Concordancia según el número de intentos en la evaluación de la altura, velocidad y vuelo en la prueba de salto vertical

Agreement according to the number of trials in the assessment of height, velocity, and flight time in the vertical jump test

Autores

Karen Julieth Quintero Diaz [1, 3.]
Oscar A. Gutiérrez-Lesmes [2]
Javier Martínez Torres [3]
Mónica Vannesa Chanagá Gelves [1]
Sergio Alejandro Quijano Duarte [4]

[1] Escuela de Cuidado, Universidad de los Llanos (Colombia)
[2] Escuela de Salud Pública, Universidad de los Llanos (Colombia)
[3] Grupo FAVISA, Universidad de los Llanos (Colombia).
[4] Docente Escuela de Ciencias de la Salud (ECISA) de la Universidad Nacional Abierta y a Distancia (UNAD), Colombia.

Autor de correspondencia:
Karen Julieth Quintero Diaz
karen.quintero.diaz@unillanos.edu.co

Recibido: 19-10-26
Aceptado: 06-11-26

Cómo citar en APA

Quintero Díaz, K. J., Gutiérrez Lesmes, O. A., Martínez Torres, J., Chanaga Gelves, M. V., & Quijano Duarte, S. A. (2026). Concordancia según el número de intentos en la evaluación de la altura, velocidad y vuelo en la prueba de salto vertical. *Retos, 85*, 329-337. https://doi.org/10.47197/retos.v85.120210

Resumen

Introducción: La prueba de salto vertical es un marcador clínico y fisiológico fundamental para evaluar la capacidad muscular de los miembros inferiores.
Objetivo: evaluar la fiabilidad y concordancia según el número de intentos en la estimación de la altura, velocidad y tiempo de vuelo del salto vertical.
Métodos: Estudio transversal analítico en 74 adultos jóvenes físicamente inactivos de 18 a 25 años (45 mujeres y 29 hombres). Cada participante ejecutó tres saltos verticales sobre una plataforma Axon Jump con intervalos de tres minutos de reposo. Se comparó el primer intento frente al promedio de dos y tres intentos. La fiabilidad relativa se determinó mediante el coeficiente de correlación intraclase (CCI), la fiabilidad absoluta mediante el error estándar de la medición (%EEM) y la concordancia a través del método de Bland-Altman.
Resultados: El rendimiento cinemático mostró una fiabilidad relativa excelente en todas las métricas analizadas, con valores de CCI superiores a 0,990; para altura (0,992), tiempo de vuelo (0,991) y velocidad (0,992). La fiabilidad absoluta confirmó un margen de error mínimo (%EEM menor al 3,7%). El análisis de Bland-Altman evidenció sesgos sistemáticos cercanos a cero y límites de acuerdo estrechos.
Conclusión: Un único intento y el promedio de múltiples mediciones presentan alta concordancia operativa. No obstante, en investigación se recomienda implementar modelos jerárquicos de medidas repetidas en lugar de promedios simples para capturar adecuadamente la variabilidad intraindividual.

Palabras clave

Adulto joven; capacidades físicas; rendimiento físico; reproducibilidad de los resultados.

Abstract

Introduction: The vertical jump test is a fundamental clinical and physiological marker for assessing lower-limb muscular capacity.
Objective: To assess the reliability and agreement according to the number of trials in estimating vertical jump height, velocity, and flight time.
Methods: An analytical cross-sectional study was conducted in 74 physically inactive young adults (aged 18 to 25 years; 45 females and 29 males). Each participant performed three vertical jumps on an Axon Jump contact mat, separated by three-minute rest intervals. The first trial was compared against the mean of two and three trials. Relative reliability was determined using the intraclass correlation coefficient (ICC), absolute reliability via the percentage standard error of measurement (%SEM), and agreement through the Bland–Altman method.
Results: Kinematic performance demonstrated excellent relative reliability across all analyzed metrics, with ICC values exceeding 0.990; for height (0.992), flight time (0.991), and velocity (0.992). Absolute reliability confirmed a minimal margin of error (%SEM < 3.7%). The Bland–Altman analysis revealed systematic biases close to zero and narrow limits of agreement.
Conclusion: A single trial and the mean of multiple measurements exhibit high operational agreement. Nevertheless, in research settings, implementing hierarchical repeated-measures models rather than simple averages is recommended to adequately capture intra-individual variability.

Keywords

Young adult; physical fitness; physical performance; reproducibility of results.

Introducción

La capacidad muscular se define como la funcionalidad de un grupo muscular específico para generar tensión mecánica y torque, orientándose ya sea a sostener contracciones repetidas y prolongadas en el tiempo, como ocurre en la resistencia muscular, o a ejecutar una contracción máxima y dinámica en un intervalo breve, característica de la fuerza explosiva (Bohannon, 2015; Ruiz et al., 2009).

Diversas investigaciones han demostrado una asociación directamente proporcional entre el incremento de la potencia muscular en los miembros inferiores y la mejora en los indicadores de salud ósea (García-Hermoso et al., 2019), salud metabólica y cardiovascular (de Lima et al., 2022). Adicionalmente, se ha documentado una relación inversamente proporcional entre la capacidad muscular de los miembros inferiores y diversos marcadores de adiposidad corporal (García-Hermoso et al., 2019).

Con el propósito de evaluar la capacidad muscular de miembros inferiores, se han implementado múltiples metodologías, caracterizadas por sus particulares ventajas y restricciones, las cuales van desde cinturones de salto, plataformas de fuerza convencionales hasta herramientas emergentes como aplicaciones móviles y dispositivos dotados de sensores de última tecnología (Vivas et al., 2016). En la actualidad, la prueba de salto vertical ha adquirido especial relevancia debido a su doble utilidad: por un lado, funciona como un parámetro de monitorización sensible para el seguimiento de procesos de entrenamiento y de rehabilitación clínica; por otro, se consolida como un marcador fiable de la capacidad física y la salud biológica general del individuo (García-Hermoso et al., 2018; Wang et al., 2020).

No obstante, más allá de la instrumentación empleada para registrar el rendimiento en el salto vertical, el rigor metodológico exige la realización de pruebas y mediciones seriadas en cada sujeto evaluado (Cagua Ardila et al., 2023) con el fin de minimizar la influencia de la aleatoriedad, el sesgo de medición y la variabilidad interindividual propia de la muestra. El protocolo exige habitualmente la realización de repeticiones seriadas (de Vet et al., 2011). Bajo este marco metodológico, los protocolos de evaluación suelen estipular la ejecución y documentación de al menos dos mediciones consecutivas en cada sesión de prueba (Martinez-Torres et al., 2025; Martínez-Torres et al., 2022). Habitualmente, la estimación del parámetro se determina a partir del promedio de los intentos registrados (de Vet et al., 2011). Asimismo, de manera alternativa se suele emplear el valor máximo registrado, reflejando con ello la máxima expresión de la capacidad neuromuscular del individuo (Cruz-Jentoft et al., 2019).

A pesar de la amplia aceptación de estas prácticas en los contextos clínico, deportivo y científico, persiste un vacío en la literatura respecto al análisis riguroso de la discrepancia observable entre mediciones sucesivas realizadas dentro de una misma sesión de evaluación (Attia et al., 2017; de Vet et al., 2011). Consecuentemente, esta carencia metodológica dificulta la correcta estimación de la consistencia intraindividual, y la justificación teórica de los criterios de selección para el puntaje final (de Vet et al., 2011). Profundizar en el estudio de estas diferencias permite optimizar los procedimientos de evaluación, incrementar la precisión en la estimación del rendimiento y garantizar una mayor comparabilidad metodológica entre las investigaciones del área.

Debido a lo descrito anteriormente, este trabajo pretende evaluar la concordancia según el número de intentos en la evaluación de la altura, velocidad y vuelo en la prueba de salto vertical.

Método

Participantes

Se realizó un estudio transversal analítico. Se incluyeron 74 voluntarios, adultos jóvenes (18 a 25 años) clasificados como físicamente inactivos según los criterios descritos por el Colegio Americano de Medicina del Deporte (Riebe et al., 2018). Se excluyó a los participantes que consumían sustancias psicoactivas o fármacos neuroactivos (anticonvulsivos, reguladores del sueño o medicamentos para el sistema nervioso central), así como a aquellos con un índice de masa corporal inferior a 20 kg/m^2 o superior a 28 kg/m^2. De igual forma, no fueron elegibles personas con diagnóstico de epilepsia, dolor musculoesquelético agudo o neuropático, antecedentes de tratamiento para la depresión, o lesiones en los miembros inferiores en los últimos seis meses.

La investigación se consideró de bajo riesgo para los participantes y se implementaron medidas orientadas a garantizar su seguridad durante las evaluaciones. Todos los procedimientos fueron supervisados por un fisioterapeuta entrenado, con el propósito de asegurar la precisión y estandarización de las mediciones. Antes de participar, todos los voluntarios fueron informados detalladamente sobre los objetivos, procedimientos del estudio y firmaron el consentimiento informado, garantizándose su derecho a retirarse en cualquier momento sin repercusiones (Martinez-Torres & Gutiérrez-Lesmes, 2017). La investigación contó con concepto técnico favorable del Comité de Bioética de la Universidad de los Llanos, según el Acta 005 del 3 de mayo de 2024, comunicado mediante Memorando 3110.02.0179 del 6 de mayo de 2024.

Procedimiento

Para la recolección de información, se diseñaron formatos de registro que permitieron documentar los datos correspondientes a las características clínicas y las variables de interés del estudio.

Instrumentos

Evaluación del Salto vertical

Para la evaluación del salto vertical, los participantes realizaron una prueba de salto vertical en la plataforma de salto Axon Jump (Azócar-Gallardo et al., 2022). Previamente, efectuaron ejercicios de activación, que incluyeron movilidad activa de miembros inferiores. Adicionalmente, recibieron instrucciones detalladas sobre la ejecución de la prueba. Se les indicó adoptar una posición inicial de pie, con los pies separados a la anchura de los hombros, los brazos relajados a los lados del cuerpo y las rodillas y caderas flexionadas aproximadamente a 90°. Desde esta posición, los participantes realizaron un movimiento explosivo hacia arriba, extendiendo simultáneamente las rodillas y las caderas para generar impulso vertical. Durante el salto, se les solicitó alcanzar la máxima altura posible mediante la extensión completa del cuerpo, evitando levantar los talones antes del despegue. En el aterrizaje, debían apoyar ambos pies de manera controlada y amortiguar el impacto mediante la flexión de las rodillas.

Cada participante realizó tres intentos de salto, con intervalos de recuperación de tres minutos entre cada ejecución. Se registraron tres indicadores para el análisis: el resultado del primer salto, el promedio de los dos primeros saltos y el promedio de los tres saltos. Estos indicadores se utilizaron para analizar las variables de velocidad, tiempo de vuelo y altura del salto.

Evaluación de las características clínicas

La masa corporal, masa grasa corporal, masa libre de grasa, porcentaje de agua corporal, masa de agua corporal, masa muscular, masa ósea, frecuencia cardíaca en reposo fueron evaluadas por medio de una báscula Tanita BC1500 pro. Antes de la medición, se le pidió al individuo que subiera a la báscula sin zapatos, con la menor cantidad de ropa posible, sin accesorios tales como correas o bolsos, sin ningún contenido en los bolsillos y sin accesorios en la cabeza. La estatura fue evaluada por autorreporte, el cual se considera una metodología válida para población universitaria (Martínez Torres et al., 2013). Con los datos de la masa y la estatura se calculó el índice de masa corporal.

Análisis de datos

Inicialmente, se desarrolló un análisis exploratorio de los datos. Posteriormente, se realizó una descripción bivariable de las características clínicas y antropométricas; la descripción de las variables cualitativas se llevó a cabo mediante frecuencias absolutas y porcentuales, mientras que las variables cuantitativas se expresaron como medias y desviaciones estándar. Para estimar la fiabilidad test-retest, se calculó el coeficiente de correlación intraclase (CCI) mediante un modelo de efectos aleatorios de dos vías con acuerdo absoluto para mediciones individuales, y se estimaron el error estándar de la medición (EEM) y el porcentaje del error estándar de la medición (EEM%) (Weir, 2005), finalmente, se empleó el gráfico de Bland-Altman (Bland et al., 2012).

Resultados

Los voluntarios presentaron un índice de masa corporal con valores equivalentes entre ambos sexos (23,36 vs. 23,48 kg/m^2) pese a las diferencias en estatura y peso. En cuanto a la composición corporal, los hombres tienden a mostrar mayores valores en los indicadores libres de grasa (masa libre de grasa, muscular, ósea y agua corporal), mientras que las mujeres tienden a registrar una mayor masa grasa absoluta (18,34 vs. 15,30 kg). La frecuencia cardíaca media en reposo fue de 86,64 ± 13,86 lpm. Los demás datos sobre características sociodemográficas y antropométricas de los participantes se muestran en la tabla 1.

Tabla 1. Características sociodemográficas y antropométricas de los participantes.

Característica	Total n = 74 X̄ (SD)	Femenino n = 45 X̄ (SD)	Masculino n = 29 X̄ (SD)
Estatura (cm)	165,81 (9,86)	160,10 (6,17)	174,66 (7,75)
Peso (kg)	65,50 (12,74)	59,95 (10,86)	74,11 (10,57)
Índice de masa corporal (kg/m2)	23,41 (4,43)	23,36 (3,56)	23,48 (5,59)
Masa grasa corporal (kg)	17,15 (6,36)	18,34 (6,41)	15,30 (5,94)
Masa libre de grasa (kg)	48,06 (9,89)	41,84 (5,06)	57,71 (7,51)
Porcentaje de agua corporal (%)	55,29 (5,91)	52,54 (4,79)	59,55 (4,91)
Masa de agua corporal (kg)	36,19 (7,69)	31,19 (3,18)	43,93 (6,00)
Masa muscular (kg)	46,05 (9,90)	39,70 (4,82)	55,90 (7,35)
Masa ósea (kg)	2,46 (0,49)	2,14 (0,24)	2,94 (0,35)
Frecuencia cardíaca en reposo (lpm)	86,64 (13,86)	87,58 (13,59)	85,26 (14,39)
Altura un intento (cm)	28,17 (9,11)	22,77 (4,37)	36,56 (8,16)
Altura dos intentos (cm)	28,04 (8,86)	22,84 (4,18)	36,11 (8,13)
Altura tres intentos (cm)	28,10 (8,82)	22,93 (4,23)	36,12 (8,05)
Vuelo un intento (ms)	473,41 (75,67)	428,98 (41,25)	542,34 (64,49)
Vuelo dos intentos (ms)	473,18 (74,84)	429,70 (39,28)	540,64 (66,53)
Vuelo tres intentos (ms)	473,46 (73,91)	430,47 (40,00)	540,17 (64,41)
Velocidad un intento (m/s)	2,32 (0,37)	2,10 (0,20)	2,66 (0,32)
Velocidad dos intentos (m/s)	2,32 (0,36)	2,11 (0,19)	2,64 (0,32)
Velocidad tres intentos (m/s)	2,32 (0,36)	2,11 (0,20)	2,65 (0,31)

Nota: X̄: promedio; SD: Desviación estándar.

La tabla 2, muestra valores de ICC superiores a 0,985 que indican una sobresaliente confiabilidad relativa. La confiabilidad absoluta confirmó un margen de error instrumental mínimo, con un %SEM consistentemente inferior al 4% en todas las métricas analizadas. Asimismo, el análisis de Bland-Altman mostró un sesgo sistemático cercano a cero y límites de acuerdo estrechos. Las figuras 1, 2 y 3 presentan los resultados del análisis de Bland-Altman.

Tabla 2. Confiabilidad y concordancia en las evaluaciones de altura, velocidad y tiempo de vuelo mediante uno, dos y tres intentos de salto.

	Concordancia			Confiabilidad Absoluta		Confiabilidad Relativa
Altura	Sesgo	SD	LI; LS	EEM (IC95%)	%EEM (IC95%)	ICC (IC95%)
Un intento vs dos intentos	0,14	1,11	(-2,04; 2,31)	0,78 (0,68-0,94)	2,79 (2,40-3,33)	
Un intento vs tres intentos	0,08	1,46	(-2,79; 2,94)	1,03 (0,89-1,23)	3,67 (3,16-4,38)	0,992 (0,988-0,995)
Dos intentos vs tres intentos	-0,06	0,66	(-1,35; 1,23)	0,46 (0,40-0,55)	1,65 (1,42-1,97)	
Vuelo						
Un intento vs dos intentos	0,23	10,94	(-21,22; 21,68)	7,74 (6,66-9,23)	1,64 (1,41-1,95)	
Un intento vs tres intentos	-0,05	12,27	(-24,10; 23,99)	8,68 (7,47-10,35)	1,83 (1,58-2,19)	0,991 (0,987-0,994)
Dos intentos vs tres intentos	-0,28	6,10	(-12,25; 11,68)	4,32 (3,72-5,15)	0,91 (0,78-1,09)	
Velocidad						
Un intento vs dos intentos	0,00	0,05	(-0,09 0,10)	0,03 (0,03-0,04)	1,43 (1,23-1,71)	
Un intento vs tres intentos	0,00	0,06	(-0,12 0,12)	0,04 (0,04-0,05)	1,81 (1,56-2,16)	0,992 (0,988-0,994)
Dos intentos vs tres intentos	0,00	0,03	(-0,06 0,06)	0,02 (0,02-0,03)	0,93 (0,80-1,11)	

Nota: SD: Desviación estándar; IC95%: Intervalo de Confianza del 95%, LI: Límite Inferior; LS: Límite Superior.

Figura 1. Bland-Altman de comparación entre las alturas estimadas con las diferentes cantidades de intentos.

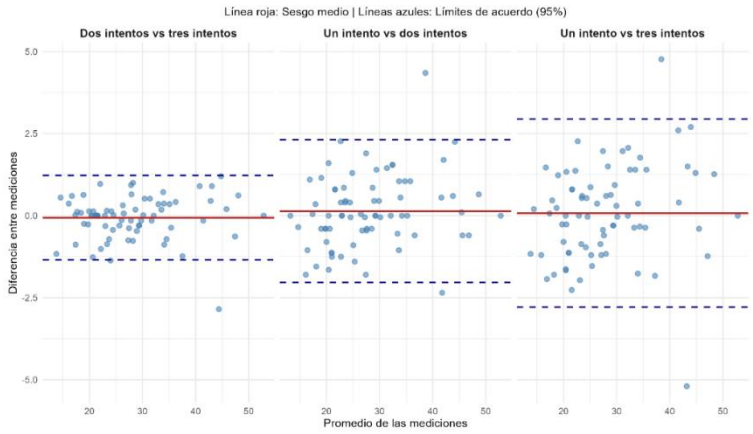


Figura 2. Bland-Altman de comparación entre los vuelos estimados con las diferentes cantidades de intentos.

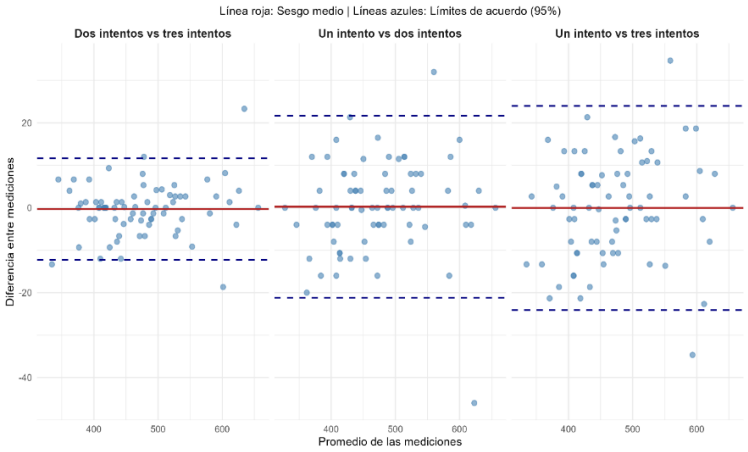
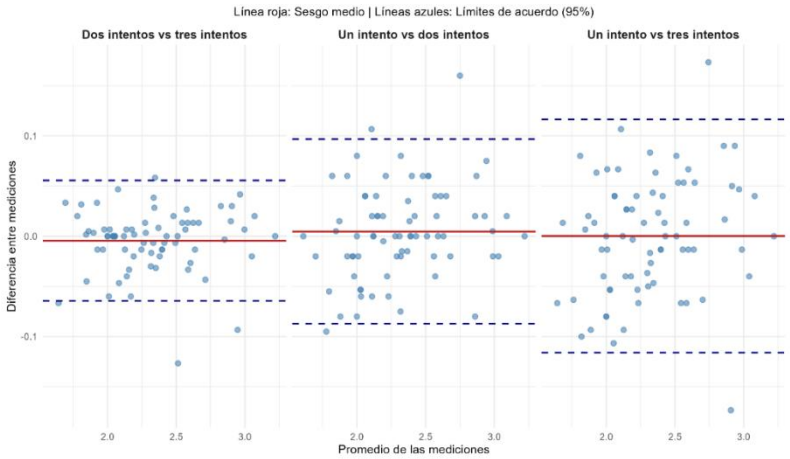


Figura 3. Bland-Altman de comparación entre las velocidades estimadas con las diferentes cantidades de intentos.



Discusión

El hallazgo central de esta investigación evidencia que la cuantificación de la altura, el tiempo de vuelo y la velocidad mediante un solo intento de salto suelen mostrar concordancia y confiabilidad operativa en la práctica con el registro del promedio de dos o tres intentos. Esta consistencia se refleja en coeficientes de correlación intraclase casi perfectos, errores estándar de medición reducidos y diferencias medias prácticamente nulas entre los protocolos evaluados.

Estos resultados son similares a los descritos previamente, donde en pruebas de salto y potencia, la fiabilidad es generalmente alta, aunque se ha descrito que los efectos de aprendizaje pueden influir en los resultados iniciales (Petré et al., 2023; Venegas-Carro et al., 2022). Por ejemplo, en atletas de élite, no se encontraron diferencias significativas entre reportar el mejor intento, el promedio de los dos mejores o el promedio de tres intentos (Petré et al., 2023). No obstante, para estudios longitudinales, el promedio de tres intentos suele producir resultados más consistentes y menor error (Venegas-Carro et al., 2022).

Con respecto a otros marcadores de las capacidades físicas como la fuerza prensil manual, se ha descrito que en adultos sanos el reporte del promedio de tres intentos o del mejor resultado son fiables (Venegas-Carro et al., 2022). Sin embargo, en individuos con discapacidad intelectual, el promedio de tres intentos es superior al mejor resultado, mostrando mayor fiabilidad (ICC=0,96 vs 0,94) y menor error de medición (SEM%=8,2-8,7% vs 9,2-10,1%) (Karatrantou et al., 2020).

Algunos trabajos han descrito que se debe utilizar el mejor intento, debido a que es considerado más apropiado por la naturaleza máxima del esfuerzo (Moir et al., 2008), siendo el que proporciona el mayor valor, pero manteniendo una excelente fiabilidad (ICC = 0.97) con las otras dos mediciones (González-García et al., 2024). Inclusive se ha recomendado que para la evaluación de la fuerza prensil manual sea el valor tenido en cuenta para la definición y diagnóstico de la sarcopenia (Cruz-Jentoft et al., 2019). Sin embargo, un único intento debería ser valorado por su eficiencia temporal y alta reproducibilidad en contextos específicos (Coffey et al., 2020).

Aunque los resultados muestran que las estimaciones obtenidas a partir de uno, dos o tres mediciones pueden ser concordantes y confiables, es importante distinguir estas propiedades de la precisión estadística. Desde una perspectiva metodológica, la precisión de una estimación está relacionada con la variabilidad del estimador y la incertidumbre asociada a su cálculo, determinadas en parte por el error aleatorio derivado del proceso de medición y del muestreo (Altman & Bland, 2014). Aumentar el número de mediciones o el tamaño de la muestra puede contribuir a reducir este error y mejorar la precisión, siempre que se considere la dependencia entre las observaciones y las fuentes de variabilidad involucradas (Weichenthal et al., 2017). En los estudios epidemiológicos, estas consideraciones son fundamentales para interpretar adecuadamente los resultados y tener en cuenta la variabilidad intraindividual (Indrayan, 2025).

El aumento del número de mediciones puede mejorar la precisión de las estimaciones al reducir la varianza y el error estándar bajo determinadas condiciones (Altman & Bland, 2014; Weichenthal et al., 2017). Sin embargo, esta precisión depende del número de sujetos, las mediciones realizadas por individuo y los componentes de varianza asociados a las diferencias intraindividuales e interindividuales (Weichenthal et al., 2017). Una estimación inadecuada de estos componentes puede distorsionar la incertidumbre de los resultados y conducir a interpretaciones erróneas de su precisión (Kwok et al., 2007). En este sentido, incrementar el tamaño muestral o el número de mediciones puede contribuir a reducir el error estándar y obtener intervalos de confianza más estrechos, siempre que se utilicen procedimientos estadísticos apropiados para la estructura de los datos (Altman & Bland, 2014). No obstante, cuando el objetivo es analizar la variabilidad de las mediciones repetidas, reducirlas a un único indicador resumen, como el promedio o el valor máximo, limita la información disponible sobre las diferencias entre intentos. Por ello, resulta pertinente conservar los registros individuales en formato largo (long) y considerar técnicas estadísticas que tengan en cuenta la correlación entre las observaciones, como los modelos de medidas repetidas (de Melo et al., 2022).

De igual manera, la ley de los grandes números vincula directamente la cantidad de las mediciones con la calidad de la estimación de los parámetros poblacionales (Landon & Freedman, 1993). Este principio establece que, a medida que aumenta el número de observaciones, el promedio muestral tiende a acercarse al valor real de la población (Cabuñas et al., 2026; Landon & Freedman, 1993). En el diseño de estudios, este comportamiento garantiza que la estimación sea más eficiente y representativa, permitiendo que el investigador obtenga resultados que se aproximen a la realidad poblacional con mayor seguridad (Landon & Freedman, 1993).

Finalmente, es fundamental distinguir entre la precisión y la exactitud de las estimaciones, ya que ambas constituyen propiedades diferentes. Mientras que la precisión se relaciona con la variabilidad aleatoria de las estimaciones, la exactitud se refiere a su proximidad al valor verdadero del parámetro de interés.

Así, aunque incrementar el número de participantes o de mediciones puede reducir el error aleatorio y mejorar la precisión bajo determinadas condiciones, esto no corrige por sí mismo los sesgos sistemáticos asociados al instrumento o al procedimiento de medición (Altman & Bland, 2014).

Conclusiones

Independientemente de que se emplee el valor de un único intento o el promedio de dos o tres mediciones como estimador único por individuo, las estimaciones suelen mostrar concordancia y confiabilidad operativa; en términos prácticos, cualquier estrategia tiende a arrojar resultados similares. No obstante, se recomienda evitar la agregación previa mediante promedios cuando se dispone de múltiples registros por sujeto. En su lugar, se aconseja implementar una estructura de medidas repetidas con anidamiento por individuo que incorpore la totalidad de las mediciones disponibles, permitiendo así modelar adecuadamente la variabilidad intraindividual e interindividual y capitalizar los elementos metodológicos detallados en esta discusión.

Financiación

La presente investigación fue financiada en su totalidad por la Universidad de los Llanos, Colombia.

Agradecimientos:

Los autores agradecen encarecidamente la valiosa contribución del fisioterapeuta Jaime Mauricio Hernández-Jiménez, quien actuó como asistente de investigación y estuvo a cargo de las evaluaciones descritas en este estudio.

Afiliación de los autores

Karen Julieth Quintero Diaz. Fisioterapeuta, Magíster en Fisioterapia. Docente Escuela de Cuidado Universidad de los Llanos, Colombia.

Oscar Alexander Gutiérrez Lesmes. Enfermero, Doctor en Epidemiología. Docente Escuela de Salud Pública Universidad de los Llanos, Colombia.

Javier Martínez Torres. Fisioterapeuta, Doctor en Epidemiología. Grupo FAVISA, Universidad de los Llanos, Colombia.

Mónica Vannesa Chanagá Gelves. Fisioterapeuta, Magíster en Neurorehabilitación. Docente Escuela de Cuidado, Universidad de los Llanos, Colombia.

Sergio Alejandro Quijano Duarte. Fisioterapeuta, Especialista en Entrenamiento Deportivo y Magíster en Ciencias de la Actividad Física y del Deporte. Docente Escuela de Ciencias de la Salud (ECISA) de la Universidad Nacional Abierta y a Distancia (UNAD), Colombia.

Referencias

Altman, D., & Bland, J. (2014). Uncertainty beyond sampling error. *BMJ (Online)*, *349*, g7065. https://doi.org/10.1136/bmj.g7065

Attia, A., Dhahbi, W., Chaouachi, A., Padulo, J., Wong, D. P., & Chamari, K. (2017). Measurement errors when estimating the vertical jump height with flight time using photocell devices: The example of Optojump. *Biology of Sport*, *34*(1), 63–70. https://doi.org/10.5114/biolsport.2017.63735

Azócar-Gallardo, J., Azócar-Arancibia, F., Gutiérrez-Laclote, G., Ávila-Saldaña, C., Olivares-Arancibia, J., & Ojeda-Aravena, A. (2022). Confiabilidad relativa entre la plataforma de contacto Axon Jump 4.0® y la plataforma Globus Ergo Tester Italia®. *Journal of Movement & Health*, *19*(2), 1–7. https://doi.org/10.5027/jmh-Vol19-Issue2(2022)art164

Bland, J. ., Altman, D. G., & Warner, D. S. (2012). Agreed statistics: Measurement method comparison. *Anesthesiology*, *116*(1), 182–185. https://doi.org/10.1097/ALN.0b013e31823d7784

Bohannon, R. W. (2015). Muscle strength: Clinical and prognostic value of hand-grip dynamometry. *Current Opinion in Clinical Nutrition and Metabolic Care*, *18*(5), 465–470. https://doi.org/10.1097/MCO.0000000000000202

Cabuñas, J. T., Montero, V. J. G., & Fillone, A. M. (2026). Stratified Sampling Framework for Household Travel Survey: A Monte Carlo Simulation Approach for Sample Size Optimization. *Future Transportation*, *6*(4), 159. https://doi.org/10.3390/futuretransp6040159

Cagua Ardila, Y. A., Portilla Díaz, M., & Martínez-Torres, J. (2023). Normative values for handgrip strength in Colombian older adults: Estimation by quantile regression. *Semergen*, *50*(2), 102123. https://doi.org/10.1016/j.semerg.2023.102123

Coffey, K., McCollum, R., Smyth, E., Casey, E., Plunkett, J., & Horner, K. (2020). Reproducibility of objective and subjective markers of exercise recovery in college aged males. *International Journal of Exercise Science*, *13*(4), 1041–1051. https://doi.org/10.70252/qafj7094

Cruz-Jentoft, A. J., Bahat, G., Bauer, J., Boirie, Y., Bruyère, O., Cederholm, T., Cooper, C., Landi, F., Rolland, Y., Sayer, A. A., Schneider, S. M., Sieber, C. C., Topinkova, E., Vandewoude, M., Visser, M., Zamboni, M., Bautmans, I., Baeyens, J. P., Cesari, M., … Schols, J. (2019). Sarcopenia: Revised European consensus on definition and diagnosis. *Age and Ageing*, *48*(1), 16–31. https://doi.org/10.1093/ageing/afy169

de Lima, T. R., Martins, P. C., Moreno, Y. M. F., Chaput, J. P., Tremblay, M. S., Sui, X., & Silva, D. A. S. (2022). Muscular Fitness and Cardiometabolic Variables in Children and Adolescents: A Systematic Review. *Sports Medicine*, *52*(7), 1555–1575. https://doi.org/10.1007/s40279-021-01631-6

de Melo, M. B., Daldegan-Bueno, D., Menezes Oliveira, M. G., & de Souza, A. L. (2022). Beyond ANOVA and MANOVA for repeated measures: Advantages of generalized estimated equations and generalized linear mixed models and its use in neuroscience research. *European Journal of Neuroscience*, *56*(12), 6089–6098.

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). Development of a measurement instrument. En H. C. W. de Vet, C. B. Terwee, L. B. Mokkink, & D. L. Knol (Eds.), *Measurement in Medicine* (pp. 30–64). Cambridge University Press. https://doi.org/10.1017/CBO9780511996214.004

García-Hermoso, A., Cavero-Redondo, I., Ramírez-Vélez, R., Ruiz, J. R., Ortega, F. B., Lee, D. C., & Martínez-Vizcaíno, V. (2018). Muscular Strength as a Predictor of All-Cause Mortality in an Apparently Healthy Population: A Systematic Review and Meta-Analysis of Data From Approximately 2 Million Men and Women. *Archives of Physical Medicine and Rehabilitation*, *99*(10), 2100-2113.e5. https://doi.org/10.1016/j.apmr.2018.01.008

García-Hermoso, A., Ramírez-Campillo, R., & Izquierdo, M. (2019). Is muscular fitness associated with future health benefits in children and adolescents? a systematic review and meta-analysis of longitudinal studies. *Sports Medicine*, *49*, 1079–1094. https://doi.org/10.1007/s40279-019-01098-6

González-García, J., Conejero, M., & Gutiérrez-Hellín, J. (2024). Assessing Jump Performance: Intra- and Interday Reliability and Minimum Difference of Countermovement Jump and Drop Jump Outcomes, Kinetics, Kinematics, and Jump Strategy. *Applied Sciences (Switzerland)*, *14*(6), 2662. https://doi.org/10.3390/app14062662

Indrayan, A. (2025). Determining the size of sample for rigorous conclusions. *Journal of Postgraduate Medicine*, *71*(4), 201–204. https://doi.org/10.4103/jpgm.jpgm_697_25

Karatrantou, K., Xagorari, A., Vasilopoulou, T., & Gerodimos, V. (2020). Does the number of trials affect the reliability of handgrip strength measurement in individuals with intellectual disabilities? *Hand Surgery and Rehabilitation*, *39*(3), 223–228. https://doi.org/10.1016/j.hansur.2020.01.004

Kwok, O., West, S. G., & Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. *Multivariate Behavioral Research*, *42*(3), 557–592. https://dx.doi.org/10.1080/00273170701540537

Landon, R., & Freedman, L. S. (1993). Estimating correlations between questionnaires and food records/true dietary intake. *Communications in Statistics-Theory and Methods*, *22*(8), 2327–2340. https://dx.doi.org/10.1080/03610929308831152

Martinez-Torres, J., Gallo-villegas, J. A., & Aguirre-acevedo, D. C. (2025). Indicadores bioquímicos y cardiovasculares asociados a nos y adolescentes. Una revisión de alcance. *Medicina de Familia. SEMERGEN*, *51*(4), 102459. https://doi.org/10.1016/j.semerg.2025.102459

Martínez-Torres, J., Gallo Villegas, J., & Aguirre-Acevedo, D. (2022). Características antropométricas y de composición corporal asociadas a la fuerza prensil manual en niños y adolescentes. Una Revisión Sistemática Exploratoria. *Andes pediatrico*, *93*(6), 906–917. https://doi.org/10.32641/andespediatr.v93i6.4408

Martinez-Torres, J., & Gutiérrez-Lesmes, O. A. (2017). Componentes éticos mínimos que debe tener la investigación clínica. *Revista Criterios*, *24*(1), 15–17. https://revistas.umariana.edu.co/index.php/Criterios/article/view/1724/1808

Martínez Torres, J., Bucheli Sánchez, L. M., Manrique Gutiérrez, L., Cruz, R., Rojas Hoyos, Z., & Pérez Buelvas, J. L. (2013). Concordancia del autorreporte de peso y talla para valoración nutricional en estudiantes universitarios de 18 a 25 años. *Perspectivas en Nutrición Humana*, *15*(1), 57–65. https://doi.org/10.17533/udea.penh.17904

Moir, G., Shastri, P., & Connaboy, C. (2008). Intersession reliability of vertical jump height in women and men. *Journal of Strength and Conditioning Research*, *22*(6), 1779–1784. https://doi.org/10.1519/JSC.0b013e318185f0df

Petré, H., Psilander, N., & Rosdahl, H. (2023). Between-Session Reliability of Strength- and Power-Related Variables Obtained during Isometric Leg Press and Countermovement Jump in Elite Female Ice Hockey Players. *Sports*, *11*(5), 96. https://doi.org/10.3390/sports11050096

Riebe, D., Ehrman, J. K., Liguori, G., & Magal, M. (2018). *ACSM's guidelines for exercise testing and prescription*. American College of Sports Medicine.

Ruiz, J. R., Castro-Piñero, J., Artero, E. G., Ortega, F. B., Sjöström, M., Suni, J., & Castillo, M. J. (2009). Predictive validity of health-related fitness in youth: A systematic review. *British Journal of Sports Medicine*, *43*(12), 909–923. https://doi.org/10.1136/bjsm.2008.056499

Venegas-Carro, M., Kramer, A., Moreno-Villanueva, M., & Gruber, M. (2022). Test–Retest Reliability and Sensitivity of Common Strength and Power Tests over a Period of 9 Weeks. *Sports*, *10*(11), 171. https://doi.org/10.3390/sports10110171

Vivas, J. C., Martín-Martínez, J. P., Hernández-Mocholí, M. Á., & Pérez-Gómez, J. (2016). Validation of the iPhone app using the force platform to estimate vertical jump height. *Journal of Sports Medicine and Physical Fitness*, *58*(3), 227–232. https://doi.org/10.23736/S0022-4707.16.06664-0

Wang, D. X. M., Yao, J., Zirek, Y., Reijnierse, E. M., & Maier, A. B. (2020). Muscle mass, strength, and physical performance predicting activities of daily living: a meta-analysis. *Journal of Cachexia, Sarcopenia and Muscle*, *11*(1), 3–25. https://doi.org/10.1002/jcsm.12502

Weichenthal, S., Baumgartner, J., & Hanley, J. A. (2017). Sample size estimation for random-effects models: balancing precision and feasibility in panel studies. *Epidemiology*, *28*(6), 817–826. https://doi.org/10.1097/EDE.0000000000000727.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of strength and conditioning research*, *19*(1), 231–240. https://doi.org/10.1519/15184.1.

Datos de los/as autores/as y traductor/a:

Karen Julieth Quintero Diaz	karen.quintero.diaz@unillanos.edu.co	Autora
Oscar Alexander Gutiérrez-Lesmes	oagutierrez@unillanos.edu.co	Autor
Javier Martínez Torres	jmartineztorres@unillanos.edu.co	Autor
Mónica Vannesa Chanagá Gelves	mvchanagag@unillanos.edu.co	Autora
Sergio Alejandro Quijano Duarte	sergioa.quijano@unad.edu.co	Autor