

Fiabilidad absoluta de 2 pruebas de valoración del rango de movimiento del tobillo en jugadores de balonmano

Absolute reliability of 2 clinical tests for assessing ankle range of motion in handball player

Cejudo, A¹; Sainz de Baranda, P²; Ayala, F³; Santonja, F⁴

¹ Doctorando en Ciencias de la Actividad Física y del Deporte. Facultad de Medicina. Universidad de Murcia. acpejudo@gmail.com

² Profesora de la Facultad de Ciencias del Deporte. Universidad de Castilla La Mancha. pilar.sainzdebaranda@uclm.es

³ Profesor del Instituto Superior de Enseñanzas. Universidad de Murcia. franciscoayalarodriguez@gmail.com

⁴ Profesor de la Facultad de Medicina. Universidad de Murcia. Servicio de Traumatología. Hospital Universitario. V. de la Arrixaca. Murcia (España). fernando@santonjatrauma.es

Resumen

El objetivo de este estudio fue examinar la fiabilidad absoluta de las pruebas de estimación de la flexibilidad de la musculatura del tríceps sural ROM-soleo y ROM-gemelo a través de un diseño de medidas repetidas. 25 jugadores senior de balonmano completaron 3 sesiones de evaluación del ROM articular de la dorsi-flexión del tobillo con rodilla flexionada (ROM-soleo) y extendida (ROM-gemelo) con un intervalo de 2 semanas entre sesiones consecutivas. La fiabilidad absoluta fue examinada mediante el cálculo de los estadísticos cambio en la media (CM) entre sesiones de valoración, porcentaje del error típico (CV_{ET}) e índice de correlación intraclass (ICC). Los resultados del actual estudio demuestran que las pruebas ROM-soleo y ROM-gemelo presentan una elevada fiabilidad absoluta ($CM < 1\%$; $CV_{ET} < 4,6\%$; $ICC > 0,92$).

Palabras clave: Coeficiente de variación, reproducibilidad, error de la medida, rango de movimiento, variabilidad.

Abstract

The purpose of this study was to examine the absolute reliability of both ROM-soleus and ROM-gastrocnemius tests for estimating triceps surae flexibility throughout a repeated measures design. 25 senior handball players completed 3 measurement sessions of dorsi-flexion ankle ROM with knee flexed (ROM-soleus) and extended (ROM-gastrocnemius) with a 2-weeks rest interval between consecutive sessions. The absolute reliability was measured through the statistics change in the mean between measures (CM), typical percentage error (CV_{ET}) and intraclass correlation coefficients (ICC). The results of the current study showed that both ROM-soleus and ROM-gastrocnemius present high absolute reliability scores ($CM < 1\%$; $CV_{ET} < 4,6\%$; $ICC > 0,92$).

Key words: coefficient of variation, reproducibility, measurement error, range of motion, variability.

1. Introducción

La valoración de la flexibilidad de la musculatura del tríceps sural (gemelo y soleo) es una práctica habitual en el ámbito de la salud físico-deportiva. La razón fundamental reside en que reducidos valores de flexibilidad del tríceps sural han sido relacionados con un incremento de la probabilidad de sufrir lesiones por sobrecarga de la extremidad inferior, tales como: fascitis plantar, tendinopatía del tendón de Aquiles, síndrome de estrés medial de la tibia, síndrome de la banda iliotibial y síndrome de dolor femoropatelar (Alter, 2004; Johanson, Baer, Hovermale y Phouthavong, 2008; Radford, Burns, Buchbinder, Landorf y Cook, 2006; Zito, Driver, Parker y Bohannon, 1997).

La valoración eficaz del estado de una musculatura requiere de la selección de pruebas diagnósticas que posean un elevado grado de fiabilidad y validez, así como un reducido gasto humano y material. Las pruebas de estimación

de la flexibilidad del tríceps sural que valoran el rango de movimiento (ROM) articular en grados de la dorsi-flexión del tobillo con rodilla flexionada (ROM-soleo) y extendida (ROM-gemelo) en posición de bipedestación son probablemente las herramientas de estimación de la flexibilidad del tríceps sural más utilizadas en el ámbito científico y físico-deportivo (American Academic of Orthopedic Association, 1965; Ekstrand, Wiktorsson, Öberg y Gillquist, 1982, Sady, Wortman y Blanke, 1982; Möller, Öberg y Gillquist, 1985), debido principalmente a su facilidad de uso y al escaso material necesario para su desarrollo, y en menor medida a la evidencia científica existente con respecto a su grado de validez y de fiabilidad.

La evaluación de la fiabilidad absoluta (definida como la estabilidad de la medida a lo largo del tiempo) de las pruebas de estimación de la flexibilidad del tríceps sural (ROM-soleo y ROM gemelo) debe ser determinada antes de que puedan ser utilizadas legítimamente en el ámbito científico y clínico (Ayala y Sainz de Baranda, 2011). En este sentido, el conocimiento de la fiabilidad absoluta de las pruebas ROM-soleo y

Dirección para correspondencia: Pilar Sainz de Baranda. Facultad de Ciencias del Deporte. Universidad de Castilla La Mancha. Avenida Carlos III, s/n. 45071 Toledo. pilar.sainzdebaranda@uclm.es

ROM-gemelo es una información muy importante para clínicos y especialistas del ámbito físico-deportivo, ya que puede ser empleada para determinar su sensibilidad y estimar la magnitud necesaria en la variación de sus niveles iniciales, lo que podría ser considerado como un “cambio real” más allá del error de la medida (debido a variación técnica y biológica). A nivel práctico, el análisis de la fiabilidad absoluta permite valorar la “eficacia real” de programas de intervención sobre los valores iniciales de flexibilidad del tríceps sural de pacientes y deportistas; así como una mejor interpretación de los resultados obtenidos en estudios previos donde se emplean dichas pruebas exploratorias. Igualmente, otro uso importante de la fiabilidad absoluta es la posibilidad de comparación entre diferentes pruebas diagnósticas, e incluso clínicos e investigadores pueden emplear esta información para determinar el tamaño muestral de sus estudios (Atkinson y Nevill, 1998; Hopkins, 2000; Hopkins, Marshall, Batterham y Hanin, 2009).

Sin embargo, a pesar del extensivo uso que en la actualidad se está haciendo de las pruebas ROM-soleo y ROM-gemelo como herramientas de estimación de la flexibilidad del tríceps sural, sorprendentemente no se han encontrado (desde el conocimiento de los autores) estudios científicos que hayan determinado la fiabilidad absoluta de las mismas. Por ello, parece clara la necesidad de estudios científicos que aborden el análisis de la fiabilidad absoluta de dichas pruebas de estimación de la flexibilidad del tríceps sural.

Por lo tanto, el objetivo principal de este estudio fue examinar la fiabilidad absoluta inter-sesión de las pruebas de estimación de la flexibilidad de la musculatura del tríceps sural ROM-soleo y ROM-gemelo en jugadores de balonmano.

2. Método

2.1. Muestra

Un total de 25 jugadores masculino senior con más de 9 años de práctica deportiva federada en balonmano (3 sesiones de entrenamiento semanal con una duración mínima de 1,5 horas por sesión) participaron voluntariamente en este estudio. Los jugadores jugaban en la 1ª y 2ª División Nacional (edad: $26,8 \pm 4,4$ años; peso: $87,08 \pm 10,6$ Kgs; talla: $181 \pm 0,07$ cms) durante la temporada 2010/11.

Como criterios de exclusión se establecieron: (a) presentar alteraciones músculo-esqueléticas, tales como desgarros de la musculatura del tríceps sural y tendón de Aquiles en los últimos 6 meses previos al presente procedimiento exploratorio; (b) presentar dolor muscular de aparición tardía (agujetas) durante cualquiera de los tres momentos de evaluación; y (c) no asistir a una o más sesiones de valoración durante todo el proceso de recogida de datos.

Todos los criterios de inclusión y exclusión fueron eva-

luados por dos investigadores con dilatada experiencia en el ámbito científico y clínico empleando para este fin un cuestionario de evaluación médica y físico-deportiva. Todos los participantes fueron verbalmente informados de la metodología a utilizar, así como de los propósitos y posibles riesgos del estudio, y un consentimiento informado fue firmado por cada uno de ellos. El presente estudio fue aprobado por el Comité Ético y Científico de la Universidad de Murcia (España).

2.2. Procedimiento

El estudio de la fiabilidad absoluta de las pruebas de valoración del ROM articular de la dorsi-flexión del tobillo con rodilla flexionada (ROM-soleo) y extendida (ROM-gemelo) fue llevado a cabo a través de un diseño de medidas repetidas y de acuerdo con las recomendaciones establecidas por Hopkins et al. (2009).

Así, una semana antes del comienzo de la fase experimental, todos los participantes fueron sometidos a una sesión de familiarización con el propósito de conocer la correcta ejecución técnica de las pruebas exploratorias mediante la realización práctica de cada una de ellas. Igualmente, otro propósito de esta sesión fue la reducción del posible sesgo de aprendizaje sobre los resultados obtenidos en las diferentes valoraciones. Tras la sesión de familiarización, cada participante fue examinado un total de 3 ocasiones, con un intervalo de tiempo de 2 semanas entre sesiones consecutivas.

Cada una de las sesiones de valoración fue llevada a cabo por los mismos dos experimentados clínicos (uno controlaba la correcta posición del participante durante todo el proceso exploratorio [estabilización de segmentos corporales] y el otro conducía el test) bajo las mismas condiciones ambientales y franja horaria para tratar de minimizar la posible influencia de la variabilidad inter-examinador y ritmos circadianos sobre los resultados (Atkinson y Nevill, 1998). Los clínicos fueron ciegos en cuanto a los objetivos del estudio y a los resultados obtenidos en las sesiones exploratorias previas. Además, los participantes fueron instados a realizar cada una de las sesiones de valoración en los mismos días y franja horaria que normalmente realizaban sus sesiones de entrenamiento para minimizar la variabilidad intra-sujeto (Hopkins, 2000).

Previamente a cada sesión de valoración, todos los participantes realizaron 5 minutos de calentamiento aeróbico (carrera ligera) unido a una serie de ejercicios de estiramientos estandarizados (Gabbe, Bennell, Wajswelner y Finch 2004), enfatizando la actividad de los músculos gemelo y sóleo, bajo la estricta supervisión de los examinadores. En este sentido, dos ejercicios de estiramiento fueron seleccionados, imitando cada uno de ellos la posición adoptada en las dos pruebas de valoración seleccionadas. Estudios previos sugieren que las modificaciones que el estiramiento provoca sobre las propiedades viscoelásticas de la musculatura permanecen estables

durante al menos 20 minutos tras la aplicación de volúmenes de estiramiento de 120-150 segundos (Ford y McChesney, 2007; Power, Behm, Cahill, Carroll y Young, 2004). Por ello, para asegurar la estabilidad en las propiedades de la musculatura durante todo el proceso de valoración, la secuencia de ejercicios de estiramiento presentó un volumen total de 180 segundos (6 series de 30 segundos por ejercicio y pierna).

El calentamiento aeróbico y la secuencia estandarizada de estiramientos fue llevada a cabo porque: (a) todas las pruebas de valoración someten a la musculatura del gemelo y sóleo a fuerzas tensionales máximas; y (b) para tratar de minimizar la variabilidad y error estándar de la medida mediante la reducción del efecto que el estiramiento y la diferente temperatura muscular poseen sobre las propiedades viscoelásticas del tejido blando (Dixon y Keating, 2000).

Una vez finalizados el calentamiento y los estiramientos, los participantes fueron instados a realizar dos intentos máximos para cada una de las pruebas de valoración y segmento corporal (izquierdo y derecho) de forma aleatoria con el propósito de eliminar el sesgo que una secuencia específica podría presentar sobre los resultados obtenidos. El valor medio de cada par de intentos para cada prueba de valoración fue seleccionado para el posterior análisis estadístico (Ayala y Sainz de Baranda, 2011; Khan et al., 2000; Gabbe et al., 2004). La aleatorización en la realización de las pruebas de

valoración se llevo a cabo a través de la extracción ciega por parte de cada participante de una carta de una baraja española, de tal forma que la primera prueba a realizar estuvo determinada por el número de la carta seleccionada, equivaliendo los números pares a la prueba ROM-soleo y los impares a la prueba ROM-gemelo.

Cada participante fue examinado con ropa deportiva y sin calzado (Castro-Piñero, Chillón, Ortega, Montesinos, Sjöström y Ruiz, 2009). Se permitió un periodo de descanso de 2-3 minutos entre las pruebas de valoración (Ayala y Sainz de Baranda, 2011) con un descanso de aproximadamente 30 segundos entre cada uno de los dos intentos para cada prueba.

2.3. Pruebas de valoración

Las pruebas de valoración ROM-soleo y ROM-gemelo fueron llevadas a cabo siguiendo las directrices metodológicas establecidas por la American Academic of Orthopedic Association (1965), Ekstrand, Wiktorsson, Oberg y Gillquist (1982) y el American College of Sports Medicine (2001). Se utilizó para ambas pruebas exploratorias un inclinómetro ISOMED Unilevel con varilla telescópica extensible. Previo a cada sesión de valoración, el inclinómetro fue calibrado a 0° con la vertical (figura 1).

Figura 1. Pruebas de valoración del rango de movimiento del tobillo.



2.4. Análisis estadístico

Previo a todo análisis estadístico, la distribución normal de los datos fue comprobada a través de la prueba Kolomogorov-Smirnov. Se realizó un análisis descriptivo de cada una de las variables cuantitativas, que incluía la media y su correspondiente desviación típica. Además, una prueba t de Student fue empleada para determinar la existencia de diferencias significativas entre los valores de la pruebas de la extremidad

inferior derecha e izquierda.

La fiabilidad absoluta de cada una de las pruebas de valoración (ROM-soleo y ROM-gemelo) fue determinada a través del cálculo de los estadísticos cambio en la media entre sesiones de valoración expresado en términos porcentuales (CM), porcentaje del error típico (expresado como coeficiente de variación $[CV_{ET}]$) y a través del índice de correlación intraclase (ICC) empleando el método previamente descrito por Hopkins (2000) y Hopkins et al. (2009). Así, la fiabilidad

absoluta fue calculada empleando el valor medio de los valores de fiabilidad de cada una de las sesiones pareadas consecutivas (2-1 y 3-2) para cada una de las pruebas de valoración (Hopkins, 2000).

Un 2 (ROM-gemelo y ROM-sóleo) x 3 (sesión de valoración 1-3) análisis de la varianza con medidas repetidas en el último factor (RMANOVA) fue empleado para identificar el cambio en los valores medios (systematic bias) y la desviación típica de la diferencia entre las sesiones de valoración pareadas consecutivas para cada una de las variables evaluadas (Bonferroni post hoc test). La esfericidad de los datos fue acreditada a través de la prueba de Mauchly.

El CM fue calculado a través del modelo RMANOVA como diferencia de medias entre sesiones consecutivas, tomando el logaritmo de los valores conseguidos por los participantes.

El CV_{ET} fue calculado mediante el uso de logaritmos empleando la siguiente ecuación: $100(e^s - 1)$. En esta ecuación, s representa el error típico (desviación estándar de la diferencia entre sesiones de valoración consecutivas / $\sqrt{2}$). La transformación de los datos en logaritmos fue realizada como medio para minimizar de forma exitosa la posible presencia de heterocedasticidad en los mismos (Atkinson y Nevill, 1998; Hopkins, 2000). Para interpretar los resultados obtenidos a través del cálculo del CV_{ET} se ha considerado la idea arbitraria de que una variabilidad menor del 10-15% para una herramienta de medida ha sido considerada como "aceptable" por la literatura científica (Atkinson y Nevill, 1998; Castro-Piñero et al., 2009; Hopkins, 2000; Stokes, 1985).

Por su parte, el ICC de la muestra fue calculado siguiendo la fórmula: $(F - 1) / (F + k - 1)$. En esta fórmula, F es el F-ratio de los sujetos y k (3) es el número total de las sesiones de valoración (Hopkins et al., 2009; Schabert, Hopkins y Hawley, 1998). Hopkins et al. (2009) categorizan a través de una escala cualitativa la magnitud de los valores obtenidos en el estadístico ICC, de tal forma que valores próximos a 0,1 se consideran bajos, 0,3 moderados, 0,5 altos, 0,7 muy altos y los cercanos a 0,9 extremadamente altos.

El análisis estadístico fue realizado mediante el paquete estadístico SPSS (Statistical Package for Social Sciences, v. 16.0 para Windows; SPSS Inc, Chicago) y el programa Microsoft Excel 2003.

3. Resultados

La prueba t de Student para muestras relacionadas informó de no diferencias significativas ($p < 0,05$) en los resultados obtenidos en la prueba ROM-gemelo y ROM-soleo entre la pierna derecha y la pierna izquierda. Por lo tanto, para el análisis de la fiabilidad absoluta se utilizó la media de ambas piernas.

La tabla 1 presenta la estadística descriptiva (resultado medio de cada sesión de valoración [$k = 3$] \pm desviación estándar) y los estadísticos para la fiabilidad absoluta (media y 90% intervalo de confianza) de la muestra de estudio para cada una de las pruebas exploratorias. No se encontraron diferencias significativas entre los resultados obtenidos para las sesiones pareadas consecutivas ($p > 0,05$).

Tabla 1. Estadísticos de fiabilidad absoluta para las pruebas de ROM-gemelo y ROM-sóleo.

Grados	Media y desviación estándar			CM	Media y 90% IC	
	Medición 1	Medición 2	Medición 3		CV_{ET}	ICC
ROM_Gemelo	40,9 \pm 6,1	41,0 \pm 5,9	41,2 \pm 5,5	0,55 (-1,32-2,41)	3,86 (3,13-5,08)	0,93 (0,87-0,97)
ROM_Sóleo	41,8 \pm 6,4	42,3 \pm 6,1	42,3 \pm 5,6	0,73 (-1,46-2,91)	4,52 (3,67-5,95)	0,92 (0,84-0,96)

CM: Cambio en la media; CV_{ET} : coeficiente de variación en función del error típico de la medida; ICC: coeficiente de correlación intraclase; IC: intervalo de confianza

Todas las pruebas exploratorias analizadas presentaron valores altos de fiabilidad absoluta (CM: 0,55-0,73%; CV_{ET} : 3,86-4,52%; ICC: 0,92-0,93), no existiendo diferencias clínicamente relevantes entre sus niveles de variabilidad inter-sesión.

4. Discusión

La fiabilidad absoluta es un componente esencial para justificar la utilización de técnicas de valoración de la condi-

ción física en el campo de las Ciencias del Deporte (Peeler y Anderson, 2008). En este sentido, las pruebas de valoración ROM-soleo y ROM-gemelo son frecuentemente empleadas como tests clínicos y de campo para la estimación de la flexibilidad de la musculatura del tríceps sural.

Además, en la actualidad, existe una incipiente evidencia científica que propone la evaluación y la posterior monitorización de la flexibilidad del tríceps sural como herramienta para: (a) identificar a deportistas con elevado riesgo de padecer lesiones de la extremidad inferior (Bozic, Pazin, Berjan,

Planic, Cuk, 2010; Bradley, Olsen y Portas, 2007; Dadebo, White y George, 2004; Kibler y Chandler, 2003; Shellock y Prentice, 1985; Wang, Whitney, Burdett y Janosky, 1993; Weerapong, Hume y Kolt, 2004); y (b) para valorar cuantitativamente el impacto de determinadas intervenciones terapéuticas y/o el efecto de un programa de entrenamiento sobre su nivel inicial de movilidad.

Sin embargo, no se han encontrado (desde el conocimiento de los autores) estudios científicos que analicen la fiabilidad absoluta inter-sesión de las pruebas de valoración ROM-gemelo y ROM-soleo, quedando entonces su utilidad en tela de juicio. Por ello, el objetivo principal de este estudio fue examinar la fiabilidad absoluta inter-sesión de las pruebas de estimación de la flexibilidad de la musculatura del tríceps sural ROM-soleo y ROM-gemelo en jugadores/as senior de balonmano.

Para determinar la fiabilidad absoluta de las pruebas de valoración ROM-soleo y ROM-gemelo, el presente estudio utilizó los estadísticos CM, CV_{ET} e ICC, así como sus respectivos intervalos de confianza del 90%. Estos estadísticos fueron elegidos basándose en la propuesta de análisis para los estudios de reproducibilidad establecida recientemente por Hopkins et al. (2009).

El estadístico CM concretamente refleja el cambio en el valor medio de cada variable entre sesiones de evaluación (Hopkins, 2000). Este estadístico es especialmente importante cuando un grupo de deportistas llevan a cabo una serie de sesiones de valoración como parte del proceso de control de la eficacia de un programa de intervención, ya que la tendencia general de la medida tiende a ser diferente en una dirección en particular (positiva o negativa) entre sesiones de valoración (Hopkins, 2000; Cameron, Adams, Maher, 2003). Es importante utilizar sesiones de familiarización y procedimientos exploratorios sencillos que minimicen la posible influencia que el sesgo de aprendizaje podría tener sobre la magnitud de los resultados obtenidos en una prueba de valoración, lo cual podría influir en el estadístico CM (Iga, George, Lees, Reilly, 2006).

Cuando se quiera realizar una interpretación de los cambios producidos en una variable después de un programa de intervención en un grupo de personas, clínicos y profesionales del campo de las Ciencias del Deporte deben decidir si dichos cambios son reales o si únicamente reflejan la magnitud del error de la medida utilizada. En este sentido, si un cambio en los valores iniciales de una variable para un grupo de personas presenta una magnitud inferior al CV_{ET} de la medida, es muy probable (68% de certeza) que dicho cambio refleje el error de la medida y con ello no sea clínicamente relevante (Hopkins, 2000). Igualmente, Hopkins (2000) sugiere que un umbral en torno a 1,5-2 veces la magnitud del error típico podría también ser apropiado para indicar que se ha producido un cambio real en los niveles previos individuales de una

persona (80-90% probabilidad).

Finalmente, el estadístico ICC refleja el grado de proximidad o correlación existente entre los valores obtenidos en diferentes momentos tras la aplicación de la misma prueba de valoración. El estadístico ICC comparte con el estadístico CV_{ET} la ventaja de ser adimensional, y por lo tanto permiten la comparación con otros estudios de fiabilidad que empleen diferentes procedimientos exploratorios, herramientas de evaluación y poblaciones objeto de estudio (Hopkins, 2000).

Los resultados del presente estudio en lo relativo al estadístico CM reflejaron diferencias no significativas ($p > 0,05$) e inferiores al 2% en los resultados obtenidos en las pruebas de valoración ROM-soleo y ROM-gemelo entre sesiones de valoración. Estos resultados tentativamente sugieren que la sesión de familiarización llevada a cabo fue efectiva para eliminar el posible sesgo error (systematic bias) asociado al aprendizaje técnico de las pruebas exploratorias por parte de los participantes.

Si se considera la idea arbitraria normalmente aceptada de que una variabilidad inter-sesión menor del 10% para una herramienta de valoración dada puede ser considerada como aceptable para su utilización en el ámbito científico y clínico (Atkinson y Nevill, 1998; Hopkins, 2000; Hopkins, 2009; Vincent, 1994), entonces el análisis de la fiabilidad absoluta llevado a cabo en este estudio informa de valores aceptables de variabilidad intra-sujetos (expresados a través del CV_{ET}) tanto para la prueba ROM-gemelo (3,86%) como para la prueba ROM-soleo (4,52%), después de 3 sesiones de medida (con un intervalo de dos semanas entre sesiones consecutivas).

Igualmente, los valores obtenidos en el estadístico ICC para ambas pruebas de estimación de la flexibilidad del tríceps sural oscilaron entre 0,92 y 0,93, lo que se traduce en una calificación cualitativa de "extremadamente altos" en la escala de baremación establecida recientemente por Hopkins et al. (2009).

Por todo ello, la interpretación global de los resultados del presente estudio científico demuestra que las pruebas de estimación de la flexibilidad de la musculatura del tríceps sural ROM-soleo y ROM-gemelo son herramientas de medida con una gran fiabilidad absoluta.

Desde un punto de vista práctico, los resultados sobre fiabilidad absoluta de ambas pruebas de valoración derivados del presente estudio proporcionan una información muy útil, pues permite la toma de decisiones justificada sobre si se ha producido un "cambio real" entre sesiones de valoración tras la aplicación de un tratamiento (ej: programa de estiramientos para la musculatura) o si por el contrario, el cambio observado es simplemente producto del error típico de la medida. Así, si se emplea la propuesta de Hopkins (2000) como umbral (1,5-2 veces el error típico), un cambio en la flexibilidad mayor de 9,04% (ROM-gemelo) y un 5,88% (ROM-soleo), tras la realización de un programa de intervención podría in-

dicar que se ha producido un probable cambio real.

Desafortunadamente, la comparación de los resultados obtenidos en este estudio con los observados en otras investigaciones no fue posible porque, desde nuestro conocimiento, el actual estudio es el primero que analiza la fiabilidad absoluta de las pruebas ROM-soleo y ROM-gemelo empleando métodos estadísticos contemporáneos, como los propuestos por Hopkins (2000).

5. Conclusiones

Los resultados del presente estudio demuestran que las pruebas de valoración del ROM articular de la dorsi-flexión del to-

billo con rodilla flexionada (ROM-soleo) y extendida (ROM-gemelo) poseen una elevada fiabilidad absoluta ($CV_{ET} < 10\%$; $ICC > 0,9$) calculada por medio del método recientemente descrito por Hopkins et al. (2009). Por ello, se recomienda el uso de las pruebas de valoración ROM-soleo y ROM-gemelo para valorar la flexibilidad de la musculatura del tríceps sural.

Desde el punto de vista del entrenamiento deportivo, un cambio en los valores iniciales de flexibilidad del tríceps sural mayor del 9% para las prueba ROM-gemelo y ROM-soleo tras la realización de un programa de intervención podría indicar que se ha producido un “cambio real” (80-90% probabilidad) y no simplemente debido al error de la medida.

6. Referencias bibliográficas

- Alter, M.J. (2004). *Los estiramientos. Desarrollo de ejercicios*. Barcelona: Paidotribo.
- American Academy of Orthopaedic Association (1965). *Joint motion: method of measuring and recording*. Chicago: Park Ridge.
- American College of Sports Medicine (2001). *ACSM's resource manual for guidelines for exercise testing and prescription* (4ª edición). Baltimore MD.:Lippincott, Williams & Wilkins.
- Atkinson, G. & Nevill, A.M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 4, 217-238.
- Ayala F. & Sainz de Baranda, P. (2011). Reproducibilidad inter-sesión de las pruebas distancia dedos planta y distancia dedos suelo para estimar la flexibilidad isquiosural en jugadores adultos de fútbol sala de primera división. *Revista Andaluza de Medicina Deporte*, 4(2), 47-51.
- Bozic, P.R., Pazin, N.R., Berjan, B.B., Planic, N.M. & Cuk, I.D. (2010). Evaluation on the field tests of flexibility of lower extremity: Reability and the concurrent and factorial validity. *The Journal of Strength & Conditioning Research*, 24 (9), 2523-2531.
- Bradley, P.S., Olsen, P.D. & Portas M.D. (2007). The effect of static, ballistic, and proprioceptive neuromuscular facilitation stretching on vertical jump performance. *The Journal of Strength & Conditioning Research*, 21(1), 223-226.
- Cameron, M., Adams, R. & Maher, C. (2003). Motor control and strength as predictors of hamstring injury in elite players of Australian football. *Journal of Orthopaedic & Sports Physical Therapy*, 4,159-166.
- Castro-Piñero, J., Chillón, P., Ortega, F.B., Montesinos, J.L., Sjöström, M. & Ruiz, J.R. (2009). Criterion-related validity of sit-and-reach and modified sit-and-reach test for estimating hamstring flexibility in children and adolescents aged 6-17 years. *International Journal Sports Medicine*, 30, 658-662.
- Dadebo, B., White, J. & George, K.P. (2004). A survey of flexibility training protocols and hamstring strains in professional football clubs in England. *British Journal of Sports Medicine*, 38, 388-394.
- Davis, D.S., Quinn, R.O., Whiteman, C.T., Williams, J.D. & Young, C.R. (2008). Concurrent validity of four clinical tests used to measure hamstring flexibility. *Journal of Strength and Conditioning Research*, 22, 583-588.
- Dixon, J. & Keating, J.L. (2000). Variability in straight leg raise measurements. *Physiotherapy*, 86, 361-370.
- Ekstrand, J., Wiktorsson, M., Oberg, B. & Gillquist, J. (1982). Lower extremity goniometric measurements: A study to determine their reliability. *Archives of Physical Medicine and Rehabilitation*, 63, 171-175.
- Ford, P. & McCesney, J. (2007). Duration of maintained hamstring ROM following termination of three stretching protocols. *Journal of Sports Rehabilitation*, 16, 18-27.
- Gabbe, B., Bennell, K., Wajswelner, H. & Finch, C. (2004). The reliability of commonly used lower limb musculoskeletal screening tests. *Physical Therapy in Sport*, 5, 90-97.
- Hopkins, W.G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, 30, 1-15.
- Hopkins, W.G. (2009). Calculating the reliability intraclass correlation coefficient and its confidence limits (Excel spreadsheet). newstats.org/xICC.xls.
- Iga, J., George, K., Lees, A. & Reilly, T. (2006). Reliability of assessing indices of isokinetic leg strength in pubertal soccer players. *Pediatric Exercise Science*, 18, 436-445.
- Johanson, M.A., Baer, J., Hovermale, H. & Phouthavong, P. (2008). Subtalar joint position during gastrocnemius stretching and ankle dorsiflexion range of motion. *Journal of Athletic Training*, 43, 172-178.
- Khan, K.M., Bennell, K., Ng, S., Matthews, B., Roberts, P., Natrass, C., Way, S. & Brown, J. (2000). Can 16-18-year-old elite Ballet Dancers improve their hip and ankle range of motion over a 12-month period? *Clinical Journal of Sport Medicine*, 10, 98-103.
- Kibler, W.B. & Chandler, T.J. (2003). Range of movement in junior tennis player participating in an injury risk modification program. *Journal of Science and Medicine in Sport*, 6(1), 51-62.
- Möller, M.H., Öberg, B.E. & Gillquist, J. (1985). Stretching exercise and soccer: effect of stretching on range of motion in the lower extremity in connection with soccer training. *International Journal of Sports Medicine*, 6(1), 50-52.
- Peeler, J.D. & Anderson, J.E. (2008). Reliability limits of the modified Thomas test for assessing rectus femoris muscle flexibility about the knee joint. *Journal of Athletic Training*, 43, 470-476.
- Power, K., Behm, D., Cahill, F., Carroll, M. & Young, W. (2004). An acute bout of static stretching: effects on force and jumping performance. *Medicine & Science in Sports & Exercise*, 36 (8), 1389-1396.
- Radford, J., Burns, J., Buchbinder, R., Landorf, K.B. & Cook, C. (2006). Does stretching increase ankle dorsiflexion range of motion? A systematic review. *British Journal of Sports Medicine*, 40, 870-875.
- Schaborn, E.J., Hopkins, W.G. & Hawley, J.A. (1998). Reproducibility of self-paced treadmill performance of trained endurance runners. *International Journal of Sports Medicine*, 19, 48-51.
- Shellock, F.G. & Prentice, W.E. (1985). Warming-up and stretching for improved physical performance and prevention of sports-related injuries. *Sports Medical*, 2, 267-278.
- Stokes, M. (1985). Reliability and repeatability of methods for measu-

- ring muscle in physiotherapy. *Physiotherapy Theory And Practice*, 1, 71-6.
29. Vincent J. (1994). *Statistics in kinesiology*. Champaign (IL): Human Kinetics Books;
30. Wang, S.S., Whitney, S.L., Burdett, R.C. & Janosky, J.E. (1993). Lower extremity muscular flexibility in long distance runners. *Journal of Orthopaedic & Sports Physical Therapy*, 17 (2), 102-107.
31. Weerapong, P., Hume, P.A. & Kolt, G.S. (2004). Stretching: mechanisms and benefits for Sport performance and injury prevention. *Physical Therapy Reviews*, 9, 189-206.
32. Zito, M., Driver, D., Parker, C. & Bohannon, R.W. (1997). Lasting effects of one bout of two 15-second passive stretches on ankle dorsiflexion range of motion. *Journal of Orthopaedic & Sports Physical Therapy*, 26, 214-221.

