Evaluating a muscle ultrasound education programme: theoretical knowledge, hands-on skills, reliability and satisfaction of critical care physiotherapists

Felipe González-Seguel (✉ feligonzalezs@udd.cl)

Servicio de Medicina Física y Rehabilitación and Departamento de Paciente Crítico, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo

Juan José Pinto-Concha

Centro de Paciente Crítico Adulto, Clínica INDISA

Francisco Ríos-Castro

Servicio de Medicina Física y Rehabilitación and Departamento de Paciente Crítico, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo

Alexis Silva-Gutiérrez

Unidad de Paciente Crítico Adulto, Hospital Clínico Herminda Martin

Agustín Camus-Molina

Servicio de Medicina Física y Rehabilitación and Departamento de Paciente Crítico, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo

Kirby P. Mayer

Department of Physical Therapy, College of Health Sciences, University of Kentucky

Selina M. Parry

Department of Physiotherapy, The University of Melbourne, Parkville, VIC

Abstract

Background: There is scarce evidence on learning outcomes of physiotherapists receiving formal muscle ultrasound training focused on the intensive care setting. The aim was to evaluate the theoretical knowledge, hands-on skills, satisfaction and reliability of critical care physiotherapists participating in a muscle ultrasound (MUS) educational programme.

Methods: A cross-sectional study involving 19 critical care physiotherapists with little to no prior experience in ultrasound who completed standardized 20-hours in MUS was performed over 2-week time period; including knobology, patient positioning, anatomical landmarks, image acquisition and limb muscle measurements. Theoretical knowledge, hands-on skills acquisition and satisfaction were assessed. Inter and intra-rater reliability on landmarks, thickness and pennation angle of quadriceps between participants was evaluated using intraclass correlation coefficient (ICC). Reliability among instructors measured prior to the course was also reported as reference.

Results: The percentage score (mean±SD) of knowledge questionnaires was 69±0.11 (pre-course), 89±0.10 (post-course) and 92±0.09 (hands-on skills). Course satisfaction scores were ranged from 90% to 100%. Participants obtained pooled inter-rater reliability (median ICC [IQR]) of 0.70 [0.59 to 0.79] for thickness, 0.47 [0.46 to 0.92] for landmarks, 0.00 [0.00 to 0.05] for pennation angle; and intra-rater reliability of 0.76 [0.51 to 0.91] for thickness, and 0.35 [0.29 to 0.52] for pennation angle. Instructors obtained inter-rater ICC values of 0.90 for thickness, 0.67 for landmarks and 0.41 for pennation angle; and intra-rater ICC values of 0.94 for thickness and 0.75 for pennation angle.

Conclusions: Increasing on theoretical knowledge, high hands-on performance acquisition, and good satisfaction of physiotherapists were observed. Reliability was moderate to excellent for thickness and landmarks, and absent to weak for pennation angle. Landmarking and pennation angle remain challenges for physiotherapist training in the application of muscle ultrasound. Further studies are needed to identify variables that could modify the reliability during MUS training.

Background

Muscle wasting occurs during a variety of disease and illness states, including patients with critical illness admitted to the intensive care unit (ICU)[1−3]. Studies using movement sensor technology have reported that critically ill patients remain inactive for 92% to 98% of their ICU stay[4−7]. Immobility and disuse are primary reasons patients in ICU suffer rapid and early muscle wasting in rectus femoris muscle size during the first 10 days of an ICU admission[1−3]. Outcome measures that allow early identification of musculoskeletal and physical dysfunction are increasingly used in critically ill patients[8−10]. However, the majority of outcome measures are volitional in nature requiring the patient to be alert and cooperative with testing which often results in a delay in identifying those individuals at highest risk of musculoskeletal and physical dysfunction[10]. In recent years, muscle ultrasound has

gained traction as a potential tool that can be utilized early during critical illness without the need for volitional patient effort to assess changes in skeletal muscle[11].

Muscle ultrasound can be performed at the bedside, and is non invasive and readily available, showing high clinical and research utility[12]. Muscle ultrasound can be used to detect changes in the trajectory of muscle mass quality and quantity[2, 12], and potentially when combined with physical functioning measures may enable tracking of progress and evaluation of the effectiveness of rehabilitation interventions for critically ill patients[11, 13]. Muscle ultrasound measurements have robust clinimetric properties in critically ill patients[2, 14–16], including excellent validity and reliability when standardized training has been performed[15–18]. Muscle ultrasound allows the evaluation of muscle quantity (muscle thickness, cross-sectional area), and quality/ biomechanical properties (fiber pennation angle, fascicle length and echogenicity)[11, 19]. Good to excellent reliability has been reported for measurement of muscle thickness, cross-sectional area and echogenicity within critically ill patients[15, 17, 18]. However, reliability of ultrasonographer landmarking and measurement of fiber pennation angle have not been reported for patients with critical illness[12]. Patient setup including standardized postioning and the correct anatomical landmarks are vital components of acquiring an accurate ultrasound image[11]. This is particuraly important for skeletal muscle as small variations in setup and landmarking may lead to significant differences in anatomical location and impact on the acquisition process[18]. Thus, standardized patient positioning and identification of landmarks are essential to minimize differences between evaluators and when assessing for change over time.

Due to operator dependence of muscle ultrasound, international recommendations suggest that the accurate use of ultrasound requires standardized training combining theoretical knowledge and hands-on practical skills[20–22]. There are eight international ultrasound training programs for the ICU setting primarily focused on physicians and radiologist, not including ultrasonography for physiotherapists[22]. Ultrasound training for acute care physiotherapists have been previously conducted[23, 24], however formal training on peripheral skeletal muscle mass assessment does not exist in Chile. For this reason, there is scarce evidence on learning outcomes of physiotherapists receiving muscle ultrasound training which can be applied within the ICU setting.

Accordingly, the aim of this study is to evaluate the change in pre and post course theoretical knowledge, practical skills, and satisfaction of ICU physiotherapists during a bespoke muscle ultrasound education programme focused on ICU patients, and to determine the inter-rater and intra-rater reliability measurement of landmarks, muscle thickness and fiber pennation angle.

Materials And Methods

Study design

A cross-sectional observational study was conducted to evaluate the results of the first Chilean muscle ultrasound education programme focused on ICU patients (eMUSICS, stand for: Education in Muscle Ultrasound for Intensive Care Setting) performed between August 19th and 31st, 2019

(https://bit.ly/2PvTWzt) in the faculty of medicine of a Chilean private university, which has a structured continuing education platform. This study received review and approval after course development by the research ethics committee of Universidad del Desarrollo (registration number 2020-106).

Instructors

Five critical care physiotherapists (ACM, JJP, FRC, ASG and FGS) with more than two years of ultrasound clinical experience led the eMUSICS. Instructors had received formal training from one of the following three alternatives: 1) Ultrasound in Emergency and Critical Care (USECC) course of the Sociedad Chilena de Medicina Crítica y Urgencias, 2) Ecografía en el Paciente Crítico as a 2017 pre-congress workshop at the Congreso Argentino of the Sociedad Argentina de Terapia Intensiva, or 3) Diagnostic ultrasound course as a 2017 post-congress workshop at the World Confederation of Physical Therapy Conference[23]. Two months before the course, all instructors carried out two 3-hour training meetings amongst themselves to standardize the methodology of image acquisition and measurement. The aim of this study was to evaluate the participants, however, the intra and inter-reliability of instructors performed prior to course initiation were presented in the results section for transparency. Reliability measurement procedure of instructors is presented in **Additional file, Table S1**.

Programme description

A 20-hour muscle ultrasound education programme was developed based on international training recommendations[20–22, 25]. Selected contents and materials from an ultrasound post-congress workshop developed in the 2017 World Confederation of Physical Therapy Conference[23] were used in this programme with prior authorization (SMP), involving ultrasound physics, knobology, muscle anatomy and physiology, patient positioning, landmarks, image acquisition and lower limb measurements of muscle quality and quantity. The course was designed for rehabilitation clinicians with at least 1 year of working experience with ICU patients. The ultrasound-training course was delivered in two parts – the first was online learning (eLearning) and the second in-person learning including a combination of lectures and hands-on practice (**Additional file, Table S2** provides a detailed programme curriculum).

eLearning training and assessment

The eLearning was available for participants 2 weeks before the in-person course using the Moodle™ online platform with an estimated dedication time of 6 hours. The objectives of the eLearning were to identify the normal anatomical structures with real ultrasound images and to understand initial concepts of muscle structure measurements using ultrasound. The online platform included recommended pre-reading material, ultrasound machines technical manuals, an instructor-led discussion forum to answer questions, and four training videos (for more details on training videos see **Additional file, Table S3**). To guarantee the participant's knowledge before the in-person course, a formative questionnaire with 20-question open-ended based on the observation of normal muscle ultrasound images was performed,

which was answered using the pre-reading material provided. Participants submitted their answers until one day before the in-person course using the online platform.

In-person course

After the eLearning, the in-person course was performed during 2 consecutive days including 5-hours of didactic lectures and 9-hours of hands-on training led by 5 trained instructors. Lectures were conducted with a projector/screen using standard slide sets or real-time ultrasound muscle scanning of a participating volunteer. During hands-on training, 10 ultrasound machines were available, including 9 wireless (Philips Lumify, Sonus SL-2C, and Sonus DUO LCP) connected to an iPad; and one portable (Philips InnoSight, including linear and curvilinear array transducer). To optimize participant learning, a maximum of 20 participants was defined to achieve an instructor to trainee ratio of 1:4 and ultrasound machine to trainee ratio of 1:2 being more than the recommended ratio of 1:5 by the Australian Society of Ultrasound Medicine (see link: http://www.asum.com.au/files/public/Education/CAHPU/CAHPUForms/CAHPU-Unit-Accreditation-Application-Form.pdf). Knobology, patient positioning, landmarks, image acquisition, muscle thickness, and fiber pennation angle measurement were studied through lectures and hands-on training. The hands-on training was designed for each participant to perform 15 landmarks identifications and 25 supervised muscles scans including all ultrasound parameters involved in this study using the quadriceps measurements of the same participants.

Participant assessments: Theoretical knowledge was evaluated using a pre-course diagnostic questionnaire at the beginning of the in-person course. At the end of the course, theoretical knowledge was re-evaluated through a post-course formative questionnaire using the same questions in a randomly assigned order. Both questionnaires included 25 multiple-choice questions of which 11 were obtained from the questionnaire of Ntoumenopoulos *et al*[23].

Hands-on assessment was performed during the last 3 hours of the course to assess the practical skills of participants. Due to class size and timing convenience, participants were split into groups to perform the hands-on assessment and inter-rater reliability. Using the ©Intemodino RNG as random number generator software, participants were distributed into 5 groups of 4 people. The participants were blind to each other's measurements to assess: quadriceps landmarks; rectus femoris (RF), vastus intermedius (VI), quadriceps complex (QC) thickness; and/or vastus lateralis (VL) pennation angle in 6 healthy individuals. The ultrasound machine to trainee ratio was 1:2, and each instructor qualitatively evaluated four participants through direct observation using a non-middle answer category Likert scale from 1 to 4 points (1 = poor, 2 = fair, 3 = good, 4 = excellent)[26]. This scale was used to evaluate the performance of the following items: patient positioning, landmarks identification, knobology operation (accurate use of gain compensation, zoom, focus, depth, freeze function and caliper), image acquisition (anatomy identification), transducer placement and quadriceps measurements. Participants measured the landmarks and muscle thickness in centimeters and pennation angle in degrees on a pre-specified written record document. Participants were previously trained to capture all ultrasound images directly on the

ultrasound machine, and subsequently instructors exported without any adjustments to a computer for analysis using a data storage device. This data was used to evaluate inter and intra-rater reliability of participants according to the assigned group. Reliability measurement procedure of participant is presented in **supplemental material, Table S1**, and was informed following the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) standards to assess the quality of studies on reliability and measurement error[27].

Course measurement protocol of landmarks, muscle thickness and pennation angle

Participants were examined in the supine position with neutral rotation and passive extension of lower limb[2]. Anatomical landmarks were marked using an erasable skin marker and measured in centimeters using a flexible tape measure. Quadriceps landmark was defined as the midpoint between the anterior superior iliac spine and the superior patella border[28, 29]. Tibialis anterior landmark was defined as one third of the distance from the tibial plateau to the inferior border of lateral malleolus[28, 29]. Depending on specific thigh size a B-mode with linear (4 to 12 MHz) and curvilinear (5 to 2 MHz) array transducers was used. When required, participants adjusted gain compensation, zoom, focus, depth, and freeze function. A generous amount of contact gel was used to minimize the required pressure of the transducer on the skin, allowing the minimal compression technique. Scans were performed with the transducer in neutral tilt using a transverse cross-sectional view for the muscle thickness and a sagittal view for the pennation angle. The transverse cross-sectional view was acquired to measure RF, VI and QC muscle thickness[30–32]. Muscle thickness were reported in centimeters using the caliper of the ultrasound machine as the inside height measured between epimysial borders of each muscle[2] (**Additional file, Figure S1**). To acquire an accurate sagittal view for VL pennation angle, the transducer was moved laterally 5 centimeters from the site where RF/VI was obtained. Pennation angle of the VL was reported in degrees as the angle between direction of muscle fibers and force line action represented by external tendon or aponeurosis i.e., the vertical inclination of fibers from the long axis of muscle[33], using the average of three consecutive separate attempts (**Additional file, Figure S2**).

Course satisfaction

At the end of the second day of the in-person course, participants were asked to voluntarily answer a standardized anonymous satisfaction survey predesigned by the local university educational programme and used for all courses related to medicine. This survey evaluates the participant perception including the overall assessment of the course, academic scope, instructor quality, eLearning, pre-reading material and course coordination. Each item was scored using a non-middle answer category Likert scale from 1 to 4 points (1 = Strongly disagree, 2 = Disagree, 3 = Agree, 4 = Strongly agree)[26]. Additionally, this survey included the following two yes-no questions: would you recommend these instructors for a future course?, and would you recommend this program to other people?.

Statistical analyses

Descriptive data were analyzed using STATA SE 15.0 (StataCorp, LLC, 2017, College Station, TX). The normality of the data of each variable was analyzed with the Shapiro-Wilk test. Data were described as mean ± standard deviation (SD) or median (interquartile range [IQR]), depending on the normality of the data. The limit of statistical significance was set at two-sided p value of ≤0.05. The scores of knowledge and practical skills questionnaires and satisfaction survey were reported as percentage correct. Reliability of participants and instructors was calculated using the intraclass correlation coefficient (ICC) and 95% CI according to Koo *et al*[34]. Repeated measurements by the same rater on the same subject were used to calculate intra-rater reliability that included a brief period with the instructor removing previous test results. Repeated measurements by different raters on the same subject were used to calculate inter-rater reliability, while raters were blinded to the test and results of the other raters. Inter-rater reliability was calculated using the 2-way random effects, absolute agreement, average measure of the number of records; and intra-rater reliability was calculated using the 2-way mixed effects, absolute agreement, single measure (**Additional file, Table S3**). Group comparisons were not performed, as this was not the focus of the study. To obtain the overall inter-rater reliability of the five groups of participants, the median and IQR of the ICC values of each ultrasound parameter was calculated. The qualitative interpretation of the ICC was classified as 0.00 (absent), 0.00−0.19 (poor), 0.20−0.39 (weak), 0.40−0.59 (moderate), 0.60−0.79 (good), and ≥0.80 (excellent)[35]. The standard error of measurement (SEM) was calculated for each ICC value as the product of standard deviation and the square-root (1 − single measures ICC). Coefficient of variation (CV) was calculated using the division of the SD by the mean value of the repeated measures (CV= [SD/mean]*100); interpreted as CV<10 (very good), 10-20 (good), 20-30 (acceptable), and CV>30 (not acceptable)[36, 37].

Results

Participants

Nineteen physiotherapists were enrolled in the course (**Figure 1**), and demographics of instructors and participants are provided in **Table 1**. Only one (5.3%) had previously received specific muscle ultrasound training and 11 (57.9%) had never participated in any ultrasound training prior to this workshop.

Table 1. Characteristics of instructors and participants of eMUSICS training programme

Variables	Instructors n = 5	Participants n = 19
Age, years	31 [29 to 32.5]	32 [29 to 34]
Female	0 (0.0)	6 (31.5)
Years working in ICU	8 [4 to 8.5]	5 [4 to 9]
Type of hospital		
Private	4 (80)	10 (52.6)
Public	1 (20)	9 (47.4)
Highest academic degree		
Bachelor	0 (0.0)	2 (10.5)
ICU diploma*	3 (60)	14 (73.7)
Master	2 (40)	2 (10.5)
PhD	0 (0.0)	1 (5.3)
ICU PT or RT certification	3 (60)	1 (5.3)
Previous ultrasound training	5 (100)	7 (36.8)
Previous muscle ultrasound training	2 (40)	1 (5.3)

Data are presented as median (interquartile range) or as count (percentage).

* Includes any 1-year postgraduate course related to intensive care.

ICU = intensive care unit, PhD = doctor of philosophy, PT = physiotherapy, RT = respiratory therapist

eLearning assessment

The eLearning questionnaire (mean ± SD) showed 89.1% ± 0.1 when participants had the option to respond using the pre-reading material before the in-person course.

In-person course assessments

The percentage score (mean ± SD) of the theoretical knowledge was 69.0% ± 0.1 and 88.9% ± 0.1 for the pre-course and post-course questionnaires, respectively. The percentage score (mean ± SD) of practical skills during hands-on assessment was 91.5% ± 0.1, and the best and worst evaluated items were positioning of the patient (100% ± 0.0) and quadriceps measurements using caliper or angle function (84.2 ± 0.8), respectively (**Table 2**). The results of the theoretical knowledge and practical skills of participants are presented in **Table 3**.

Table 2. Practical skills evaluated during hands-on assessment (n = 19)

Evaluated parameter	Likert scale[*]				Percentage score[†]
	Poor	Fair	Good	Excellent	
Patient positioning	0 (0.0)	0 (0.0)	0 (0.0)	19 (100)	100 ± 0.0
Landmarks	0 (0.0)	0 (0.0)	1 (5.3)	18 (94.7)	98.6 ± 0.2
Knobology	0 (0.0)	2 (10.5)	5 (26.3)	12 (63.2)	88.1 ± 0.7
Image acquisition	0 (0.0)	0 (0.0)	9 (47.4)	10 (52.6)	88.1 ± 0.5
Transducer placement	0 (0.0)	1 (5.3)	8 (42.1)	10 (52.6)	86.8 ± 0.6
Quadriceps measurements	1 (5.3)	1 (5.3)	7 (36.8)	10 (52.6)	84.2 ± 0.8

Data are presented as count (percentage) for the Likert scale and as mean ± standard deviation for percentage score.

[*] Practical skills were evaluated through subjective direct observation of one instructor using the following Likert scale: 1 = Poor (the skill is not fulfilled or appears full of imperfections throughout the process); 2 = Fair (the skill is partially accomplished, with numerous imperfections limiting the process); 3 = Good (the skill is fulfilled almost entirely, with some imperfections that can be corrected); 4 = Excellent (the skill is accomplished in an outstanding and sustained way throughout the process).

[†] Percentage score was calculated by dividing the score obtained by the total score, multiplied by 100 (percetage correct).

Table 3. Percentage scores of theoretical knowledge and practical skills per group and overall during eMUSICS training programme[*]

	Group 1 (n = 4)	Group 2 (n = 4)	Group 3 (n = 4)	Group 4 (n = 4)	Group 5 (n = 3)	Overall (n = 19)
eLearning	86.0%	93.3%	91.5%	87.0%	87.3%	89.1%
Theoretical knowledge						
Pre-course questionnaire	72.5%	73.0%	72.3%	63.5%	62.0%	69.0%
Post-course questionnaire	82.0%	86.0%	96.8%	88.3%	92.7%	88.9%
Post-pre difference	9.5%	12.8%	24.8%	24.5%	31.0%	19.9%
Practical skills (Likert scale)	88.8%	81.0%	100.0%	93.0%	96.0%	91.5%
Final course score	85.8%	85.0%	97.3%	90.3%	93.0%	90.1%

* Percentage score was calculated by dividing the score obtained by the total score, multiplied by 100 (percetage correct).

Reliability measurements

The on-site reliability measurement process during practical evaluation is presented schematically in **Figure 2**. The highest inter and intra-rater reliability of instructors was for the QC thickness, which was 0.90 and 0.94, respectively. The lowest inter and intra-rater ICC values of instructors were for VL pennation angle, which was 0.41 and 0.75, respectively. The reliability of participants ranged widely according to the assigned group. All inter and intra-rater reliability results of instructors and participants are presented in **Table 4**. There was moderate to excellent inter-rater reliability (median [IQR]) for VI thickness (0.87 [0.73 to 0.91]), QC thickness (0.70 [0.59 to 0.79]), landmarks (0.47 [0.46 to 0.92]) and RF thickness (0.41 [0.13 to 0.66]), and only absent for VL pennation angle (0.00 [0.00 to 0.05]). There was good to excellent intra-rater reliability (median [IQR]) for RF thickness (0.84 [0.69 to 0.87]), VI thickness (0.78 [0.73 to 0.85]) and QC thickness (0.76 [0.51 to 0.91]), and only weak for VL pennation angle (0.35 [0.29 to 0.52]). Overall, moderate to excellent reliability was obtained for muscle thickness and moderate for landmarks. Similar to instructors, the lowest inter and intra-rater reliability ICC values of participants were obtained for VL pennation angle, ranged from 0.00 to 0.10 and 0.23 to 0.69, respectively.

Table 4. Inter and intra-rater reliability of instructors and participants of eMUSICS training programme

	Muscle ultrasound parameter	Mean ± SD	CV (%)	Inter-rater Reliability ICC (95%CI)†	SEM	Intra-rater Reliability ICC (95%CI)†	SEM
Instructors*	Landmarks (cm)	23.19 ± 0.8	3.5	0.67 [0.48 to 0.87]	0.10	-	-
	RF thickness (cm)	2.21 ± 0.3	14.5	0.83 [0.71 to 0.95]	0.06	0.93 [0.90 to 0.96]	0.02
	VI thickness (cm)	1.93 ± 0.6	31.0	0.89 [0.81 to 0.97]	0.04	0.92 [0.88 to 0.95]	0.02
	QC thickness‡ (cm)	4.35 ± 0.9	19.9	0.90 [0.83 to 0.98]	0.04	0.94 [0.92 to 0.97]	0.01
	VL pennation angle (degrees)	16.43 ± 3.1	18.8	0.41 [0.13 to 0.70]	0.14	0.75 [0.66 to 0.85]	0.05
Group 1	Landmarks (cm)	25.12 ± 4.2	16.6	0.12 [0.00 to 0.58]	0.23	-	-
	RF thickness (cm)	1.8 ± 0.3	14.7	0.00 [0.00 to 0.00]	0.00	0.86 [0.78 to 0.93]	0.04
	VI thickness (cm)	1.94 ± 0.3	16.8	0.41 [0.04 to 0.79]	0.19	0.75 [0.62 to 0.87]	0.06
	QC thickness‡ (cm)	3.94 ± 0.6	14.6	0.39 [0.00 to 0.77]	0.20	0.89 [0.83 to 0.95]	0.03
	VL pennation angle (degrees)	15.28 ± 4.6	30.3	0.00 [0.00 to 0.00]	0.00	0.69 [0.54 to 0.84]	0.08
Group 2	Landmarks (cm)	22.91 ± 1.5	6.4	0.92 [0.84 to 0.99]	0.4	-	.
	RF thickness (cm)	1.98 ± 0.2	8.8	0.17 [0.00 to 0.60]	0.2	0.34 [0.06 to 0.62]	0.14
	VI thickness (cm)	1.77 ± 0.3	19.4	0.84 [0.70 to 0.98]	0.07	0.66 [0.49 to 0.82]	0.08
	QC thickness‡ (cm)	3.88 ± 0.5	12.3	0.74 [0.53 to 0.95]	0.12	0.19 [0.00 to 0.52]	0.17
	VL pennation angle (degrees)	14.73 ± 2.9	20.2	0.00 [0.00 to 0.00]	0.00	0.35 [0.07 to 0.63]	0.14
Group 3	Landmarks (cm)	23.66 ± 1.4	5.8	0.96 [0.91 to 1.00]	0.02	-	-
	RF thickness (cm)	2.03 ± 0.4	20.5	0.69 [0.45 to 0.94]	0.12	0.90 [0.84 to 0.95]	0.03
	VI thickness (cm)	1.67 ± 0.6	36.1	0.95 [0.89 to 1.00]	0.03	0.96 [0.93 to 0.98]	0.01
	QC thickness‡ (cm)	3.83 ± 1.0	26.4	0.95 [0.90 to 1.00]	0.02	0.96 [0.94 to 0.98]	0.01
Group 4	Landmarks (cm)	23.23 ± 1.7	7.4	0.46 [0.09 to 0.82]	0.18	-	-
	RF thickness (cm)	2.12 ± 0.3	14.9	0.65 [0.37 to 0.92]	0.14	0.81 [0.72 to 0.91]	0.05
	VI thickness (cm)	1.64 ± 0.3	19.0	0.89 [0.79 to 0.99]	0.05	0.81 [0.71 to 0.90]	0.05
	QC thickness‡ (cm)	3.84 ± 0.6	14.5	0.66 [0.40 to 0.92]	0.13	0.62 [0.46 to 0.78]	0.08
Group 5	Landmarks (cm)	23.77 ± 1.1	4.8	0.47 [0.05 to 0.90]	0.21	-	-
	VL pennation angle	19.55	11.2	0.10 [0.00 to	0.28	0.23 [0.01 to	0.11

	(degrees)	± 2.2	0.65]	0.45]

* Four instructors evaluated their reliability two months before muscle ultrasound education programme implementation.

† ICCs values range from 0.00 to 1.00

‡Quadriceps complex thickness includes the rectus femoris and vastus intermedius thickness.

RF = rectus femoris, VI = vastus intermedius, QC = quadriceps complex, VL = vastus lateralis, ICC = intraclass correlation coefficient, CI = confident interval, SEM = standard error of measurement, SD = standard deviation, CV = coefficient of variation

Course satisfaction

The mean percentage scores of the individual evaluated items in the satisfaction survey ranged from 90%±0.6 to 100%±0.0 on the Likert scale (**Additional file, Table S4**). The lowest evaluations were for the friendliness of the eLearning platform and for the pre-reading material. Additionally, when participants were asked if they would recommend the course instructors, all 19 (100%) answered yes. When the participants were asked if they would recommend this course to other people, 16 (84%) answered yes, 1 (5.3%) answered no and 2 (10.5%) left it blank.

Discussion

This study reported the acquisition of theoretical knowledge and hands-on skills of 19 ICU physiotherapists with little to no prior experience in muscle ultrasound, who received training through a dedicated bespoke muscle ultrasound education programme. This is the first report of ICU physiotherapists evaluating the quadriceps landmarks reliability and pennation angle. Compared with the median reliability values of participants, the instructors had higher reliability in the measurement of all muscle ultrasound parameters. Overall, the reliability of participating physiotherapists was moderate to excellent for muscle thickness, moderate for landmarks, and absent to poor for pennation angle. Instructors had excellent reliability for muscle thickness (inter-rater ICC ranged from 0.83 to 0.90; intra-rater ICC ranged from 0.92 to 0.94), good for landmarks (inter-rater ICC: 0.67), and moderate for pennation angle (inter-rater ICC: 0.41; intra-rater ICC: 0.75).

The eMUSICS training programme in this study was developed based on available international recommendations[20–22, 25], including online/digital media, hands-on sessions led by instructors, theoretical assessments, face-to-face practical assessments and a logbook of scans performed by participants. The results of ultrasound learning for acute care physiotherapists have been published in only two articles[23, 24]. In a one-day course dedicated to thoracic diagnostic ultrasound for physiotherapists there was a 13% increase on the post-pre training theoretical knowledge[24], while in other one-day course dedicated to lung, diaphragm and limb muscles ultrasound there were no

increase[23]. The eMUSICS course included limb muscle ultrasound exclusively, obtaining almost 20% increase on the theoretical knowledge of participants. According to a systematic review of international ultrasound training competencies and programmes, the first step in acquiring any new skill is the need to assess subsequent practical learning with theoretical knowledge as a foundation, but little practical assessments has been reported in published courses[22]. Compared to a general critical care ultrasonography course[38], eMUSICS programme obtained a similar mean percentage practical score (86% vs. 92%, respectively), as a result of supervised practical sessions. Although the number of measurements or scans in training courses is not yet standardized, critical care ultrasonography training of various national professional societies and organizations have reported a supervised or no-supervised number of scans ranged widely from 10 to 100[22]. During the eMUSICS programme 25 supervised muscle scans, including all muscle ultrasound parameters per participant, were achieved during training. By adding the muscle measurement performed during the whole practical assessment, at least 70 muscle scans per participant were performed during the course. This number of scans is consistent with non-physiotherapists ICU ultrasound training courses[22].

Ultrasound for the assessment of peripheral skeletal muscle architecture in critical illness has revealed excellent reliability levels, including parameters as thickness, cross-sectional area and echogenicity[12]. Particularly, the reliability of ultrasound measurement of quadriceps thickness in ICU patients has been widely reported ranged from 0.76 to 1.00[15–18, 39–42], and consistently in the eMUSICS programme was 0.70 by participants and 0.90 by instructors. Although reliability of pennation angle has not been reported for ICU patients, studies in non-ICU population have reported weak to moderate reliability and high variability, ranged from 0.38 to 0.74[43–45]. In the eMUSICS programme, instructors and participants achieved an inter-rater reliability of VL pennation angle between 0.00 and 0.41, showing the lowest reliability values of this study. Pennation angle in this study had high inter-rater variability in healthy muscles, and thus we suggest it could be worse in ICU patients due to loss of muscle echogenicity increasing the difficulty of visual interpretation of the muscle fibers.

Mourtzakis *et al*. highlighted the importance of landmarking on training, standardization, reporting in articles and reliability assessment[11]. Not all ultrasound studies have reported the quadriceps landmarks used, which could modify the accuracy of muscle ultrasound parameter depending on the measurement site. For example, Pardo *et al*. reported an inter-rater reliability of quadriceps thickness of 0.76 using the midpoint site and 0.81 using the two-thirds site[18]. In the eMUSICS programme, the quadriceps landmarks were evaluated using the midpoint in centimeters and the inter-rater reliability was 0.47 and 0.67 for participants and instructors, respectively. Thus, ultrasound measurements could be modified by the select of a specific landmark or by the accuracy of the evaluators to identify that landmark. More studies are needed to evaluate the reliability of landmarking because this could influence the consecutive measurements of thickness, cross-sectional area, echogenicity, and pennation angle.

Although inter and intra-rater reliability obtained by the instructors was consistent with preceding reliability studies of muscle architecture[15–18, 39–42], there was varied reliability obtained between participants. Overall, the highest inter and intra-rater reliability was found in the group 3 ranged from 0.69

to 0.96 and 0.90 to 0.96, respectively. The lowest inter-rater reliability was found in the group 1 ranged from 0.00 to 0.41; and the lowest intra-rater reliability was obtained in the group 2, ranged from 0.19 to 0.66. The differences between groups on reliability results could be explained by the ultrasound devices used, by the muscle ultrasound parameters assessed and by the number of repeated measures used per group (**supplemental material, Table S3**). Another potential rationale for overall lower inter- and intra-reliability may be the relative years of experience of our instructors, which is slightly lower than previously reported studies[15, 17]. Additionally, the learning performance was varied between groups. Although all participants received the same training and were randomized into five groups, the groups of participants with lowest scores on the theoretical and practical assessments (groups 1 and 2) had lowest reliability values as presented in **Table 3** and **Table 4**. Specifically, randomization concentrated the participants with the best scores in theoretical and practical assessments in the group 3, which likely led to the best reliability values. Further studies should explore if the level of knowledge acquisition during muscle ultrasound training may modify the reliability of participants.

Limitations

This study has limitations that deserve to be mentioned. The reliability assessment was performed only on healthy models, which does not guarantee that the participants have the same reliability when evaluating ICU patients. This was done to facilitate the early learning using normal anatomy as a reference. The small sample sizes per group limits the external validity of the reliability data. However, this study provides novel information on the reliability of landmarks and pennation angle. A longer washout period would strengthen the methodology on intra-rater reliability; however, the timing in this study provides preliminary data supporting further analyses on intra-rater reliability. Another limitation was the absence of certified expert instructors, which could have resulted in higher reliability measurement in the participants. To counter this, an exhaustive training process of instructors was carried out to ensure accurate instruction during in-person course. This study did not explore the durability of participants training and if the acquired knowledge was translated into effective clinical practice to assess muscle mass in ICU patients during ICU stay.

Conclusions

A novel muscle ultrasound course focused on ICU patients demonstrated increasing on theoretical knowledge, high hands-on performance acquisition and good satisfaction in a group of critical care physiotherapists with little to no prior experience in ultrasound. Despite the high theoretical results of the participants, this did not ensure high reliability in all ultrasound parameters. Overall, instructors and participants reached moderate to excellent inter and intra-rater reliability for quadriceps thickness, but reliability of landmarks and pennation angle remains challenging for physiotherapist training even for instructors. Further studies are needed to identify the variables that could modify the reliability results during muscle ultrasound training programmes.

List Of Abbreviations

ICU: intensive care unit; SD: standard deviation; ICC: intraclass correlation coefficient; CV: coefficient of variation; RF: rectus femoris; VI: vastus intermedius; QC: quadriceps complex; VL: vastus lateralis; PA: pennation angle; MUS: muscle ultrasound

Declarations

Author Contributions

Concept/idea/research design: FGS, JJPC

Writing: FGS, KPM, SMP

Data collection: FGS, JJPC, FRC, ACM, ASG

Data analysis: FGS, JJPC

Project management: FGS, JJPC

Providing institutional liaisons: FGS, KPM

Consultation (including review of manuscript before submitting): FGS, JJPC, FRC, ACM, ASG, KPM, SMP

Acknowledgements

Ethics approval and consent to participate

This study received review and approval after course development by the research ethics committee of Universidad del Desarrollo (registration number 2020-106). All participants provided written consent before commencing the in-person course.

Funding

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Puthucheary ZA, Rawal J, McPhail M, Connolly B, Ratnayake G, Chan P, et al. Acute Skeletal Muscle Wasting in Critical Illness. JAMA. 2013;310:1591.

2. 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. Journal of Critical Care. 2015;30:1151.e9-1151.e14.

3. 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. Critical Care. 2020;24:637.

4. Camus-Molina A, González-Seguel F, Castro-Ávila AC, Leppe J. Construct Validity of the Chilean-Spanish Version of the Functional Status Score for the Intensive Care Unit: A Prospective Observational Study Using Actigraphy in Mechanically Ventilated Patients. Archives of Physical Medicine and Rehabilitation. 2020;101:1914–21.

5. Schujmann DS, Teixeira Gomes T, Lunardi AC, Zoccoler Lamano M, Fragoso A, Pimentel M, et al. Impact of a Progressive Mobility Program on the Functional Status, Respiratory and Muscular Systems of ICU Patients. Critical Care Medicine. 2019;:1.

6. Baldwin CE, Rowlands A V., Fraysse F, Johnston KN. The sedentary behaviour and physical activity patterns of survivors of a critical illness over their acute hospitalisation: An observational study. Australian Critical Care. 2020;33:272–80.

7. Beach LJ, Fetterplace K, Edbrooke L, Parry SM, Curtis R, Rechnitzer T, et al. Measurement of physical activity levels in the Intensive Care Unit and functional outcomes: An observational study. Journal of Critical Care. 2017;40:189–96.

8. González-Seguel F, Corner EJ, Merino-Osorio C. International Classification of Functioning, Disability, and Health Domains of 60 Physical Functioning Measurement Instruments Used During the Adult Intensive Care Unit Stay: A Scoping Review. Physical Therapy. 2019;99:627–40.

9. Parry SM, Huang M, Needham DM. Evaluating physical functioning in critical care: considerations for clinical practice and research. Critical Care. 2017;21:249.

10. Parry SM, Granger CL, Berney S, Jones J, Beach L, El-Ansary D, et al. Assessment of impairment and activity limitations in the critically ill: a systematic review of measurement instruments and their clinimetric properties. Intensive Care Medicine. 2015;41:744–62.

11. Mourtzakis M, Parry S, Connolly B, Puthucheary Z. Skeletal muscle ultrasound in critical care: A tool in need of translation. Annals of the American Thoracic Society. 2017;14:1495–503.

12. 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. Critical Care Medicine. 2015;43:897–905.

13. Formenti P, Umbrello M, Coppola S, Froio S, Chiumello D. Clinical review: peripheral muscular ultrasound in the ICU. Annals of Intensive Care. 2019;9.

14. Cartwright MS, Demar S, Griffin LP, Balakrishnan N, Harris JM, Walker FO. Validity and reliability of nerve and muscle ultrasound. Muscle & Nerve. 2013;47:515–21.

15. Sarwal A, Parry SM, Berry MJ, Hsu F-C, Lewis MT, Justus NW, et al. Interobserver Reliability of Quantitative Muscle Sonographic Analysis in the Critically Ill Population. Journal of Ultrasound in Medicine. 2015;34:1191–200.

16. Segers J, Hermans G, Charususin N, Fivez T, Vanhorebeek I, Van den Berghe G, et al. Assessment of quadriceps muscle mass with ultrasound in critically ill patients: intra- and inter-observer agreement and sensitivity. Intensive Care Medicine. 2015;41:562–3.

17. Mayer KP, Dhar S, Cassity 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. Physical Therapy. 2020;100:1701–11.

18. 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 Anesthesiology. 2018;18:205.

19. Parry SM, Burtin C, Denehy L, Puthucheary ZA, Bear D. Ultrasound Evaluation of Quadriceps Muscle Dysfunction in Respiratory Disease. Cardiopulmonary Physical Therapy Journal. 2019;30:15–23.

20. Whittaker JL, Ellis R, Hodges PW, OSullivan C, Hides J, Fernandez-Carnero S, et al. Imaging with ultrasound in physical therapy: What is the PT's scope of practice? A competency-based educational model and training recommendations. British Journal of Sports Medicine. 2019;:bjsports-2018-100193.

21. McKiernan S, Chiarelli P, Warren-Forward H. A survey of diagnostic ultrasound within the physiotherapy profession for the design of future training tools. Radiography. 2011;17:121–5.

22. Wong A, Galarza L, Duska F. Critical Care Ultrasound. Critical Care Medicine. 2019;47:e256–62.

23. 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. Australasian Journal of Ultrasound in Medicine. 2018;21:104–14.

24. Ntoumenopoulos G, Ong HK, Toh HC, Saclolo RP, Sewa WD. Evaluation of a pilot programme on diagnostic thoracic ultrasound curriculum for acute care physiotherapists. Australasian Journal of Ultrasound in Medicine. 2017;20:147–54.

25. McKiernan S, Chiarelli P, Warren-Forward H. The content of a training package in diagnostic ultrasound for physiotherapists. Sonography. 2015;2:1–7.

26. Baka A, Figgou L, Triga V. "Neither agree, nor disagree": a critical analysis of the middle answer category in Voting Advice Applications. International Journal of Electronic Governance. 2012;5 3/4:244.

27. 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 Medical Research Methodology. 2020;20:293.

28. Arts IMP, Pillen S, Schelhaas HJ, Overeem S, Zwarts MJ. Normal values for quantitative muscle ultrasonography in adults. Muscle & Nerve. 2010;41:32–41.

29. Witteveen E, Sommers J, Wieske L, Doorduin J, van Alfen N, Schultz MJ, et al. Diagnostic accuracy of quantitative neuromuscular ultrasound for the diagnosis of intensive care unit-acquired weakness: a cross-sectional observational study. Annals of Intensive Care. 2017;7.

30. van Alfen N, Mah JK. Neuromuscular Ultrasound: A New Tool in Your Toolbox. Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques. 2018;45:504–15.

31. Bönnemann CG, Brockmann K, Hanefeld F. Muscle ultrasound in Bethlem myopathy. Neuropediatrics. 2003;34:335–6.

32. Pillen S, Arts IMP, Zwarts MJ. Muscle ultrasound in neuromuscular disorders. Muscle & Nerve. 2008;37:679–93.

33. Ikegawa S, Funato K, Tsunoda N, Kanehisa H, Fukunaga T, Kawakami Y. Muscle Force per Cross-sectional Area is Inversely Related with Pennation Angle in Strength Trained Athletes. Journal of Strength and Conditioning Research. 2008;22:128–31.

34. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine. 2016;15:155–63.

35. 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. PLOS ONE. 2019;14:e0219057.

36. Lenz ST. Michael J. Campbell, David Machin and Stephen J. Walters (2007): Medical Statistics, a Textbook for the Health Sciences, 4th edition. Statistical Papers. 2009;50:217–8.

37. Shechtman O. The Coefficient of Variation as an Index of Measurement Reliability. In: Doi S. WG (eds), editor. Methods of Clinical Epidemiology. Berlin, Heidelberg: Springer Series on Epidemiology and Public Health. Springer; 2013. p. 39–49.

38. Greenstein YY, Littauer R, Narasimhan M, Mayo PH, Koenig SJ. Effectiveness of a Critical Care Ultrasonography Course. Chest. 2017;151:34–40.

39. Vieira L, Rocha LPB, Mathur S, Santana L, Melo PF de, Silva VZM da, et al. Reliability of skeletal muscle ultrasound in critically ill trauma patients. Revista Brasileira de Terapia Intensiva. 2019;31.

40. Tourel C, Burnol L, Lanoiselé J, Molliex S, Viallon M, Croisille P, et al. Reliability of standardized ultrasound measurements of quadriceps muscle thickness in neurological critically ill patients: a comparison to computed tomography measures. Journal of Rehabilitation Medicine. 2020;52.

41. Hadda V, Khilnani GC, Kumar R, Dhunguna A, Mittal S, Khan MA, et al. Intra- and Inter-observer Reliability of Quadriceps Muscle Thickness Measured with Bedside Ultrasonography by Critical Care Physicians. Indian Journal of Critical Care Medicine. 2017;21:448–52.

42. Sabatino A, Regolisti G, Bozzoli L, Fani F, Antoniotti R, Maggiore U, et al. Reliability of bedside ultrasound for measurement of quadriceps muscle thickness in critically ill patients with acute kidney injury. Clinical Nutrition. 2017;36:1710–5.

43. Staehli S, Glatthorn JF, Casartelli N, Maffiuletti NA. Test–retest reliability of quadriceps muscle function outcomes in patients with knee osteoarthritis. Journal of Electromyography and Kinesiology. 2010;20:1058–65.

44. 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.

45. Oranchuk DJ, Nelson AR, Storey AG, Cronin JB. Variability of Regional Quadriceps Architecture in Trained Men Assessed by B-Mode and Extended-Field-of-View Ultrasonography. International Journal of Sports Physiology and Performance. 2020;15:430–6.
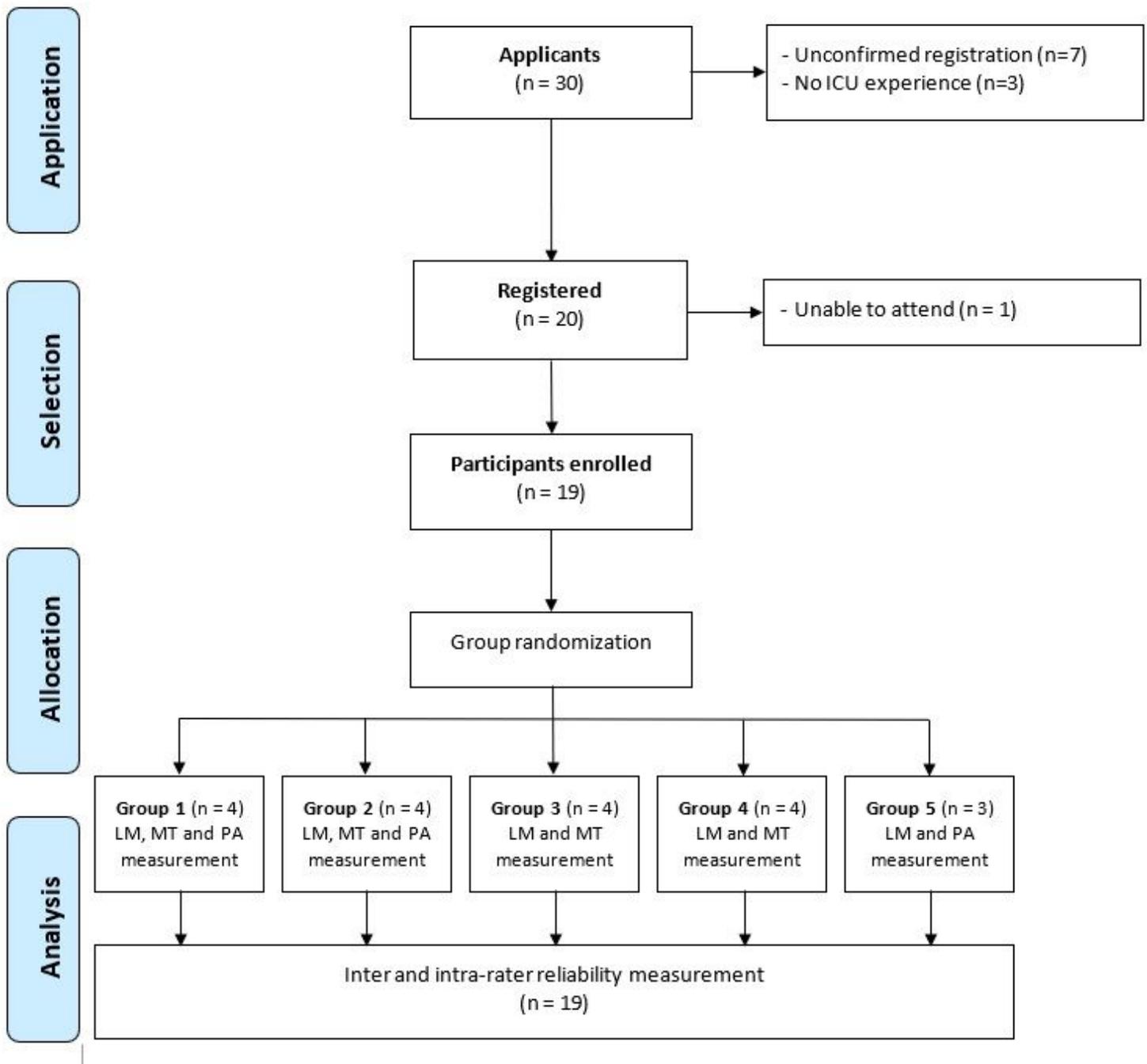
Figures

Figure 1

Flow diagram of participants during eMUSICS training programme. ICU = intensive care unit, LM = landmarks, MT = muscle thickness, PA = pennation angle

Figure 2

Reliability measurement process during hands-on assessment. (A) Landmarking of a left quadriceps. (B) Identification of anatomical structures in a transverse cross-sectional view using a curvilinear transducer. (C) Measurement of quadriceps muscle thickness using the ultrasound calliper. (D) Measurement of pennation angle of vastus lateralis. (E) Transcription of measurement values from ultrasound machine to the pre-specified written record document.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile.docx