estudio no fue la misma, como tampoco lo fueron las posiciones de evaluación y la población evaluada (9,33,34,35).

Los MCDs obtenidos a partir de la transformación logarítmica se pueden transferir directamente a cualquier estudio que realice una transformación del logaritmo natural. Además, los factores antilogaritmo de multiplicación/división que se muestran en la tabla 3 pueden proporcionar información relevante. Teniendo en cuenta esto, los factores antilogarítmicos podrían aplicarse a las mediciones iniciales de un paciente con el fin de anticipar el grado en que la fuerza tendrá que cambiar en un individuo para representar un cambio verdadero y no atribuible al error por haber sido evaluado por diferentes evaluadores. Por ejemplo, para un individuo que al inicio presenta una fuerza media de HG de 10 Kg/F de ReD, nuestro factor antilogarítmico de 1.41 (ver tabla 3) informa que, en la próxima medición, la fuerza de HG debería aumentar a 14.1 Kg/F o disminuir a 7,1 Kg/F, para poder decir con un 95% de confianza que el cambio fue real.

Otra alternativa es expresar los valores de seguimiento en porcentaje de los valores basales. Si la relación es mayor que el factor antilogarítmico de 1.41, indica una mejora de la fuerza y si es menor a 0.71, representa una disminución de la fuerza. Por ejemplo, para un sujeto con una fuerza inicial de 15 Kg/F en la ReD y que se incrementa a 20 Kg/F al seguimiento, la relación 20/15 es igual a 1.33 y no alcanza el MCD de 1.41, por lo que no se puede asumir que la mejora sea real y podría ser atribuible a un error de medición por haber utilizado diferentes evaluadores. En otro ejemplo, si el paciente empeora su fuerza de 10 a 6 Kg/Fza, la relación 6/10 es igual a 0.6 y es inferior al MCD de 0.71, por lo que se puede asumir que la disminución de la fuerza es real y no atribuible a un error de medición.

Este trabajo cuenta con limitaciones. En primer lugar, es de destacar los amplios Intervalos de Confianza como así también los elevados EEm y MCD. Una posible explicación podría ser que estas estimaciones son directamente proporcionales a la variabilidad (error) e inversamente proporcionales al tamaño muestral.

Por tal motivo, incrementar la cantidad de participantes podría haber mejorado la precisión de estas estimaciones. No obstante, el tamaño de nuestra muestra fue el estimado según fórmula de tamaño muestral, similar al utilizado en otros estudios, y el mínimo sugerido para estudios de confiabilidad (16,17). Otra explicación podría ser la relacionada a las características clínicas inexploradas entre los participantes reclutados, el haber utilizado diferentes evaluadores, un tiempo de pausa posiblemente insuficiente entre mediciones, y errores durante los procedimientos de medición. En segundo lugar, podemos mencionar que la evaluación fue realizada en individuos sanos por lo cual la confiabilidad obtenida podría no ser extrapolable en pacientes sintomáticos. En tercer lugar, al comparar las medianas entre la primera y segunda medición, hemos observado diferencias estadísticamente significativas en la RiD. Este incremento durante la segunda evaluación de la fuerza de RiD no sugiere relevancia clínica y no consideramos que haya habido un efecto de arrastre. No obstante, no se puede desestimar un posible efecto de aprendizaje o acondicionamiento entre la primera y segunda medición.

Este es, a nuestro entender, el primer estudio que brinda un análisis de la confiabilidad de la medición de fuerza de Re y Ri de hombro con la herramienta HG. La excelente confiabilidad para la medición de la fuerza de Re podría ser alentadora para incluirla dentro de la evaluación clínica. Por este motivo consideramos que este método de evaluación podría introducirse en la práctica debido a su mayor confiabilidad respecto de las pruebas manuales subjetivas.

Cabe destacar, que la medida del cambio mínimo detectable obtenido en la medición de fuerza de los rotadores de hombro podría ser utilizada como parámetro de cambio clínico de la misma en la evaluación de un programa de rehabilitación. Estos resultados justifican que futuros estudios evalúen la validez de esta herramienta y la confiabilidad en pacientes con afecciones de hombro.

Conclusión

El estudio mostró que la confiabilidad interobservador de la herramienta HG para medir fuerza de hombro es excelente para los rotadores externos y de regular a buena para los rotadores internos.

Contribución de los/as autores: PCD, GPC, SGS y MAG: búsqueda bibliográfica, recolección de datos, diseño del estudio, análisis de los datos, preparación del manuscrito, revisión del manuscrito. MFA: búsqueda bibliográfica, recolección de datos, diseño del estudio, análisis estadístico, análisis de los datos, revisión del manuscrito.

Referencias

1. Linaker CH, Walker-Bone K. Shoulder disorders and occupation. *Best Pract Res Clin Rheumatol* 2015 Jun;29(3):405-23. DOI: 10.1016/j.berh.2015.04.001
2. Greving K, Dorrestijn O, Winters JC, Groenhof F, van der Meer K, Stevens M, et al. Incidence, prevalence, and consultation rates of shoulder complaints in general practice. *Scand J Rheumatol.* 2012 Mar;41(2):150-5. DOI: 10.3109/03009742.2011.605390
3. Luime JJ, Koes BW, Hendriksen IJ, Burdorf A, Verhagen AP, Miedema HS, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. *Scand J Rheumatol* 2004;33(2):73-81.
4. Muhafara, G. J., Villarruel, M., Cura, A., Tozzi, W., Vicente, A., Esperón, L. F., Salzberg, S., Soliño, S., Navarro, E., Argento, F., & Policastro, P. (2021). Estudio epidemiológico de afecciones musculoesqueléticas de hombro en la Ciudad de Buenos Aires. *Argentinian Journal of Respiratory & Physical Therapy*, 3(2), 13-21
5. Cools AM, Michener LA. Shoulder pain: can one label satisfy everyone and everything? *Br J Sports Med.* 20: 16 Mar;51(5):416-417. DOI: 10.1136/bjsports-2016-096772
6. Cools AM, Vanderstukken F, Vereecken F, Duprez M, Heyman K, Goethals N, et al. Eccentric and isometric shoulder rotator cuff strength testing using a hand-held dynamometer: reference values for overhead athletes. *Knee Surg Sports Traumatol Arthrosc.* 2016 Dec;24(12):3838-3847. DOI: 10.1007/s00167-015-3755-9
7. Warner JJ, Micheli LJ, Arslanian LE, Kennedy J, Kennedy R. Patterns of flexibility, laxity, and strength in normal shoulders and shoulders with instability and impingement. *Am J Sports Med.* 1990 Jul-Aug; 18(4):366-75. DOI: 10.1177/036354659001800406
8. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. *PM R* 2011; 3:472-9. DOI: 10.1016/j.pmrj.2010.10.025
9. Cools AM, De Wilde L, Van Tongel A, Ceysens C, Ryckewaert R, Cambier DC. Measuring shoulder external and internal rotation strength and range of motion: comprehensive intra-rater and inter-rater reliability study of several testing protocols. *J Shoulder Elbow Surg.* 2014 Oct;23(10):1454-61. DOI: 10.1016/j.jse.2014.01.006
10. Hayes K, Walton JR, Szomor ZL, Murrell GA. Reliability of 3 methods for assessing shoulder strength. *J Shoulder Elbow Surg.* 2002 Jan-Feb;11(1):33-9. DOI: 10.1067/mse.2002.119852
11. Nagatomi T, Mae T, Nagafuchi T, Yamada SI, Nagai K, Yoneda M. Shoulder manual muscle resistance test cannot fully detect muscle weakness. *Knee Surg Sports Traumatol Arthrosc* 2017 Jul;25(7):2081-2088. DOI: 10.1007/s00167-016-4380-y
12. Innes E. Handgrip Strength testing: a review of the literature. *Australian Occupational Therapy Journal.* 1999; 46 (3): 120-140.
13. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Int J Nurs Stud.* 2011 Jun;48(6):661-71
14. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. *Stat Methods Med Res* 2004; 13(4):251-271
15. Terwee CB, Bot SD, De Boer MR, Van Der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007; 60(1):34-42
16. Fleiss JL, Reliability of measurement In: *The Design and Analysis of Clinical Experiments.* 2011. New York: Wiley; 1-32. DOI:10.1002/9781118032923
17. Hobart JC, Cano SJ, Warner TT, Thompson AJ. What sample sizes for reliability and validity studies in neurology? *J Neurol.* 2012 Dec;259(12):2681-94.
18. Johansson FR, Skillgate E, Lapauw ML, Clijmans D, Deneulin VP, Palmans T, et al. Measuring Eccentric Strength of the Shoulder External Rotators Using a Handheld Dynamometer: Reliability and Validity. *J Athl Train.* 2015 Jul;50(7):719-25
19. Riemann BL, Davies GJ, Ludwig L, Gardenhour H. Hand-held dynamometer testing of the internal and external rotator musculature based on selected positions to establish normative data and unilateral ratios. *J Shoulder Elbow Surg* 2010 Dec; 19(8):1175-83. DOI: 10.1016/j.jse.2010.05.021
20. Horsley I, Herrington L, Hoyle R, Prescott E, Bellamy N. Do changes in hand grip strength correlate with shoulder rotator cuff function? *Shoulder Elbow.* 2016 Apr;8(2):124-9. DOI: 10.1177/1758573215626103
21. Portney LG, Watkins MP. *Foundations of clinical research applications to practice.* 3rd ed. Upper Saddle River (New Jersey): Pearson Education Inc.; 2009.

22. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; 19:231–40.
23. Damron LA, Dearth DJ, Hoffman RL, Clark BC. Quantification of the corticospinal silent period evoked via transcranial magnetic stimulation. *J Neurosci Methods*. 2008 Aug 15;173(1):121-8
24. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med*. 1998 Oct;26(4):217-38
25. Bland JM, Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Comput Biol Med*. 1990;20(5):337-40.
26. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol*. 2012;8(1):23–34.
27. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*. 1979; 86(2):420–428. DOI: 10.1037/0033-2909.86.2.420
28. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. 1994; 6(4):284–290. DOI: 10.1037/1040-3590.6.4.284
29. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307–10.
30. De Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006 Oct;59(10):1033-9
31. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010 Jul;63(7):737-45
32. Remigio W, Tsai N, Layos L, Chavez M. Inter-rater and intra-rater reliability of the fluid goniometer for measuring active knee flexion in painful knees; correlations do not mean agreement. *J Phys Ther Sci*. 2017 Jun;29(6):984-988
33. Chen B, Liu L, Bin Chen L, Cao X, Han P, Wang C, Qi Q. Concurrent Validity and Reliability of a Handheld Dynamometer in Measuring Isometric Shoulder Rotational Strength. *J Sport Rehabil*. 2021 Jan 19;30(6):965-968. DOI: 10.1123/jsr.2020-0021.
34. Fieseler G, Molitor T, Irlenbusch L, Delank KS, Laudner KG, Hermassi S, Schwesig R. Intrarater reliability of goniometry and hand-held dynamometry for shoulder and elbow examinations in female team handball athletes and asymptomatic volunteers. *Arch Orthop Trauma Surg*. 2015 Dec;135(12):1719-26. DOI: 10.1007/s00402-015-2331-6.
35. Cadogan A, Laslett M, Hing W, McNair P, Williams M. Reliability of a new hand-held dynamometer in measuring shoulder range of motion and strength. *Man Ther*. 2011 Feb;16(1):97-101. DOI: 10.1016/j.math.2010.05.005