

Clinical evaluation of AI-assisted muscle ultrasound for monitoring muscle wasting in ICU patients

Phung Tran Huy Nhat

nhat.phung@kcl.ac.uk

King's College London

Hao Nguyen Van

Hospital for Tropical Diseases

Minh Yen Lam

Oxford University Clinical Research Unit

Hoang Anh Nguyen

Hospital for Tropical Diseases

Phu Khiem Dong

National Hospital for Tropical Diseases

Hamideh Kerdegari

King's College London

Thanh Phuong Le

Oxford University Clinical Research Unit

Tan Hoang Vo

Oxford University Clinical Research Unit

Thanh Ngoc Nguyen

Oxford University Clinical Research Unit

Ngoc Minh Thu Le

Hospital for Tropical Diseases

Ngoc Trung Truong

Hospital for Tropical Diseases

Luigi Pisani

Mahidol Oxford Tropical Medicine Research Unit

Reza Razavi

King's College London

Sophie Yacoub

Oxford University Clinical Research Unit

Van Vinh Chau Nguyen

Hospital for Tropical Diseases

Andrew P. King

King's College London

Louise Thwaites

Oxford University Clinical Research Unit

Linda Denehy

University of Melbourne

Alberto Gomez

King's College London

Research Article

Keywords: Muscle ultrasound, muscle wasting, intensive care unit, artificial intelligence, real-time

Posted Date: October 23rd, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3456993/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on June 26th, 2024. See the published version at <https://doi.org/10.1038/s41598-024-64564-w>.

Abstract

Background

Muscle ultrasound has been shown to be a valid and safe imaging modality to assess muscle wasting in critically ill patients in the intensive care unit (ICU). This typically involves manual delineation to measure the Rectus Femoris cross-sectional area (RFCSA), which is a subjective, time-consuming, and laborious task that requires significant expertise. We aimed to develop and evaluate an AI tool to support non-expert operators in measurement of the RFCSA using muscle ultrasound.

Method

This is a prospective study conducted in the ICU at the Hospital of Tropical Diseases (HTD), Ho Chi Minh city, Vietnam. Patients diagnosed with severe tetanus underwent three muscle ultrasound examinations of their Rectus Femoris muscle (on day 1, day 7 and ICU discharge). Patients were randomized to undergo the examinations performed by a group of non-expert users, with or without an AI tool for assistance.

Results

Twenty patients were recruited at the Adult ICU at HTD between Feb 2023 and July 2023 and were randomized sequentially to operators using AI (n = 10) or non-AI (n = 10). The median (IQR) ICU stay was 23 days (IQR 20–30). Muscle loss during ICU stay was similar for both methods: $26 \pm 15\%$ for AI and $23 \pm 11\%$ for the non-AI, respectively ($p = 0.13$). In total 59 ultrasound examinations were carried out (30 without AI and 29 with AI). When assisted by our AI tool, the operators showed less variability between measurements with higher intraclass correlation coefficients (ICCs 0.999 95%CI 0.998–0.999 vs. 0.982 95%CI 0.962–0.993) and lower Bland Altman limits of agreement ($\pm 1.9\%$ vs. $\pm 6.6\%$) compared to not using the AI tool. The time spent on scans reduced significantly from a median of 19.6 mins (IQR 16.9–21.7) to 9.4 mins (IQR 7.2–11.7) compared to when using the AI tool ($p < 0.001$).

Conclusions

AI-assisted muscle ultrasound removes the need for manual tracing, increases reproducibility and saves time. This system may aid monitoring muscle size in ICU patients assisting rehabilitation programmes.

Background

There is a large body of research demonstrating that patients admitted to the intensive care unit (ICU) suffer significant morbidity including functional impairments and early and rapid loss of muscle mass [1]. Loss of muscle mass contributes to muscle dysfunction and may impact overall function, but other

variables are also implicated. The reasons for these changes are multifactorial and may include impaired muscle protein synthesis associated with sepsis [2]; patient comorbidity, organ dysfunction, duration of mechanical ventilation and length of ICU stay.

Measurement of muscle changes in ICU is challenging due to patient sedation and subsequent difficulties with following commands when using traditional volitional techniques such as the Medical Research Council sum score [3]. The use of a non-volitional measure such as point-of-care ultrasound (POCUS) offers the potential to examine muscle changes [4].

Patients admitted with tetanus, of whom 50% are intubated and ventilated, administered muscle relaxant drugs and benzodiazepines, spend approximately three weeks in the ICU [5]. Patient index admission diagnosis and sequelae from the ICU admission (such as sepsis) are associated with loss of muscle mass, weakness and impaired functional outcomes [6]. Assessment of function is difficult in these patients. The use of POCUS has gained traction and in patients with muscle failure such as tetanus allows serial monitoring.

Muscle ultrasound is commonly used to monitor muscle wasting by measuring the Rectus Femoris cross-sectional area (RFCSA) [6, 7]. In a recent systematic review, muscle ultrasound was used in 85% (28/33) of the studies to assess muscle mass in the ICU [15]. The effectiveness of ultrasound for assessing muscle mass is comparable to, but safer, cheaper, more accessible and more easily repeatable than any other imaging modality in the ICU [9–13]. In addition, recent evidence suggests that the assessment of changes in muscle size over time may improve prognostication and enhance the choice of rehabilitation interventions that may address muscle wasting [11, 14–16]. However, the process of measurement of RFCSA from ultrasound is a time-consuming task and often suffers from significant intra and interobserver variability [10–12, 14, 17], hindering its use to inform patient management. When applied to the exact same stand views, repeated RFCSA muscle recordings always introduce image differences due to variations in probe positioning, angulation, and tilt, as well as the manual delineation of muscle by the operator. Conventionally, to enhance the accuracy of the measurement of the muscle size, the RFCSA is measured three times consecutively and an average of the measurements is calculated and saved, which increases the time taken for scanning and data acquisition.

To address the challenge of intra- and interobserver variability in muscle ultrasound measurement, AI techniques have been proposed [18–20]. However, the main limitation with these tools is that they have typically been designed and evaluated for offline use, i.e, they are not suitable for bedside use and real-time analysis. Moreover, to date these tools have been subject to limited validation in clinical settings.

In this study, we aimed to investigate the feasibility of using a real-time AI-assisted tool for RFCSA measurement from muscle ultrasound that would be suitable for clinical use, particularly in an LMIC setting. We hypothesized that this tool would have improved reproducibility and reduced interobserver variability compared to current methods. We tested our tools in a cohort of patients with tetanus as this

is a group of patients in whom muscle wasting is an important problem with long ICU stays and in whom such a tool would be used if proved reliable.

Methods

This was a prospective observational study to test the reliability of AI-assisted measurement of RFCSA from muscle ultrasound at the patient's bedside compared to standard ultrasound. The study was conducted in the adult ICU at the Hospital for Tropical Diseases Ho Chi Minh City, Vietnam.

Measurements were performed in adult patients with severe tetanus (Ablett Grade 3 or 4) admitted to the Adult ICU at HTD expected to stay at least 5 days. Patients were receiving standard treatment including mechanical ventilation, muscle relaxation and neuromuscular blockers following the hospital guideline [5]. Written informed consent was obtained from participant's legally authorized representatives prior to enrollment. Ethics approval was obtained from the HTD Ethics Committee, and the Oxford Tropical Research Ethics Committee (OxTREC). The study was registered retrospectively at ClinicalTrials.gov number NCT06034093, Date registered August 30, 2023.

RFCSA measurements were by using a standard technique or the AI-assisted method based on patient allocations – where patients were randomly assigned to receive either standard RFCSA measurements or AI-enabled measurements at regular intervals during their ICU stay (see Fig. 2).

All measurements were carried out according to a standard operating procedure where patients were in the supine position with the leg in neutral rotation. Patients not receiving muscle relaxants were reminded to relax the muscle. Measurements were taken using a 12L-RS linear probe and a Venue Go ultrasound machine (General Electric Healthcare, London, UK). Measurements were taken from a location which was three-fifths of the distance from the anterior superior iliac spine to the superior patella pole. This position was used as a landmark for subsequent measurements to provide consistency and allow reliable comparisons to be made over time. The transducer was placed perpendicular to the skin and transversally in relation to the longitudinal axis of the thigh to observe the cross-sectional area of the muscle. An excess of ultrasound gel was placed when performing the muscle scan and the pressure on the skin was kept minimal to ensure good image quality [6, 21, 22]. For each examination 3 separate measurements (scan-rescan) were made for each leg (removing the probe between each one). The examinations' durations were recorded.

The operators selected were 3 clinicians and 2 nurses, all 5 with limited training in muscle ultrasound as our target users for the AI tool are non-expert operators to reflect the common setting in a LMIC. We provided muscle ultrasound training and RAIMUS software training to allow operators to use the tool effectively. Each operator was asked to scan five patients (2 legs, each leg 3 repeat scans) as part of their muscle training with the AI software. The images were saved and manually delineated. with the AI software. The images were saved and manually delineated.

For patients in the standard measurement arm RFCSA was determined by manual delineation of the cross-sectional image of the muscle. For patients in the AI-assisted imaging arm measurements were

made in real time using the automated AI tool.

The AI-enabled technique using RAIMUS software is described in detail in the Additional file 1 and illustrated in Fig. 2. The dataset and model architecture used to develop the AI model is shown in Table S1 and Figure S1, respectively. Our RAIMUS software enables automated detection and measurement of RFCSA which can be visualized in real time using the Plugin-based REal Time UltraSound software platform (PRETUS) which connects a laptop to the ultrasound machine duplicating the ultrasound image display and superimposing the AI delineation measurement tool [23]. Once the probe is in the right position and the clinician is satisfied with the quality of the image/AI-generated segmentation mask, they store the image as a standard view. Next, the RFCSA is delineated by the AI tool or by the operator using a movable cursor following the inner echogenic line of the rectus fascia.

Evaluation of reproducibility

We assessed three types of variability including (1) scan-rescan variability, (2) intraobserver variability in delineation and (3) interobserver variability in delineation (Fig. 1). To assess (1) scan-rescan variability, the operators were asked to scan each leg three times for each of the two allocated methods. To assess (2) intraobserver variability in delineation over time, the same operator subsequently delineated each image one further time 2–4 weeks after the images were acquired (from stored raw images). To assess (3) interobserver variability in delineation, each image acquired by the first operator was delineated by 2 additional operators.

The examination durations (measured from when the operators put the probe on the leg of patient to when they finished delineations and measurements) were compared between the 2 methods. To evaluate usability, a questionnaire was administered to the operators at the end of the procedure (Additional file 1).

Sample size

The sample size for this study was estimated following Walter et al [24] with the minimum acceptable reliability (Intraclass correlation coefficient ICC) (ρ_0) of 0.9, expected reliability (ICC) (ρ_1) of 0.96, significance level two-tailed (α) of 0.05, Power ($1 - \beta$) of 80% and the number of raters/measurements per subjects (k) of 3. After using the formula, the sample size was 27 examinations. With the expected dropout of 10% the total sample size used was 30 examinations for each group.

Statistical analysis

All statistical analyses were performed with R version 4.0.4. Continuous variables are expressed as mean \pm standard deviation (SD) or as median (interquartile range), according to the symmetry of the data distribution, and compared using an unpaired Student's t-test or Wilcoxon rank sum test, as appropriate. Categorical data, presented as numbers and percentages, were compared using the χ^2 test. P values less than 0.05 were considered statistically significant.

The variability in RFCSA measurements were assessed using the two-way random effects for multiple raters/measurements ICC with 95% CI (ICC (2, k) [25]. The ICC is a measure of reliability, specifically the reliability of two different raters to measure subjects similarly, with numbers closest to 1 representing a high similarity of measurements between measurements. The standard error of measurement (SEM) was also calculated to make a judgment about the degree that measurements vary for an individual. The SEM values indicate the precision of the measurement and were calculated based on the ICC and the SD of the mean of differences between the two measurements $SEM = SD\sqrt{1 - ICC}$. There was no measure-remeasure variation for the automated AI software because the model always outputs the same measurement and hence the same RFCSA result. A modified Bland-Altman analysis for multiple observers in a single plot proposed by [26] was used to assess the agreement between RFCSA measurements. The examination duration was compared between the two methodologies of measurement using an unpaired Student's t-test.

Results

Patients

We enrolled 20 patients with tetanus at the Adult ICU at HTD between Feb 2023 and July 2023. The mean \pm SD age of patients in the AI group and non-AI group were 67 ± 13 years and 56 ± 17 years, respectively. Two (20%) patients in each group were female. 17 (85%) patients had at least one episode of hospital acquired infection (HAI) during ICU stay. The ICU stay and hospital stay were comparable between the two groups (Table 1).

Ultrasound examinations and reproducibility

In total 59 muscle ultrasound examinations were carried out, 29 examinations with the AI tool and 30 examinations without the AI tool. After visual inspection of 29 examinations with AI, 28 examinations were successfully delineated by the AI tool and one examination was rejected by the expert. All examinations without the AI tool were accepted by the expert. The average muscle loss during ICU was similar in the two groups, $26 \pm 15\%$ for the AI arm and $23 \pm 11\%$ for the non-AI arm. The full characteristics of the patients are shown in Table 1.

Table 1
 Characteristics of patients in the study (n = 20)

	AI arm (n = 10)	Non-AI arm (n = 10)
Age	67 ± 13	56 ± 17
Sex (female)	2 (20%)	2 (20%)
Comorbidities (1 or more)	6 (60%)	4 (40%)
Sedative use during ICU	10 (100%)	10 (100%)
Use of non-depolarising neuromuscular blocking agents during ICU stay	10 (100%)	10 (100%)
Length of ICU stay (days)	26 ± 11	24 ± 5
Length of hospital stay (days)	32 ± 13	31 ± 6
Mechanical ventilation duration (days)	19.7 ± 8.2	17.8 ± 5.5
ANSD duration (days)	12.4 ± 6.7	12.0 ± 2.8
Total dose of Pipecuronium	438 ± 190	430 ± 250
Enteral nutrition	10 (100%)	10 (100%)
Rehabilitation duration (days)	10 ± 9	8 ± 4
HAI during ICU stay	8 (80%)	9 (90%)
RF CSA D1 (cm ²)	4.37 ± 1.08	4.76 ± 1.50
RF CSA D7 (cm ²)	4.09 ± 1.01	4.51 ± 1.59
RFCSA Discharge (cm ²)	3.25 ± 1.24	3.65 ± 1.15
% change in RFCSA during ICU stay (%)	26 ± 15	23 ± 11
ICU survival	10 (100%)	10 (100%)

(MV: Mechanical ventilation, ANSD: Autonomous Nervous System Dysfunction, HAIs: Hospital-Acquired Infections, RFCSA D1: Rectus Femoris Cross sectional Area on ICU admission, RFCSA D7: Rectus Femoris Cross sectional Area on Day 7, RFCSA Discharge: Rectus Femoris Cross sectional Area at ICU Discharge)

Scan-rescan variability

The scan re-scan variability of the AI group was lower compared to the non-AI group (ICC 0.999 95%CI 0.998–0.999, vs ICC 0.982 95%CI 0.962–0.993). (Additional file Table S2)

Figure 3. Plot of scan-rescan agreement in RFCSA. A) Without AI. B) With AI. Horizontal dotted lines indicate the limits of agreement from the mean (LoA) of the three measurements. Some symbols are

superimposed. The percentage differences of all measurements with the mean (y-axis) are plotted against the mean RFCSA for all participants (x-axis). The horizontal dashed lines indicate the limits of agreement with the mean of the three repeated measurements and ranged from - 6.6 to 6.6% for the non-AI group and - 1.9 to 1.9% for the AI group.

Intraobserver variability in delineation over time

The manual intraobserver in delineation (initial vs 2–4 weeks later) resulted in good reliability with an ICC of 0.984 (95% CI 0.973, 0.990). The modified Bland Altman plot for intraobserver agreement results is shown in Fig. 4 (left). The intraobserver agreement without AI was $\pm 5.9\%$.

Interobserver variability in delineation

The manual interobserver ICC was 0.974 (95% CI 0.965–0.981). The interobserver agreement results are shown in Fig. 4 (right). The limits of agreement without AI were $\pm 8.2\%$ and there was no interobserver variation for the AI group for the reason stated in the statistical analysis section.

Examination duration

Examination duration (including acquisition and measurement) was shorter in the AI group compared to the non-AI group: a median of 9.4 minutes (IQR 7.2–11.7) compared to 19.6 minutes (IQR 16.9–21.7) ($p < 0.001$).

Discussion

The study presents, for the first time, a prospective study to evaluate the impact of an AI tool for real-time RFCSA estimation compared to the traditional manual measurement technique for monitoring muscle mass in the ICU. The AI tool succeeded in supporting operators in assessing muscle wasting in patients with a range of RFCSAs and varying image qualities by correctly delineating the RF muscle and measuring the RFCSA with less variability than standard non-AI measurements. Furthermore, the time spent on measuring RFCSA using the AI tool was approximately half that of standard measurements.

The study showed the reduction in scan-rescan variability, this may involve both acquisition variability (taking the probe back to the same plane every time) and intraobserver delineation variability. In addition, the scan-rescan variability, intraobserver and interobserver delineation variability are similar which indicates that the main variability of the standard technique is manual delineation. This suggest that it may be possible if using the AI tool to perform a 1-scan measurement instead of standard average of 3-scan measurement, further reducing time.

Thus, with the help of the automated AI tool, monitoring of muscle changes in ICU patients could be more practical and feasible than before. The reliability of operators with limited training in our study was already good without the AI tool but the time spent on manual measurements was twice that when using

the AI tool. However, further research should be focused on the use of AI to guide accurate and reproducible probe placement.

Limitations

It is important to emphasize that removing the measure-remeasure variability and interobserver variability completely by the use of a deterministic AI model makes repeated measurements more reproducible but does not necessarily make them more accurate. Although the performance of our AI tool is good, future work should more thoroughly assess its robustness and bias, especially in the presence of poor image quality. Further research should also focus on validation of the AI tool in different patient populations, using different machine manufacturers and different settings, for example, monitoring muscle wasting in patients with cancer.

Echogenicity was not investigated in this study. When muscle echogenicity increases, determining the muscle boundaries is very challenging because muscle tissue is replaced by intramuscular fibrous and fat tissue. As a result, the contrast between the muscle boundaries and other structures decreases. Future work should develop AI based methods for assessing muscle echogenicity as also muscle echogenicity can provide useful information on both quantity and quality of the muscle.

Conclusions

Real-time AI-assisted muscle ultrasound removes the need for manual tracing, increases reproducibility, and saves time. Our study has shown that much of the variability between measurements is related to manual delineation of the muscle and hence potentially an even faster single-scan protocol could be adopted for AI-assisted RFCSA measurement. Such a system would significantly assist routine clinical monitoring of muscle changes in ICU patients and help in assessing the effectiveness of interventions.

Declarations

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The codes and model weights needed to deploy the RAIMUS tool for the study are available at <https://github.com/vital-ultrasound/public-muscle>.

Ethics approval and consent to participate

This study was approved by the Oxford Tropical Research Ethics Committee (OxTREC) and the HTD Institutional Review Boards. Written informed consent was obtained from participant's legally authorized representatives prior to enrollment.

Funding

This work was supported by the Wellcome Trust under grant 217650/Z/19/Z.

Author contributions

P.T.H.N., L.M.Y., A.G., L.T., A.K., L.P., L.D., and N.V.H. conceptualised the study. Implementation was by P.T.H.N., L.T.P., L.M.Y., L.N.M.T., N.T.N., N.T.H., with support from T.N.T., H.K., L.P., S.Y., N.H.A., N.V.V.C., and R.R., A.M. Statistical analysis was performed by P.T.H.N. and A.K. Initial manuscript drafting was by P.T.H.N., A.G., L.T. with support from D.P.K., L.D., L.P. and R.R.

Conflict of interest

All other authors declare no conflict of interest.

Acknowledgments

We thank all the patients, nurses and clinicians who generously donated their time during the study.

References

1. Needham DM, Davidson J, Cohen H, Hopkins RO, Weinert C, Wunsch H, et al. Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. *Crit Care Med.* 2012;40:502–9.
2. Puthuchery ZA, Rawal J, McPhail M, Connolly B, Ratnayake G, Chan P, et al. Acute skeletal muscle wasting in critical illness. *JAMA.* 2013;310:1591–600.
3. Turan Z, Topaloglu M, Ozyemisci Taskiran O. Medical Research Council-sumscore: a tool for evaluating muscle weakness in patients with post-intensive care syndrome. *Crit Care.* 2020;24:1–2.
4. 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.
5. Van Hao N, Yen LM, Davies-Foote R, Trung TN, Duoc NVT, Trang VTN, et al. The management of tetanus in adults in an intensive care unit in Southern Vietnam. *Wellcome Open Res.* 2021;6.
6. Trung TN, Duoc NVT, Nhat LTH, Yen LM, Hao N V, Truong NT, et al. Functional outcome and muscle wasting in adults with tetanus. *Trans R Soc Trop Med Hyg.* 2019;113.
7. Puthuchery ZA, Phadke R, Rawal J, McPhail MJW, Sidhu PS, Rowlerson A, et al. Qualitative ultrasound in acute critical illness muscle wasting. *Crit Care Med.* 2015;43:1603–11.
8. Fazzini B, Märkl T, Costas C, Blobner M, Schaller SJ, Prowle J, et al. The rate and assessment of muscle wasting during critical illness: a systematic review and meta-analysis. *Crit Care.* 2023;27:1–26.
9. Sponbeck JK, Frandsen CR, Ridge ST, Swanson DA, Swanson DC, Johnson AW. Leg muscle cross-sectional area measured by ultrasound is highly correlated with MRI. *J Foot Ankle Res.* 2021;14:1–7.

10. 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;18:1–8.
11. 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;195:136–8.
12. 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;18:1–8.
13. Parry SM, Burtin C, Denehy L, Puthuchery ZA, Bear D. Ultrasound evaluation of quadriceps muscle dysfunction in respiratory disease. *Cardiopulm Phys Ther J.* 2019;30:15–23.
14. 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;195:136–8.
15. Zhang W, Wu J, Gu Q, Gu Y, Zhao Y, Ge X, et al. Changes in muscle ultrasound for the diagnosis of intensive care unit acquired weakness in critically ill patients. *Sci Rep.* 2021;11:18280.
16. 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.
17. Lima KMM, da Matta TT, de Oliveira LF. Reliability of the rectus femoris muscle cross-sectional area measurements by ultrasonography. *Clin Physiol Funct Imaging.* 2012;32:221–6.
18. Chen X, Xie C, Chen Z, Li Q. Automatic Tracking of Muscle Cross-Sectional Area Using Convolutional Neural Networks with Ultrasound. *Journal of Ultrasound in Medicine.* 2019;38.
19. Ritsche P, Wirth P, Cronin NJ, Sarto F, Narici M V, Faude O, et al. DeepACSA: Automatic Segmentation of Cross-sectional Area in Ultrasound Images of Lower Limb Muscles Using Deep Learning. *Med Sci Sports Exerc.* 2022.
20. Katakis S, Barotsis N, Kakotaritis A, Tsiganos P, Economou G, Panagiotopoulos E, et al. Muscle Cross-Sectional Area Segmentation in Transverse Ultrasound Images Using Vision Transformers. *Diagnostics.* 2023;13:217.
21. 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;14:1495–503.
22. Fazzini B, Märkl T, Costas C, Blobner M, Schaller SJ, Prowle J, et al. The rate and assessment of muscle wasting during critical illness: a systematic review and meta-analysis. *Crit Care.* 2023;27:1–26.
23. Gomez A, Zimmer VA, Wheeler G, Toussaint N, Deng S, Wright R, et al. PRETUS: A plug-in based platform for real-time ultrasound imaging research. *SoftwareX.* 2022;17.
24. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med.* 1998;17:101–10.

25. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15:155–63.
26. Christensen HS, Borgbjerg J, Børty L, Bøgsted M. On Jones et al.'s method for extending Bland-Altman plots to limits of agreement with the mean for multiple observers. *BMC Med Res Methodol.* 2020;20:1–8.

Figures

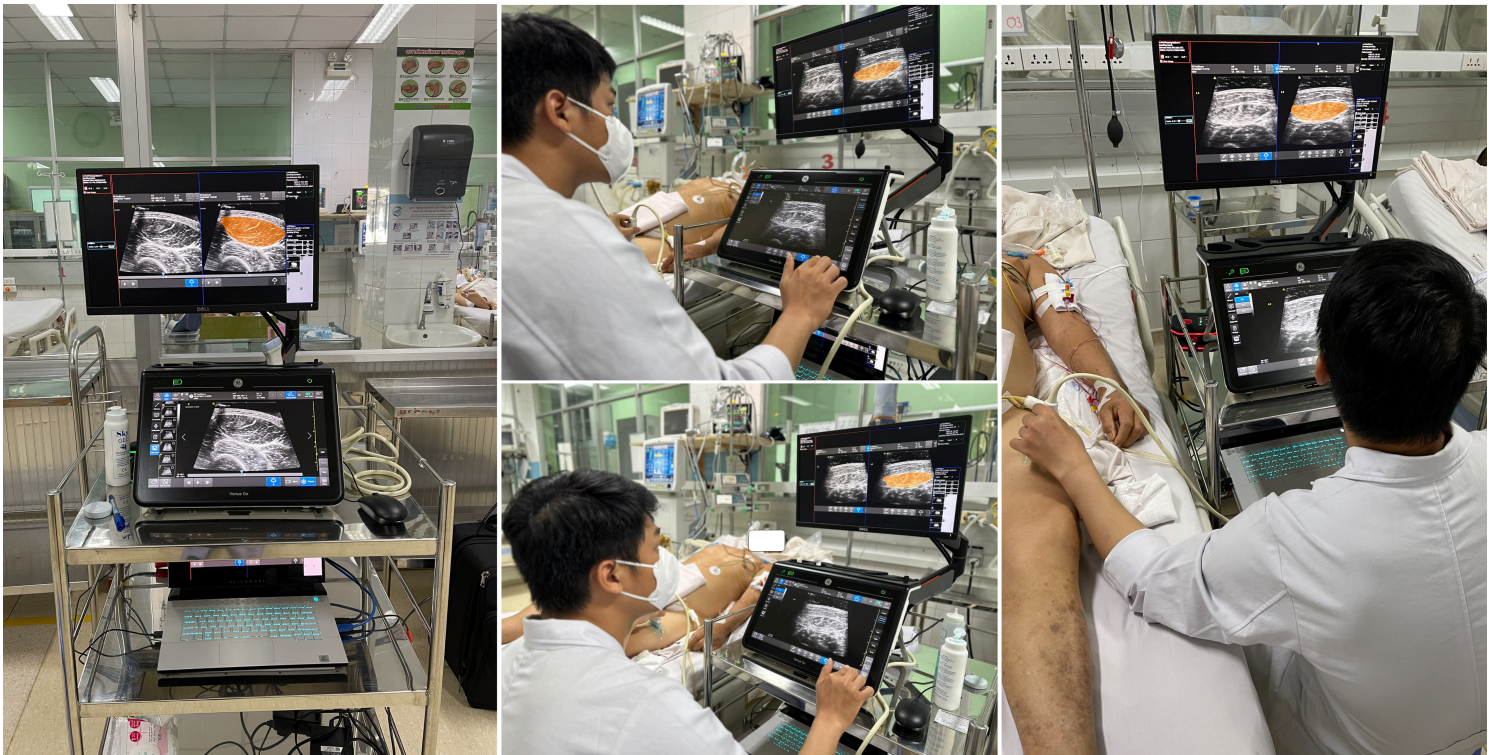


Figure 1

Real-time AI-assisted muscle ultrasound (RAIMUS) system

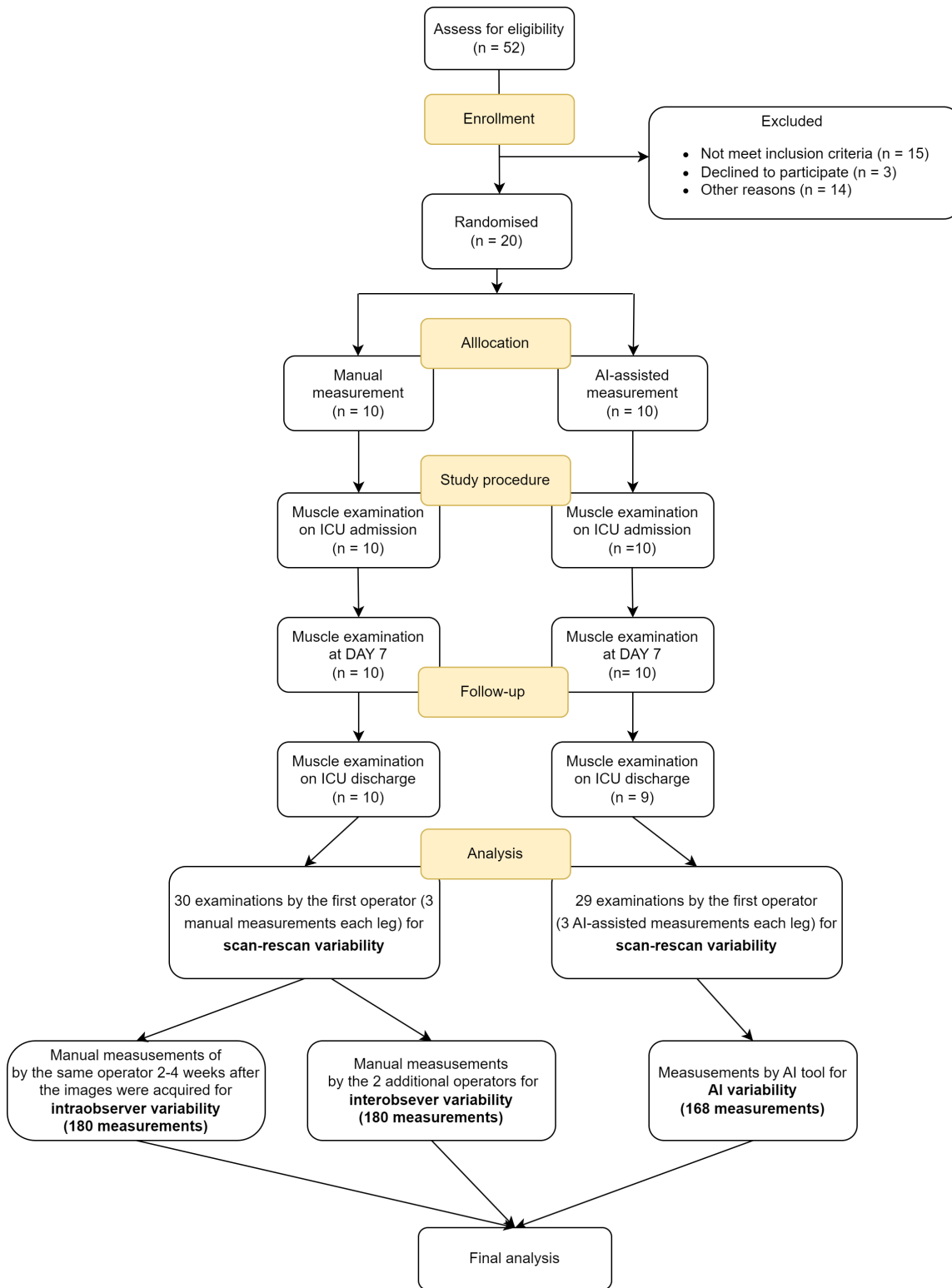


Figure 2

Study flowchart

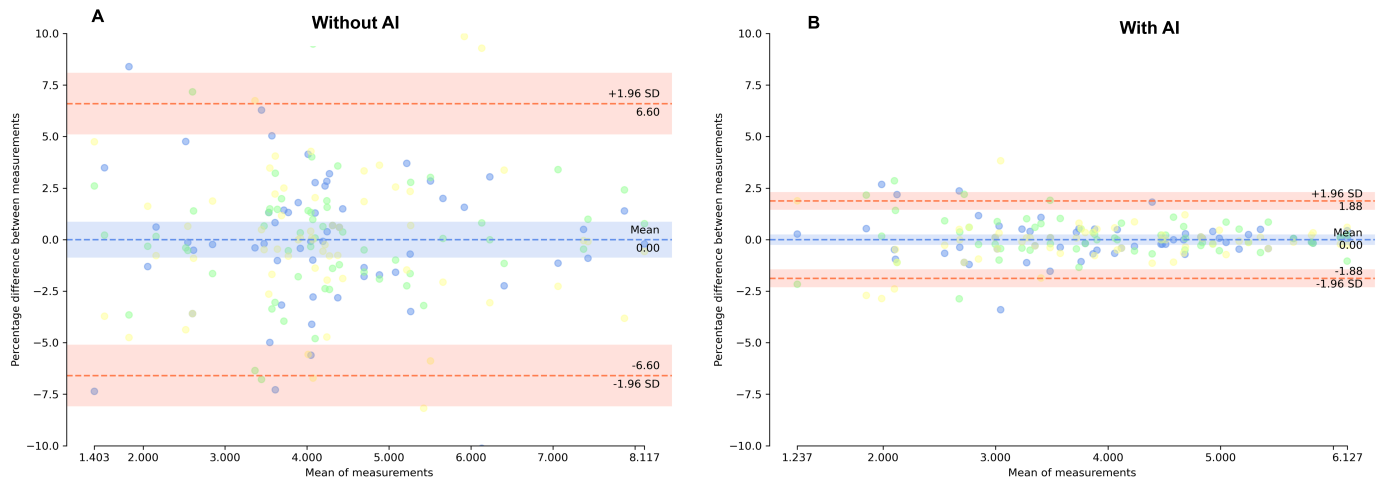


Figure 3

Plot of scan-rescan agreement in RFCSA. A) Without AI. B) With AI. Horizontal dotted lines indicate the limits of agreement from the mean (LoA) of the three measurements. Some symbols are superimposed. The percentage differences of all measurements with the mean (y-axis) are plotted against the mean RFCSA for all participants (x-axis). The horizontal dashed lines indicate the limits of agreement with the mean of the three repeated measurements and ranged from -6.6 to 6.6 % for the non-AI group and -1.9 to 1.9% for the AI group.

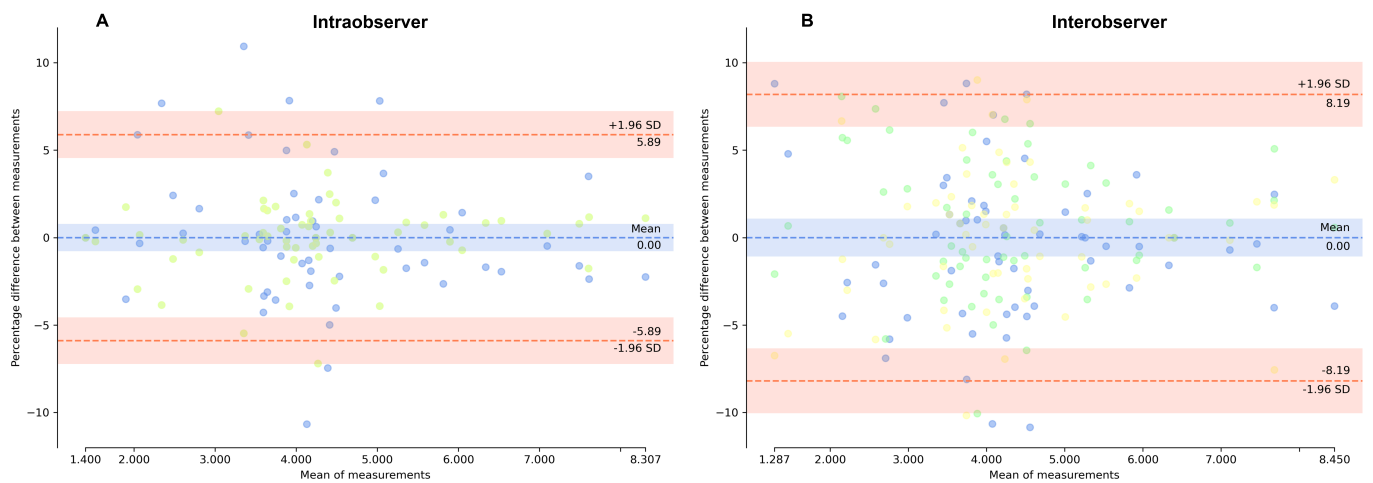


Figure 4

Intraobserver and Interobservers agreement plots of RFCSA measurements: (A) of the same operator over time and (B) interobserver agreement plot between 3 observers. Observers represent different symbols. The percentage differences of all measurements with the mean (y-axis) are plotted against the mean RFCSA for all participants (x-axis). The horizontal dashed lines indicate the limits of agreement

with the mean of the three observers and ranged from -5.9 to 5.9 % for interobserver (left) and from -8.2 to 8.2 % for interobserver (right).

Supplementary Files

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

- [RAIMUSsupplementarymaterial17OCT2023.docx](#)