

# Confiabilidad en ecografía muscular esquelética por evaluadores experimentados y novatos de cuidados críticos

FELIPE GONZÁLEZ-SEGUEL<sup>1,2,a</sup>, JORGE MOLINA<sup>1,b</sup>, FRANCISCO RÍOS-CASTRO<sup>1,2,c</sup>

## Muscle ultrasound: reliability across experience levels in critical care physiotherapists

**Background:** Muscle ultrasound is a valid tool to monitor muscle mass loss in critically ill patients. The level of experience is essential to the accuracy of the measurements. **Aim:** To evaluate the interobserver reliability of experienced and novice raters measuring muscle thickness and echo intensity of the quadriceps and tibialis anterior. **Material and Methods:** Cross-sectional observational study. Twenty-four critical care physiotherapists participated (5 experienced and 19 novice). Following a standardized ultrasound protocol, each rater measured the thickness (centimeters) of the quadriceps and tibialis anterior of 10 healthy and young models using linear and convex probes of portable devices. The Intraclass Correlation Coefficient and the Minimal Detectable Change (95% confidence interval) were calculated. Additionally, the novices scored the echo intensity of 19 muscle ultrasound images of critically ill patients using the Heckmatt score (qualitative assessment). The agreement with experienced raters was evaluated (Spearman Rho). **Results:** 960 muscle thickness measurements were performed (experienced = 200 and novice = 760). The mean thickness of the quadriceps and tibialis anterior was  $4.4 \pm 0.77$  and  $2.4 \pm 0.35$  centimeters for the experienced and  $4.2 \pm 0.80$  and  $2.2 \pm 0.39$  centimeters for the novices, respectively. Quadriceps' and tibialis' anterior reliability were 0.82 and 0.86 for experienced and 0.76 and 0.41 for novices, respectively. The Minimal Detectable Change ranged from 0.14-0.33 centimeters. The mean Heckmatt score was  $2.6 \pm 0.83$  points, with a reliability of 0.68 and an agreement with the experimenters of 0.78 [ $p < 0.001$ ]. **Conclusions:** Interobserver reliability was excellent for experienced raters and moderate to good for novice raters. The level of experience could determine the reliability of the results.

(Rev Med Chile 2023; 151: 1153-1163)

**Keywords:** Critical Care; Muscular Atrophy; Ultrasonography.

### RESUMEN

**Introducción:** La ecografía muscular es una herramienta válida para monitorizar la pérdida de masa muscular de personas críticamente enfermas. El nivel de experiencia es clave para la precisión de las mediciones. **Objetivo:** Evaluar

<sup>1</sup>Carrera de Kinesiología, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago, Chile.

<sup>2</sup>Servicio de Medicina Física y Rehabilitación, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago, Chile.

<sup>a</sup>Kinesiólogo, Magíster en Terapia Física y Rehabilitación.

<sup>b</sup>Kinesiólogo, Magíster en Gestión en Salud.

<sup>c</sup>Kinesiólogo, ©Magíster en Terapia Física y Rehabilitación.

Financiamiento: Este trabajo no recibió financiamiento.

Conflicto de interés: Francisco Ríos-Castro y Felipe González-Seguel son docentes del curso de formación en ecografía muscular de la Universidad del Desarrollo, sin embargo, no recibieron remuneración para el desarrollo de este estudio. Para las mediciones ecográficas, tanto Andover como TechMedizin proporcionaron ecógrafos. Ni la universidad ni estas empresas tuvieron participación en el diseño o resultados de este estudio.

ORCID de autores:

F. González-Seguel <https://orcid.org/0000-0001-9099-2389>

J. Molina <https://orcid.org/0000-0003-1625-4657>

F. Ríos-Castro <https://orcid.org/0000-0002-8192-383X>

Recibido el 16 de octubre de 2022, aceptado el 9 de noviembre de 2023.

Correspondencia a:

Felipe González-Seguel  
Avenida Plaza #680, Santiago, Chile.  
[feligonzalez@udd.cl](mailto:feligonzalez@udd.cl)

la confiabilidad interobservador de evaluadores experimentados y novatos midiendo grosor y eco-intensidad del cuádriceps y tibial anterior. **Métodos:** Estudio observacional transversal. Participaron 24 kinesiólogos de cuidados críticos; 5 experimentados (> 4 años de experiencia ecográfica) y 19 novatos (sin experiencia previa, por lo que recibieron entrenamiento de 16 horas). De forma estandarizada, cada evaluador midió ecográficamente el grosor (centímetros) del cuádriceps y tibial anterior de 10 modelos sanos y jóvenes usando equipos portátiles (transductores lineales y convexos). Se calculó el Coeficiente de Correlación Intraclase y el Cambio Mínimo Detectable (95% intervalo de confianza). Además, los novatos calificaron la eco-intensidad de 19 ecografías musculares de personas críticamente enfermas con la escala Heckmatt (calificación visual cualitativa) y se midió el acuerdo con los experimentados (Spearman Rho). **Resultados:** Se realizaron 960 mediciones de grosor muscular (experimentados = 200 y novatos = 760). La media del grosor del cuádriceps y tibial anterior fue  $4,4 \pm 0,77$  y  $2,4 \pm 0,35$  centímetros para los experimentados y  $4,2 \pm 0,80$  y  $2,2 \pm 0,39$  centímetros para los novatos, respectivamente. La confiabilidad de cuádriceps y tibial anterior fue 0,82 y 0,86 para los experimentados y 0,76 y 0,41 para los novatos. El Cambio Mínimo Detectable osciló entre 0,14-0,33 centímetros. La puntuación media de Heckmatt fue  $2,6 \pm 0,83$  puntos, con una confiabilidad de 0,68 y una concordancia con los experimentados de 0,78 [ $p < 0,001$ ]. **Conclusiones:** La confiabilidad interobservador de los experimentados fue excelente y la de los novatos moderada a buena. El nivel de experiencia podría determinar los resultados de confiabilidad.

**Palabras clave:** Atrofia Muscular; Cuidados Críticos; Ultrasonografía.

Las personas críticamente enfermas pueden perder alrededor de un 30% de masa muscular durante los primeros diez días desde el ingreso a la unidad de cuidados intensivos (UCI)<sup>1-3</sup>. La disminución de la masa muscular se asocia a mayor mortalidad, prolongación de la ventilación mecánica y de la hospitalización<sup>4,5</sup>. Además, la evaluación de la masa muscular puede predecir la aparición de debilidad adquirida en la UCI cuando no es posible identificarla en pacientes no cooperadores<sup>6</sup>. Por lo tanto, la monitorización de los cambios en la masa muscular es esencial desde los inicios de la estadía en la UCI para guiar la rehabilitación<sup>7</sup>.

La ecografía muscular ha mostrado un creciente uso en el contexto de la UCI, siendo una herramienta con múltiples ventajas por sobre la resonancia magnética nuclear y tomografía axial computarizada, ya que es de bajo costo y puede ser usada a la cabecera del paciente, no genera daño y permite hacer un análisis rápido de las características musculares<sup>6,8</sup>. La ecografía permite evaluar la pérdida de masa muscular principalmente a través del grosor muscular, área de sección transversal y longitud del fascículo<sup>6,8</sup>, parámetros que pueden

guiar el uso de electroestimulación y cicloergometría en etapas iniciales de UCI. Comúnmente, en UCI se miden músculos de extremidades inferiores por sobre de extremidades superiores, especialmente cuádriceps<sup>9</sup>. Además de la medición de la pérdida de la cantidad muscular, mediante la adquisición de imágenes ecográficas se logra la valoración de la eco-intensidad, que corresponde al método establecido para evaluar la calidad del músculo<sup>6,8</sup>, siendo un potencial marcador del estado del tejido muscular<sup>10</sup>. Una de las alternativas para ponderar la eco-intensidad es mediante la escala Heckmatt desarrollada por Heckmatt y Dubowitz, quienes propusieron esta escala en base al criterio cualitativo para valorar el músculo<sup>11</sup>.

Las propiedades clinimétricas de la ecografía muscular han sido ampliamente reportadas en sujetos sanos<sup>12</sup>. Sin embargo, los datos reportados en el contexto de UCI siguen siendo escasos. Las investigaciones en torno a la confiabilidad y validez de la ecografía muscular son de vital importancia para confirmar que las mediciones hechas son consistentes y permiten la comparación con otras poblaciones<sup>8</sup>. Una de las barreras actuales en el uso de la ecografía, corresponde a los

potenciales errores de medición e interpretación por las diferencias en el nivel de experiencia de los evaluadores<sup>13,14</sup>. Consecuentemente, el objetivo de este estudio fue determinar la confiabilidad interobservador de evaluadores experimentados y novatos al medir ecográficamente el grosor muscular del recto anterior, vasto intermedio, cuádriceps y tibial anterior de modelos sanos y la eco-intensidad muscular de imágenes ecográficas de personas críticamente enfermas.

## Material y Método

### Diseño

Se realizó un estudio de confiabilidad mediante un diseño observacional transversal proveniente del estudio chileno eMUSICS (*Education in Muscle Ultrasound for Intensive Care Setting*, por sus siglas en inglés)<sup>15</sup>. Las mediciones se realizaron en la Facultad de Medicina de la Universidad del Desarrollo. Para la presentación de los resultados de confiabilidad se siguieron las recomendaciones COSMIN (*Consensus-based Standards for the selection of health status Measurement Instruments*)<sup>(16)</sup>. Este estudio siguió las normas de la Declaración de Helsinki y fue aprobado por el Comité Ético Científico de la Universidad del Desarrollo (#2020-106). Todos los participantes entregaron su consentimiento informado.

### Participantes

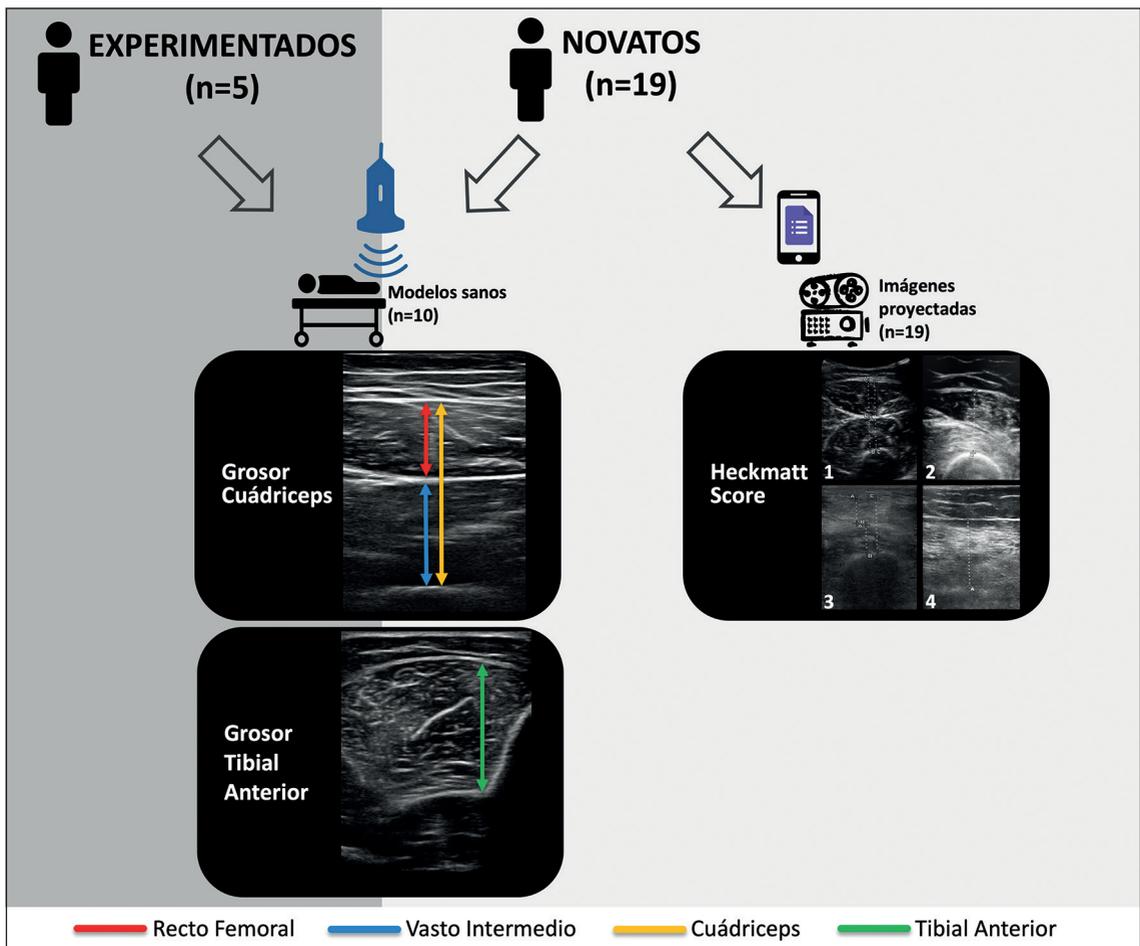
En total participaron 24 kinesiólogos de cuidados críticos seleccionados por conveniencia. Cinco evaluadores se definieron como experimentados para este estudio, ya que tenían más de 4 años de experiencia con ecografía en UCI y habían recibido capacitación formal en cursos nacionales o internacionales de ecografía; y 19 evaluadores como novatos, ya que tenían poca o nula experiencia clínica en ecografía previamente, por lo que antes de las mediciones recibieron entrenamiento teórico (10 horas) y práctico (6 horas) en un curso de ecografía muscular<sup>15</sup> guiado por los cinco evaluadores experimentados. El entrenamiento consistió en 5 horas de e-learning y 5 horas de clases teóricas, incluyendo física del ultrasonido, knobología (familiarización con el ecógrafo), anatomía y fisiología muscular, posicionamiento del paciente, puntos de referencia, adquisición de imágenes y mediciones de calidad y cantidad

muscular de extremidades inferiores. Finalmente, los novatos recibieron 6 horas de entrenamiento práctico supervisado justo antes de las mediciones de este estudio.

Para las mediciones musculares participaron 10 modelos, quienes fueron estudiantes de kinesiología de ambos sexos sin antecedentes de lesiones musculoesqueléticas de extremidades inferiores o enfermedades neurológicas.

### Procedimiento

Se evaluó cuantitativamente el músculo a través de la medición del grosor muscular<sup>(17)</sup> y cualitativamente a través de la observación de la eco-intensidad muscular usando la escala Heckmatt<sup>11</sup> (Figura 1). Las mediciones se realizaron el 9 de julio de 2022 entre las 15:00 y las 18:00 horas los novatos y entre las 18:00 y las 19:30 horas los experimentados. Primero, siguiendo un protocolo estandarizado de 8 minutos, todos los evaluadores (experimentados y novatos) midieron en centímetros el grosor del recto femoral, vasto intermedio, cuádriceps (grosor del recto femoral + vasto intermedio) y tibial anterior. Cada uno realizó estas mediciones en la extremidad inferior derecha de los 10 modelos usando transductores lineales o convexos de ecógrafos portátiles (Lumify, Philips Healthcare y Clarius Mobile Health Corp). Tanto los experimentados como los novatos evaluaron a los mismos modelos y realizaron una medición por músculo. Los modelos fueron evaluados en posición supina con extensión pasiva de la extremidad inferior y evitando rotaciones, siguiendo un protocolo de medición previamente reportado<sup>(15)</sup>. Segundo, cada evaluador novato calificó la eco-intensidad de 19 imágenes ecográficas musculares usando la escala Heckmatt<sup>11</sup>. La escala Heckmatt califica la eco-intensidad muscular en base a 4 grados: 1) eco-intensidad normal; 2) eco-intensidad muscular aumentada con reflejo óseo normal, 3) eco-intensidad muscular aumentada con reflejo óseo reducido y 4) eco-intensidad muscular notablemente aumentada con pérdida del reflejo óseo. Para esta evaluación los novatos estuvieron todos juntos en una sala oscura y sin distracciones, teniendo un minuto por imagen proyectada para puntuar en un formulario Google. Las imágenes ecográficas fueron seleccionadas por los investigadores desde un banco de imágenes anonimizadas de cuádriceps y tibial anterior de personas críticamente enfermas reales. Previamente, los



**Figura 1.** Procedimiento de las mediciones del estudio de confiabilidad en ecografía muscular por los evaluadores experimentados y novatos. Siguiendo el protocolo de Selina Parry<sup>15</sup>, todos los evaluadores evaluaron a los mismos modelos y realizaron una medición única por músculo. Los modelos fueron evaluados en posición supina con extensión pasiva de la extremidad inferior y evitando rotaciones. Los puntos de referencia anatómicos se marcaron con un marcador de piel y se midieron en centímetros con una cinta métrica flexible. El punto de referencia para el cuádriceps se definió como el punto medio entre la espina iliaca anterosuperior y el borde superior de la patela. El punto de referencia para el tibial anterior se definió como el punto ubicado en el tercio proximal de la distancia entre la meseta tibial hasta el borde inferior del maléolo lateral. Se utilizó una cantidad generosa de gel de contacto para minimizar la presión del transductor sobre la piel. Las mediciones se realizaron con el transductor en inclinación neutra utilizando una vista de sección transversal para medir el grosor muscular, utilizando el límite interno de los bordes epimisiales de cada músculo.

evaluadores expertos calificaron y consensuaron el puntaje Heckmatt de cada imagen (se debieron resolver 5 desacuerdos de las 19 imágenes). Los evaluadores novatos fueron ciegos a la puntuación acordada por los evaluadores expertos.

#### Análisis estadístico

Los datos se presentan como media  $\pm$  desviación estándar o mediana (rango intercuartílico),

dependiendo de la distribución de los datos. El límite de significancia estadística se fijó en un valor  $p$  bilateral  $\leq 0,05$ . El tamaño muestral para la medición de la confiabilidad se calculó en base a 5 evaluadores experimentados, con un  $\alpha$  de 0,05 y potencia de 0,8, necesitando 10 mediciones por evaluador; y para los 19 evaluadores novatos, la estimación fue de 6 evaluaciones<sup>18</sup>. Por lo tanto, se fijaron 10 mediciones por evaluador. La con-

fiabilidad interobservador del grosor muscular se reportó con el Coeficiente de Correlación Intra-clase modelo 2-way random effect, single measure (ICC<sub>2,1</sub>) entre los experimentados, novatos y entre todos los evaluadores. Por su naturaleza cualitativa, las puntuaciones Heckmatt se ponderaron para ser usadas como variables continuas y usar ICC para la medición de la eco-intensidad de los novatos. La confiabilidad se interpretó como 0,40-0,59 (moderada), 0,60-0,79 (buena) y  $\geq 0,80$  (excelente)[19]. El error estándar de medición (SEM) se calculó para cada valor de ICC como:  $DE \times \sqrt{1 - ICC}$ , siendo DE la desviación estándar. El Cambio Mínimo Detectable (MDC<sub>95</sub>) se calculó en base a un intervalo de confianza del 95% como:  $1,96 \times SEM \times \sqrt{2}$ , definido como la diferencia clínicamente significativa más allá del rango de error<sup>20</sup>. Además, se midió la concordancia entre las puntuaciones de Heckmatt de los evaluadores novatos y el consenso de los experimentados usando Spearman Rho. Todos los análisis fueron realizados con STATA SE 15.0.

## Resultados

Las características de los evaluadores se presentan en la Tabla 1 y las características de los modelos en la Tabla 2. Se realizaron 960 mediciones de grosor muscular en total, incluyendo las evaluaciones de los experimentados (n = 200) y las de los novatos (n = 760).

La media  $\pm$  desviación estándar del grosor del cuádriceps y tibial anterior fue  $4,4 \pm 0,77$  y  $2,4 \pm 0,35$  centímetros para los experimentados y  $4,2 \pm 0,80$  y  $2,2 \pm 0,39$  centímetros para los novatos, respectivamente. Los grosores musculares medidos por los novatos fueron entre 3,5 y 6,4% menor que el medido por los experimentados ( $p < 0,03$ ) (Tabla 3). Los evaluadores novatos tuvieron menores valores de ICC, mayores valores de SEM y de MDC<sub>95</sub> comparado con los experimentados (Tabla 4). Los evaluadores experimentados obtuvieron excelente confiabilidad interobservador en la medición del grosor muscular (ICC entre 0,80-0,87); mientras los

**Tabla 1. Características de los evaluadores experimentados y novatos**

Variables	Todos (n = 24)	Experimentados (n = 5)	Novatos (n = 19)
Edad, años	29 [27-32]	34 [32-37]	29 [27-31]
Género femenino	8 (33%)	0 (0%)	8 (42%)
Trabaja en UCI	22 (92%)	5 (100%)	17 (90%)
Experiencia laboral, años	5 [3-7]	11 [7-13]	4 [3-5]
Experiencia en hospitalizado, años	3 [2-7]	11 [7-13]	3 [2-4]
Experiencia con pacientes críticos, años	2 [2-6]	11 [7-13]	2 [2-3]
Nivel académico más alto			
Pre-grado	1 (4%)	0 (0%)	1 (5%)
Curso post-título	1 (4%)	0 (0%)	1 (5%)
Diplomado	17 (71%)	1 (20%)	16 (84%)
Magíster	5 (21%)	4 (80%)	1 (5%)
Certificación DENAKE o CLACPTER	3 (13%)	3 (60%)	0 (0%)
Entrenamiento previo ecografía general	8 (33%)	5 (100%)	3 (16%)
Entrenamiento previo ecografía muscular	6 (25%)	5 (100%)	1 (5%)
Experiencia con ecografía muscular, años	0 [0-4]	5 [4-7]	0 [0-0]

UCI = unidad de cuidados intensivos, DENAKE = Departamento de Acreditación de Especialidades y Especialistas de Kinesología, CLACPTER = Consejo Latinoamericano para la Certificación Profesional en Terapia Respiratoria.

**Tabla 2. Características de los modelos**

VARIABLES	Modelos (n = 10)
Edad, años	24 [23-27]
Género femenino	5 (50%)
Altura, centímetros	173 [161-177]
Peso, kilogramos	72 [62-90]
Índice masa corporal	24,6 [23-29]
Nivel de actividad física	
Físicamente activo	5 (50%)
Físicamente inactivo	5 (50%)
Frecuencia actividad deportiva	
1 o 2 veces por semana	2 (20%)
≥ 3 veces por semana	4 (40%)
No realiza	4 (40%)

novatos obtuvieron moderada a buena confiabilidad (ICC entre 0,41-0,76). La confiabilidad interobservador de todos los evaluadores osciló entre 0,48 y 0,77. El MDC<sub>95</sub> de las mediciones de grosor muscular osciló entre 0,14 y 0,19 centímetros en los experimentados y entre 0,22 y 0,33 centímetros en los novatos.

La media  $\pm$  desviación estándar de las puntuaciones de Heckmatt de los evaluadores novatos fue de  $2,6 \pm 0,83$  puntos. Los novatos obtuvieron buena confiabilidad interobservador en las puntuaciones Heckmatt (ICC = 0,68), mostrando concordancia con el puntaje consensuado de los experimentados (Rho = 0,78 [ $p < 0,001$ ]).

**Tabla 3. Diferencias absolutas y relativas de la medición del grosor muscular entre los evaluadores experimentados y novatos**

Parámetro	Diferencia experimentados - novatos centímetros (%) *	Intervalo de confianza al 95%	Valor p
Grosor recto femoral	0,08 (3,5%)	0,00 – 0,17	0,03
Grosor vasto intermedio	0,13 (6,4%)	0,03 – 0,24	0,01
Grosor cuádriceps	0,24 (5,4%)	0,11 – 0,38	0,00
Grosor tibial anterior	0,12 (5,1%)	0,03 – 0,21	0,01

\*El porcentaje fue calculado como: la diferencia en centímetros dividido por grosor medido por los evaluadores experimentados por 100.

**Tabla 4. Confiabilidad interevaluador de los participantes**

Evaluadores	Parámetro	Media $\pm$ DE	Confiabilidad interevaluador		
			ICC <sub>2,1</sub> (IC 95%)	SEM	MDC <sub>95</sub>
Todos (n = 24)	Grosor recto femoral, cm	2,19 $\pm$ 0,37	0,51 (0,29-0,74)	0,12	0,33
	Grosor vasto intermedio, cm	1,93 $\pm$ 0,61	0,73 (0,55-0,90)	0,09	0,25
	Grosor cuádriceps, cm	4,25 $\pm$ 0,80	0,77 (0,61-0,93)	0,08	0,22
	Grosor tibial anterior, cm	2,27 $\pm$ 0,38	0,48 (0,25-0,71)	0,12	0,33
Experimentados (n = 5)	Grosor recto femoral, cm	2,27 $\pm$ 0,35	0,80 (0,66-0,95)	0,07	0,19
	Grosor vasto intermedio, cm	2,03 $\pm$ 0,58	0,87 (0,77-0,97)	0,05	0,14
	Grosor cuádriceps, cm	4,44 $\pm$ 0,77	0,82 (0,69-0,96)	0,07	0,19
	Grosor tibial anterior, cm	2,36 $\pm$ 0,35	0,86 (0,75-0,97)	0,06	0,17
Novatos (n = 19)	Grosor recto femoral, cm	2,18 $\pm$ 0,37	0,46 (0,23-0,69)	0,12	0,33
	Grosor vasto intermedio, cm	1,90 $\pm$ 0,62	0,69 (0,51-0,89)	0,09	0,25
	Grosor cuádriceps, cm	4,20 $\pm$ 0,80	0,76 (0,60-0,92)	0,08	0,22
	Grosor tibial anterior, cm	2,24 $\pm$ 0,39	0,41 (0,18-0,64)	0,12	0,33
	Heckmatt score	2,61 $\pm$ 0,83	0,68 (0,54-0,82)	0,07	0,19

DE = desviación estándar, ICC<sub>2,1</sub> = coeficiente de correlación intraclassa modelo 2-way random effect single measurement, IC = intervalo de confianza, SEM = error estándar de la medición, MDC<sub>95</sub> = cambio mínimo detectable con intervalo de confianza al 95%, cm = centímetros.

## Discusión

Los hallazgos de este estudio confirman que evaluadores experimentados de cuidados críticos pueden obtener excelente confiabilidad en la medición del grosor de músculos sanos (ICC = 0,82-0,87). Además, los kinesiólogos con poca o ninguna experiencia previa en ecografía pudieron realizar mediciones musculares con moderada a buena confiabilidad interobservador luego de seis horas de entrenamiento práctico, especialmente al medir grosor de cuádriceps de modelos sanos (ICC = 0,76) y calidad muscular de imágenes de personas críticamente enfermas usando la escala Heckmatt (ICC = 0,68).

Todos los valores de grosor muscular que obtuvieron los evaluadores estuvieron dentro de los valores de referencia de personas sanas y jóvenes<sup>17,21</sup>. La confiabilidad obtenida en nuestro estudio fue ligeramente inferior a la reportada por estudios previos en personas sanas (entre 0,73-0,99)<sup>22,23</sup> o personas críticamente enfermas (entre 0,76-0,96)<sup>24</sup>. Esto podría explicarse porque en nuestro estudio solo se utilizaron ecógrafos portátiles, los cuales pueden empeorar los resultados de confiabilidad comparado con dispositivos de alta gama<sup>25</sup>. Además, la mayoría de los estudios de confiabilidad interobservador han incluido entre 2 y 4 evaluadores<sup>22,26-28</sup>. En nuestro estudio, la confiabilidad se midió para 24 evaluadores, lo que podría explicar una mayor posibilidad de varianza y mayor dificultad en alcanzar acuerdos entre los evaluadores<sup>18</sup>. Por otro lado, se sabe que la confiabilidad del grosor muscular puede variar según el músculo evaluado, edad, lado del cuerpo y orientación del transductor<sup>29</sup>. El grosor muscular del cuádriceps se ha reportado en diferentes poblaciones, mostrando buena a excelente confiabilidad interobservador en adultos mayores<sup>30</sup>, pacientes con diabetes<sup>28</sup> y personas críticamente enfermas<sup>9,26,27</sup>. Adicionalmente, la información para el grosor del tibial anterior es limitada<sup>29,31,32</sup>, especialmente para personas críticamente enfermas<sup>26</sup>. No obstante, la medición del grosor del tibial anterior también ha mostrado ser una alternativa confiable para evaluar la pérdida de masa muscular<sup>31</sup>.

El grosor muscular medido por los evaluadores novatos fue sistemáticamente menor que el medido por los evaluadores experimentados, de hasta 6,4% menos de grosor muscular. Esto es

consistente con estudios previos que han mostrado que los evaluadores con mayor experiencia miden un mayor grosor<sup>22,26</sup> y mejor confiabilidad<sup>13,15</sup> que evaluadores novatos o estudiantes de kinesiología. En nuestro estudio, los novatos tuvieron menor confiabilidad, particularmente en recto femoral (ICC = 0,46) y tibial anterior (ICC = 0,41), los cuales son músculos más cercanos a la superficie. Hipotetizamos que el grosor muscular medido por los novatos fue menor por una mayor presión del cabezal sobre la piel comparada con los experimentados. La técnica de compresión mínima es parte de las recomendaciones de la mayoría de los protocolos de ecografía muscular esquelética<sup>6,8,27</sup>. Consecuentemente, Schiefflers y colaboradores mostraron que el uso de la técnica de compresión mínima para medir el grosor del cuádriceps se asoció a valores más altos de ICC y más bajos de SEM y MDC<sub>95</sub> comparado con la técnica de compresión máxima<sup>33</sup>. Por lo tanto, recomendamos fuertemente que en el entrenamiento en ecografía muscular esquelética se enfatice el aprendizaje de la técnica de compresión mínima, ya que la ausencia de ésta podría influir en la precisión de las mediciones. Aunque los programas de entrenamiento en ecografía de UCI aún no incluyen el aprendizaje en ecografía muscular esquelética<sup>14</sup>, progresivamente están emergiendo alternativas para sistematizar los protocolos de medición<sup>8,15,34</sup>.

En base a nuestro conocimiento, no existen otros estudios que reporten la confiabilidad del análisis cualitativo muscular usando la escala Heckmatt en personas críticamente enfermas. En nuestro estudio, los evaluadores novatos obtuvieron buena confiabilidad al puntuar imágenes de cuádriceps y tibial anterior (ICC = 0,68), lo que es comparable con otros estudios que evaluaron la confiabilidad de la calidad muscular usando la escala Heckmatt en gastrocnemio y sóleo de pacientes con parálisis cerebral (entre 0,60-0,73)<sup>35</sup>, en músculos de extremidades superiores de pacientes con esclerosis lateral amiotrófica (entre 0,70-0,91)<sup>36</sup>, en músculos flexores digitales de pacientes con accidente cerebrovascular (entre 0,73-0,74)<sup>37</sup> y en músculos de extremidad superior e inferior de pacientes con espasticidad (entre 0,43-0,67)<sup>38</sup>. En base a esta información, la escala Heckmatt podría utilizarse en la práctica clínica con buena confiabilidad independiente de la experiencia del evaluador.

Este estudio tiene limitaciones que merecen ser

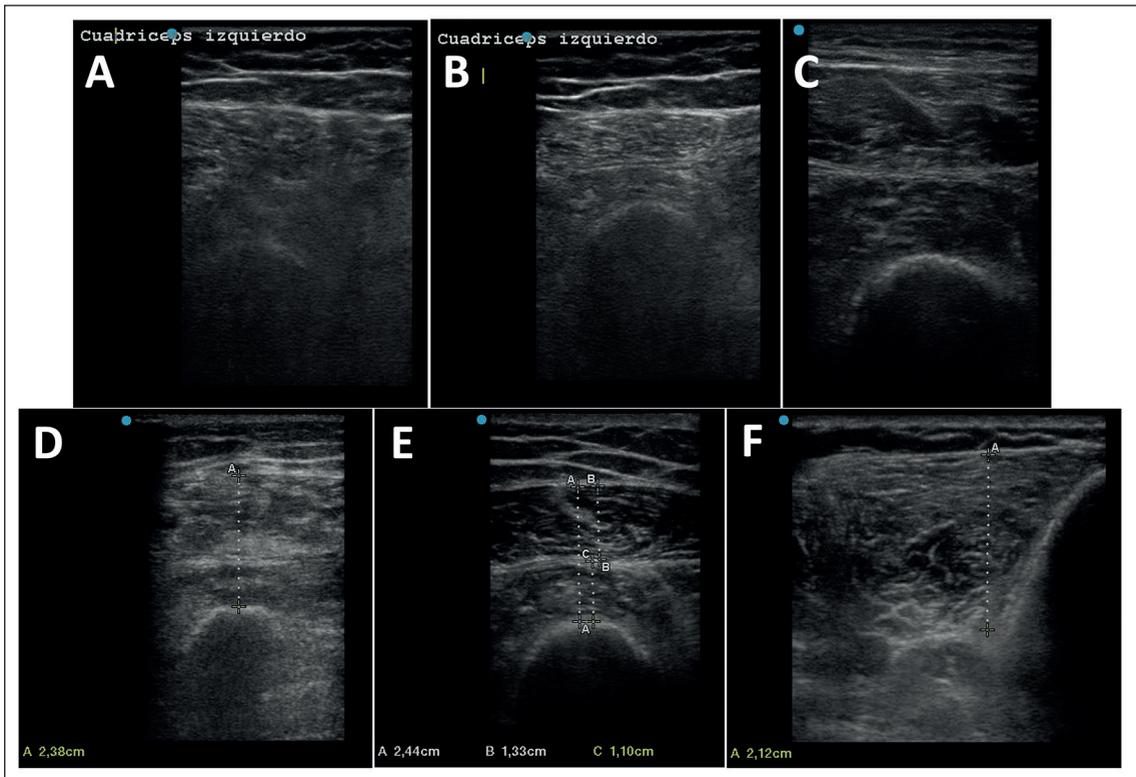
nombradas. Primero, las evaluaciones del grosor muscular se realizaron en individuos sanos, lo que no garantiza que los evaluadores tengan la misma confiabilidad en personas críticamente enfermas, quienes pueden presentar edema, inflamación o estar en posiciones como prono o decúbito lateral, limitando las mediciones. Sin embargo, la literatura disponible no ha mostrado grandes diferencias entre la confiabilidad en sanos y en personas críticamente enfermas<sup>26</sup>. Segundo, debido a la capacidad técnica de los ecógrafos y transductores disponibles en el estudio, la cantidad muscular solo se midió con el grosor muscular, por lo que los resultados de confiabilidad de este estudio no se pueden extrapolar a las mediciones del ángulo peniforme (corte longitudinal) y área de sección transversal que ha demostrado mejor correlación con fuerza y movilidad de personas críticamente enfermas<sup>39</sup>. Además, aunque utilizamos transductores lineales y convexos pudiendo haber modificado la precisión de las mediciones, experimentados como novatos estuvieron expuestos al uso de los mismos transductores. Tercero, la cantidad de evaluadores experimentados no fue la misma que la de evaluadores novatos, lo que podría haber modificado los valores de ICC. Sin embargo, el cálculo de tamaño muestral según las recomendaciones de Bonett<sup>18</sup> aseguró buena representatividad y varianza. Cuarto, la selección de evaluadores fue por conveniencia y la definición de los experimentados no está ampliamente aceptada como tal, lo que limita la generalización de los resultados. Finalmente, nuestro estudio agrega nueva información sobre la confiabilidad (precisión) de las mediciones ecográficas, pero no reporta la validez (exactitud) de las mediciones.

Los hallazgos de este estudio sobre la medición del grosor de cuádriceps y tibial anterior en personas sanas, podría contribuir para comparaciones con poblaciones de pacientes con pérdida de masa muscular. Actualmente, los valores de referencia de los parámetros ecográficos están limitados a

pocos países y poblaciones<sup>17,21</sup>. Particularmente, la ausencia de valores de referencia de grosor muscular limita la comparación de los resultados de las personas críticamente enfermas a lo largo de su estadía hospitalaria, ya que la masa muscular medida al ingreso de la UCI no siempre representa el nivel basal del paciente. Próximos esfuerzos nacionales deberían concentrarse en generar valores de referencia de los parámetros ecográficos considerando las variables que modifican la cantidad y calidad muscular.

En conclusión, la confiabilidad interobservador de los experimentados fue excelente y la de los novatos fue moderada a buena en las mediciones de grosor muscular y puntuaciones Heckmatt. Comparado con los evaluadores experimentados, los evaluadores novatos sistemáticamente midieron grosos musculares menores, tuvieron menores valores de ICC y mayores valores de SEM y MDC<sub>95</sub>. Futuros estudios deberían explorar si la técnica de compresión mínima reduce el error, ya que una mayor presión sobre la piel podría afectar la reproducibilidad de los resultados en la práctica clínica.

**Agradecimientos:** Los autores quieren agradecer la importante contribución de todos los evaluadores novatos de este estudio, ya que sin su participación no habría sido posible. Agradecemos a la Carrera de Kinesiología de la Universidad del Desarrollo por facilitar los espacios físicos para las mediciones. Agradecemos a Juan José Pinto-Concha, Alexis Silva-Gutiérrez y Agustín Camus-Molina por su contribución como evaluadores experimentados y docentes del entrenamiento de los evaluadores novatos de este estudio. Agradecemos el apoyo de Andover y TechMedicine por facilitar los ecógrafos usados en las mediciones. Un agradecimiento especial al Profesor Asociado Gabriel Cavada por su ayuda con los análisis estadísticos.



**Figura Suplementaria.** Ejemplos de ecografías musculares del banco de imágenes. **A)** Heckmatt grado 4 de Cuádriceps; **B)** Heckmatt grado 3 de Cuádriceps; **C)** Heckmatt grado 1 de Cuádriceps; **D)** Heckmatt grado 2 de Cuádriceps; **E)** Heckmatt grado 1 de Cuádriceps; **F)** Heckmatt grado 2 de Tibial Anterior.

## Referencias

1. Puthuchery Z, Rawal J, McPhail M, Connolly B, Ratnayake G, Chan P, et al. Acute skeletal muscle wasting in critical illness. *JAMA - Journal of the American Medical Association*. 2013;310: 1591-1600. doi:10.1001/jama.2013.278481
2. Mayer KP, Thompson Bastin ML, Montgomery-Yates AA, Pastva AM, Dupont-Versteegden EE, Parry SM, et al. Acute skeletal muscle wasting and dysfunction predict physical disability at hospital discharge in patients with critical illness. *Crit Care*. 2020;24: 637. doi:10.1186/s13054-020-03355-x
3. Puthuchery ZA, McNelly AS, Rawal J, Connolly B, Sidhu PS, Rowleron A, et al. Rectus Femoris Cross-Sectional Area and Muscle Layer Thickness: Comparative Markers of Muscle Wasting and Weakness. *Am J Respir Crit Care Med*. 2017;195: 136-138. doi:10.1164/rccm.201604-0875LE
4. Weijs PJM, Looijaard WGPM, Dekker IM, Stapel SN, Girbes AR, Straaten HMO, et al. Low skeletal muscle area is a risk factor for mortality in mechanically ventilated critically ill patients. *Crit Care*. 2014;18: 1-7. doi:10.1186/cc13189
5. Moisey LL, Mourtzakis M, Cotton BA, Premji T, Heyland DK, Wade CE, et al. Skeletal muscle predicts ventilator-free days, ICU-free days, and mortality in elderly ICU patients. *Crit Care*. 2013;17: R206. doi:10.1186/cc12901
6. Formenti P, Umbrello M, Coppola S, Froio S, Chiumello D. Clinical review: peripheral muscular ultrasound in the ICU. *Ann Intensive Care*. 2019;9. doi:10.1186/s13613-019-0531-x
7. Parry SM, El-Ansary D, Cartwright MS, Sarwal A, Berney S, Koopman R, et al. Ultrasonography in the intensive care setting can be used to detect changes in the quality and quantity of muscle and is related to muscle strength and function. *J Crit Care*. 2015;30: 1151.e9-1151.e14. doi:10.1016/j.jcrc.2015.05.024
8. Mourtzakis M, Parry S, Connolly B, Puthuchery Z.

- Skeletal muscle ultrasound in critical care: A tool in need of translation. *Ann Am Thorac Soc.* 2017;14: 1495-1503. doi:10.1513/AnnalsATS.201612-967PS
9. Connolly B, MacBean V, Crowley C, Lunt A, Moxham J, Rafferty GF, et al. Ultrasound for the Assessment of Peripheral Skeletal Muscle Architecture in Critical Illness. *Crit Care Med.* 2015;43: 897-905. doi:10.1097/CCM.0000000000000821
  10. Santos R, Armada-da-Silva PAS. Reproducibility of ultrasound-derived muscle thickness and echo-intensity for the entire quadriceps femoris muscle. *Radiography.* 2017;23: e51-e61. doi:10.1016/j.radi.2017.03.011
  11. Heckmatt JZ, Leeman S, Dubowitz V. Ultrasound imaging in the diagnosis of muscle disease. *J Pediatr.* 1982;101: 656-660. doi:10.1016/S0022-3476(82)80286-2
  12. Nijholt W, Scafoglieri A, Jager-Wittenaar H, Hobbelen JSM, van der Schans CP. The reliability and validity of ultrasound to quantify muscles in older adults: a systematic review. *J Cachexia Sarcopenia Muscle.* 2017;8: 702-712. doi:10.1002/jcsm.12210
  13. Carr JC, Gerstner GR, Voskuil CC, Harden JE, Dunnick D, Badillo KM, et al. The Influence of Sonographer Experience on Skeletal Muscle Image Acquisition and Analysis. *J Funct Morphol Kinesiol.* 2021;6: 91. doi:10.3390/jfmk6040091
  14. Wong A, Galarza L, Duska F. Critical Care Ultrasound. *Crit Care Med.* 2019;47: e256-e262. doi:10.1097/CCM.0000000000003626
  15. González-Seguel F, Pinto-Concha JJ, Ríos-Castro F, Silva-Gutiérrez A, Camus-Molina A, Mayer KP, et al. Evaluating a Muscle Ultrasound Education Program: Theoretical Knowledge, Hands-on Skills, Reliability, and Satisfaction of Critical Care Physiotherapists. *Arch Rehabil Res Clin Transl.* 2021;3: 100142. doi:10.1016/j.arrct.2021.100142
  16. Mokkink LB, Boers M, van der Vleuten CPM, Bouter LM, Alonso J, Patrick DL, et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Med Res Methodol.* 2020;20: 293. doi:10.1186/s12874-020-01179-5
  17. Arts IMP, Pillen S, Schelhaas HJ, Overeem S, Zwarts MJ. Normal values for quantitative muscle ultrasonography in adults. *Muscle Nerve.* 2010;41: 32-41. doi:10.1002/mus.21458
  18. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med.* 2002;21: 1331-1335. doi:10.1002/sim.1108
  19. Castro J, Livino de Carvalho K, Silva PE, Fachin-Martins E, Babault N, Marqueti R de C, et al. Intra- and inter-rater reproducibility of ultrasound imaging of patellar and quadriceps tendons in critically ill patients. Patman S, editor. *PLoS One.* 2019;14: e0219057. doi:10.1371/journal.pone.0219057
  20. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes.* 2006;4: 54. doi:10.1186/1477-7525-4-54
  21. Agyapong-Badu S, Warner M, Samuel D, Narici M, Cooper C, Stokes M. Anterior thigh composition measured using ultrasound imaging to quantify relative thickness of muscle and non-contractile tissue: a potential biomarker for musculoskeletal health. *Physiol Meas.* 2014;35: 2165-2176. doi:10.1088/0967-3334/35/10/2165
  22. Tillquist M, Kutsogiannis DJ, Wischmeyer PE, Kummerlen C, Leung R, Stollery D, et al. Bedside Ultrasound Is a Practical and Reliable Measurement Tool for Assessing Quadriceps Muscle Layer Thickness. *Journal of Parenteral and Enteral Nutrition.* 2014;38: 886-890. doi:10.1177/0148607113501327
  23. Abiko T, Ohmae K, Murata S, Shiraiwa K, Horie J. Reliability of muscle thickness and echo intensity measurements of the quadriceps: A novice examiner. *J Bodyw Mov Ther.* 2022;31: 164-168. doi:10.1016/j.jbmt.2022.03.004
  24. Hansell L, Ntoumenopoulos G. Clinimetrics: Quadriceps muscle ultrasound. *J Physiother.* 2021;67: 313. doi:10.1016/j.jphys.2021.05.001
  25. Ritsche P, Schmid R, Franchi M v., Faude O. Agreement and reliability of lower limb muscle architecture measurements using a portable ultrasound device. *Front Physiol.* 2022;13. doi:10.3389/fphys.2022.981862
  26. Mayer KP, Dhar S, Cassidy E, Denham A, England J, Morris PE, et al. Interrater Reliability of Muscle Ultrasonography Image Acquisition by Physical Therapists in Patients Who Have or Who Survived Critical Illness. *Phys Ther.* 2020;100: 1701-1711. doi:10.1093/ptj/pzaa068
  27. Pardo E, el Behi H, Boizeau P, Verdonk F, Alberti C, Lescot T. Reliability of ultrasound measurements of quadriceps muscle thickness in critically ill patients. *BMC Anesthesiol.* 2018;18: 205. doi:10.1186/s12871-018-0647-9
  28. Silva CR de S, Costa A dos S, Rocha T, Lima DAM de, Nascimento T do, Moraes SRA de. Quadriceps muscle architecture ultrasonography of individuals with type 2 diabetes: Reliability and applicability. *PLoS One.* 2018;13: e0205724. doi:10.1371/journal.pone.0205724
  29. Hagoort I, Hortobágyi T, Vuillerme N, Lamoth CJC, Murgia A. Age- and muscle-specific reliability of muscle architecture measurements assessed by two-dimensional

- panoramic ultrasound. *Biomed Eng Online*. 2022;21: 15. doi:10.1186/s12938-021-00967-4
30. Nijholt W, Scafoglieri A, Jager-Wittenaar H, Hobbelen JSM, van der Schans CP. The reliability and validity of ultrasound to quantify muscles in older adults: a systematic review. *J Cachexia Sarcopenia Muscle*. 2017;8: 702-712. doi:10.1002/jcsm.12210
  31. Leigh M, de Sire A, Colangelo M, Zagaria D, Grassi FA, Rena O, et al. Sarcopenia Diagnosis: Reliability of the Ultrasound Assessment of the Tibialis Anterior Muscle as an Alternative Evaluation Tool. *Diagnostics*. 2021;11: 2158. doi:10.3390/diagnostics11112158
  32. Miyachi R, Kanazawa Y, Fujii Y, Ohno N, Miyati T, Yamazaki T. Reliability of lower leg muscle thickness measurement along the long axis of the muscle using ultrasound imaging, in a sitting position. *J Phys Ther Sci*. 2022;34: 2022-029. doi:10.1589/jpts.34.515
  33. Schieffeler DR, Dombrecht D, Lafaire C, de Cuyper L, Rose T, Meirte J, et al. Reliability and feasibility of skeletal muscle ultrasound in the acute burn setting. *Burns*. 2022. doi:10.1016/j.burns.2022.03.003
  34. Ntoumenopoulos G, Parry SM, Neindre A le. Impact of an intensive education programme of diagnostic lung and lower limb ultrasound on physiotherapist knowledge: A pilot study. *Australas J Ultrasound Med*. 2018;21: 104-114. doi:10.1002/ajum.12089
  35. Battisti N, Milletti D, Miceli M, Zenesini C, Cersosimo A. Usefulness of a Qualitative Ultrasound Evaluation of the Gastrocnemius-Soleus Complex with the Heckmatt Scale for Clinical Practice in Cerebral Palsy. *Ultrasound Med Biol*. 2018;44: 2548-2555. doi:10.1016/j.ultrasmedbio.2018.08.006
  36. Kenis-Coskun O, Gunes T, Isak B, Yagci I. Intrarater and Interrater Reliability of Heckmatt Scoring System in Amyotrophic Lateral Sclerosis. *Journal of Clinical Neurophysiology*. 2021;38: 293-298. doi:10.1097/WNP.0000000000000690
  37. Kenis-Coskun O, Giray E, Gencer-Atalay ZK, Yagci I, Karadag-Saygi E. Reliability of quantitative ultrasound measurement of flexor digitorum superficialis and profundus muscles in stroke. *J Comp Eff Res*. 2020;9: 1293-1300. doi:10.2217/cer-2020-0105
  38. Moreta MC, Fleet A, Reebye R, McKernan G, Berger M, Farag J, et al. Reliability and Validity of the Modified Heckmatt Scale in Evaluating Muscle Changes With Ultrasound in Spasticity. *Arch Rehabil Res Clin Transl*. 2020;2: 100071. doi:10.1016/j.arrct.2020.100071
  39. Palakshappa JA, Reilly JP, Schweickert WD, Anderson BJ, Khoury V, Shashaty MG, et al. Quantitative peripheral muscle ultrasound in sepsis: Muscle area superior to thickness. *J Crit Care*. 2018;47: 324-330. doi:10.1016/j.jcrc.2018.04.003