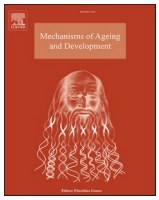
Mechanisms of Ageing and Development 215 (2023) 111860

Contents lists available at ScienceDirect

Mechanisms of Ageing and Development

journal homepage: www.elsevier.com/locate/mechagedev



Assessment and risk prediction of frailty using texture-based muscle ultrasound image analysis and machine learning techniques

Rebeca Mirón-Mombiela [a,b,c,1], Silvia Ruiz-España [d,1], David Moratal [d,2,*], Consuelo Borrás [a,e,f,2,**]

[a] Department of Physiology, Universitat de València/INCLIVA, Avda. Blasco Ibáñez, 15, 46010 Valencia, Spain
[b] Hospital General Universitario de Valencia (HGUV), Valencia, Spain
[c] Herlev og Gentofte Hospital, Herlev, Denmark
[d] Center for Biomaterials and Tissue Engineering, Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain
[e] INCLIVA Health Research Institute, Av/ de Menéndez y Pelayo, 4, 46010 Valencia, Spain
[f] Center for Biomedical Network Research on Frailty and Healthy Aging (CIBERFES), CIBER-ISCIII, Valencia, Spain

ARTICLE INFO

Keywords:
Frailty
Muscle
Ultrasound
Machine-learning
Texture analysis
Image biomarkers

ABSTRACT

The purpose of this study was to evaluate texture-based muscle ultrasound image analysis for the assessment and risk prediction of frailty phenotype. This retrospective study of prospectively acquired data included 101 participants who underwent ultrasound scanning of the anterior thigh. Participants were subdivided according to frailty phenotype and were followed up for two years. Primary and secondary outcome measures were death and comorbidity, respectively. Forty-three texture features were computed from the *rectus femoris* and the *vastus intermedius* muscles using statistical methods. Model performance was evaluated by computing the area under the receiver operating characteristic curve (AUC) while outcome prediction was evaluated using regression analysis. Models developed achieved a moderate to good AUC ($0.67 \leq AUC \leq 0.79$) for categorizing frailty. The stepwise multiple logistic regression analysis demonstrated that they correctly classified 70–87% of the cases. The models were associated with increased comorbidity ($0.01 \leq p \leq 0.18$) and were predictive of death for pre-frail and frail participants ($0.001 \leq p \leq 0.016$). In conclusion, texture analysis can be useful to identify frailty and assess risk prediction (i.e. mortality) using texture features extracted from muscle ultrasound images in combination with a machine learning approach.

1. Introduction

Over the last few decades, geriatrics and gerontology researchers have devoted an increasing amount of effort to developing and implementing preventive interventions against frailty. The accomplishment of such a task has been hampered by the lack of standardized, and universally agreed definitions for frailty (Rodríguez-Mañas et al., 2013). These definitional ambiguities are also reflected by the absence of reliable biomarkers that can identify frailty, track its progression, and monitor their response to interventions (Calvani et al., 2015).

There is a lack of methodology that could be used across many populations to optimize frail patient care, and available diagnostic tools are not fully exploited. Modern imaging techniques have a high potential to help fill this gap and facilitate frailty assessment. Ultrasound echo intensity is an attractive candidate to explore further as a biomarker of frailty (Miron Mombiela et al., 2017), as it can be used to evaluate objectively muscle quality and is relatively cheap (Fukumoto et al., 2012; Watanabe et al., 2013; Akima et al., 2017; Miron Mombiela et al.,

Abbreviations: AUC, area under the receiver operating characteristic curve; BMI, body mass index; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; GLSZM, gray-level size-zone matrix; KNN, k-nearest neighbors; MLP, multilayer perceptron; NB, naïve Bayes; QIBs, quantitative imaging biomarkers; RF, random forest; ROI, region of interest; SVM, support vector machine; SVM-RFE, support vector machine-recursive feature elimination.

* Corresponding author.
** Correspondence to: Department of Physiology, Faculty of Medicine, University of Valencia, Avenida Blasco Ibañez 15, 46010 Valencia, Spain.
E-mail addresses: rebeca.miron.mombiela@regionh.dk (R. Mirón-Mombiela), silruies@upv.es (S. Ruiz-España), dmoratal@eln.upv.es (D. Moratal), consuelo.borras@uv.es (C. Borrás).

[1] These authors contributed equally to this work.
[2] These authors contributed equally to this work.

https://doi.org/10.1016/j.mad.2023.111860
Received 25 April 2023; Received in revised form 8 August 2023; Accepted 30 August 2023
Available online 4 September 2023

2020). Mirón Mombiela and colleagues demonstrated that echo intensity was negatively correlated with muscle strength and was significantly greater in frail individuals, even after adjustment for sex and BMI (Mirón Mombiela et al., 2017). Two other studies have aimed to determine whether morphological and qualitative characteristics classified by quadriceps are associated with muscle strength, physical function, and sarcopenia in community-dwelling older adults (Kawai et al., 2018; Yamada et al., 2017), but only one of them tested ROC curves for echo intensity and founded an AUC of 0.66 for both men and women (Yamada et al., 2017). However, further investigation evaluating the accuracy, reproducibility, and predictive value is required to determine if echo intensity could become a useful aid in the diagnosis of frailty.

The ability of muscle ultrasound echo intensity, to prospectively predict clinical outcomes for adult and elderly patients has been even less investigated (Ticinesi et al., 2017). In addition, the ultrasonographic measurement of the rectus femoris cross-sectional area was demonstrated as an independent predictor of hospital length of stay, mortality, and nursing home discharge in a group of adults admitted to an intensive care unit (Mueller et al., 2016). A similar study found that loss of muscle mass shows a negative correlation with length of stay, and seems to be higher during the first 2–3 weeks of immobilization in the intensive care unit stay (Gruther et al., 2008). Despite these reports, the relevance of muscle ultrasound measures in terms of clinical outcomes needs further investigation.

The usual tools to assess frailty show, among other characteristics, a low sensitivity and a low positive predictive value. Many biomarkers of frailty have been identified but few of them have been assessed as clinical markers and there are controversial results (Rodríguez-Mañas, 2015). Bearing these considerations in mind, a shift of paradigm is needed, moving from the quest for a single biomarker to the development of multivariate/multidimensional modelling of a panel of complementary biomarkers. Single or isolated inspection of variables can result in a partial or incorrect picture. A recent study combining echo intensity and fractal analysis for the diagnosis of frailty concluded that fractal analysis was useful to characterize echo intensity, but did not improve its diagnostic performance (Mirón Mombiela et al., 2021). Authors suggest that combinations of different texture features, also known as texture analysis, may be required for accurate diagnosis and prognosis of frailty.

In the past few years, texture analysis has been proven to be an excellent source of imaging biomarkers in different areas of medicine (Rodriguez-Mañas et al., 2015, Chitalia et al., 2019, Ortiz-Ramón et al., 2018) including aging and neuromuscular disorders (Fritz et al., 2018; Nodera et al., 2018; Sogawa et al., 2017; Nodera et al., 2019). Texture analysis refers to the application of different mathematical methods to quantify the gray-level patterns and pixel inter-relationships within an image (Ríos-Díaz et al., 2019). These methods allow the computation of a considerable number of texture features, which can be combined with predictive models to increase precision in diagnosis or obtain reliable diagnostic tools (Larroza et al., 2016; Materka and Strzelecki, 1998; Wang and Summers, 2012). In addition, the use of texture analysis in medical images to characterize tissues allows us to quantify the intrinsic heterogeneous properties that are normally imperceptible to the human eye. We hypothesize that the texture analysis of ultrasound images in combination with a machine learning approach can be used for the identification of reliable frailty biomarkers.

The purpose of this study was to investigate the potential of 2D features extracted from muscle ultrasound images as a source of imaging biomarkers for the assessment and risk prediction of frailty phenotype. We also explored the output and impact of including clinical indicators in the process.

2. Materials and methods

2.1. Study design and setting

This was a diagnostic performance study of machine learning applied to texture analysis-derived features for frailty characterisation at muscle ultrasound, using a simple and commonly available clinical determination of Frailty as the reference standard[26]. The study was a secondary analysis of ultrasound images obtained between November 2014 and February 2015 from a prospective trial. The study was conducted on humans and adhered to the principles of bioethics included in the Declaration of Helsinki and the relevant Spanish legislation. The approval of the research committee and the clinical research ethics committee (CEIC by its Spanish acronym) of the Consorcio Hospital General Universitario de Valencia were obtained on the 30th of September 2014 and on the 27th of October 2014, respectively. All participants were informed of the experimental procedures and the purpose of the study. Each patient gave written informed consent before entering the study. Patients were referred from primary care to the Radiology Department, where the ultrasound was performed. Information about epidemiological data, frailty criteria, and quality of life was also obtained. Comorbidities were assessed at baseline and were followed up for the development of other comorbidities or death for up to two years (until March 31st, 2017). In the primary analysis, the objective was to determine whether echo intensity might be useful in a primary care setting to assess frailty, concluding that higher levels of echo intensity were associated with lower levels of strength, gait speed, and greater frailty (Miron Mombiela et al., 2017).

The analysis performed in this work comprised a process consisting of five major steps: image acquisition, region of interest (ROI) definition and preprocessing, feature extraction, feature selection, and classification (Fig. 1).

2.2. Study subjects

The inclusion criteria were: individuals aged $\geq$ 60 years old, able to walk independently, including with the help of a cane, walker, or similar assistance device. Frailty phenotype was determined following the Fried criteria (Fried et al., 2001), and it was used to divide the participants into robust, pre-frail, or frail based on their responses. In addition, participants between 20 and 59 years old were recruited to the control group, given the peak of muscular development within this age group[27], but also to have a true comparison of healthy individuals for the robust group and to add rigor to the study. The exclusion criteria were: participants suffering from neuromuscular disorders, oncological patients undergoing chemotherapy or radiotherapy, or dementia affecting the patient's capacity to understand the informed consent.

We invited 142 subjects, and 121 accepted to participate. Of those, only 112 individuals fulfilled all inclusion criteria and no exclusion criteria. Eleven healthy subjects were excluded from the statistical analysis either because frailty evaluation or ultrasound images classified them as frail, leaving 101 participants for the study. The flowchart of subject selection is shown in Fig. 2, and examples of control subjects excluded are shown in Supplemental Fig. S1.

2.3. Image acquisition

The anterior compartment of the right thigh was scanned by ultrasonography (LOGIC S7 Expert, General Electric, USA) in B mode with a 6–15 MHz linear transducer. Images were acquired on the halfway point between the lateral condyle of the femur and the anterosuperior iliac spine, where the *rectus femoris* and the *vastus intermedius* muscles are visible (Fig. 1). The rectus femoris and vastus lateralis muscle were selected for the study for the following reasons: a) they are large muscles, in comparison with muscles of the upper extremities and can be easily measured even when they become hypotrophic; b) they are easily

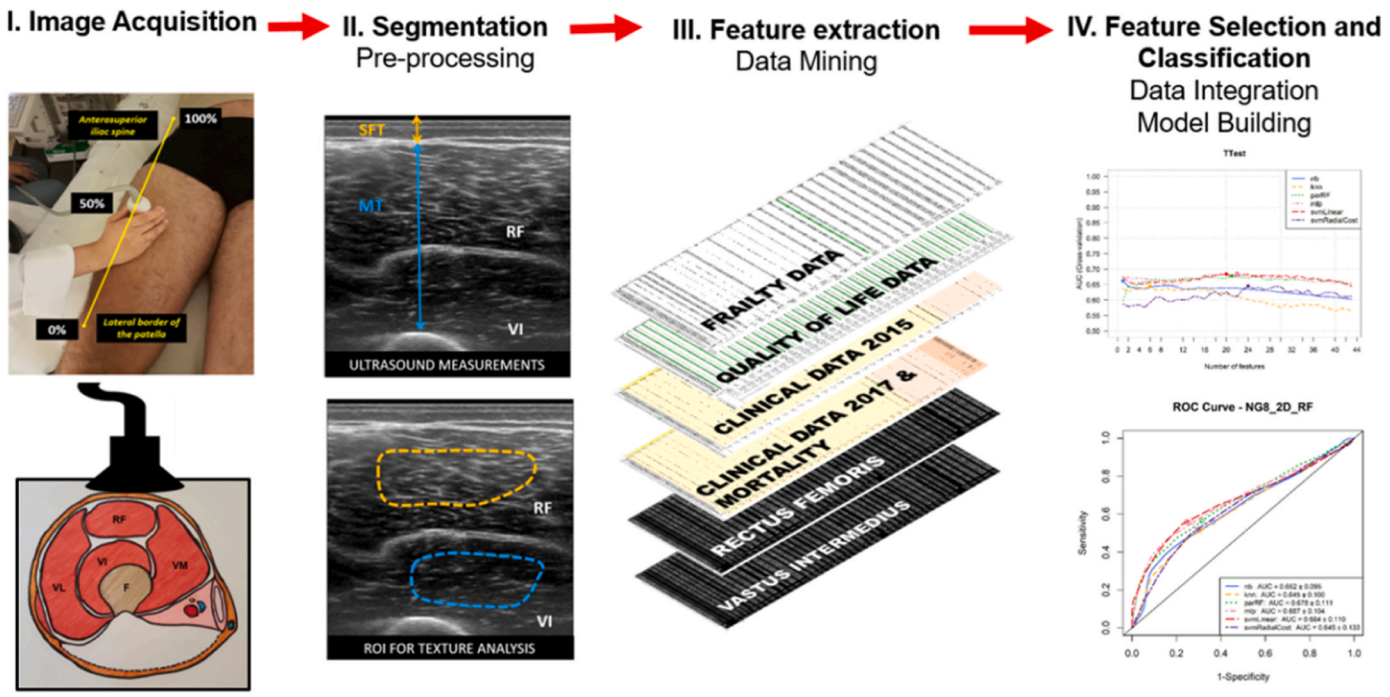


Fig. 1. Flowchart showing the main steps for ultrasound classification using texture analysis. Abbreviations: MT: muscle thickness; SFT: subcutaneous fat thickness; RF: rectus femoris muscle; VI: vastus intermedius.

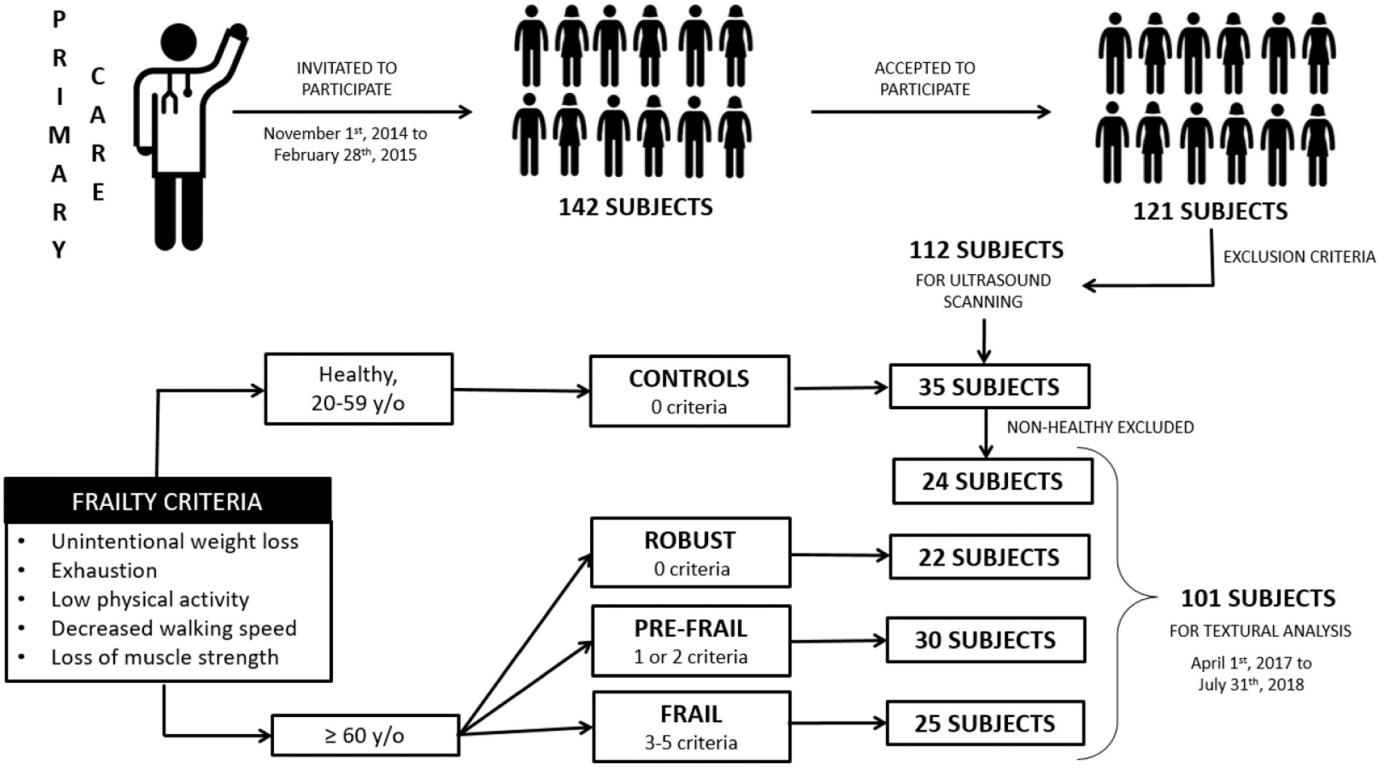


Fig. 2. Flowchart of subject selection and classification performed in the study.

accessible for an ultrasound examination when the patient is already lying for another ultrasound examination of the body; and c) the grand majority of the published body of articles on ultrasound parameters studied for frailty and/or sarcopenia has been performed in the quadriceps muscle which allows for comparisons of current results (Ticinesi et al., 2017; Mirón Mombiela et al., 2020).

Ultrasound measurements and echo intensity were performed according to previously developed standards[5]. Examples of ultrasound images corresponding to each of the different groups of participants are shown in Fig. 3.

To assess the test–retest reliabilities of the person scanning the patients, the Intra-class Correlation Coefficient (ICC) was calculated for

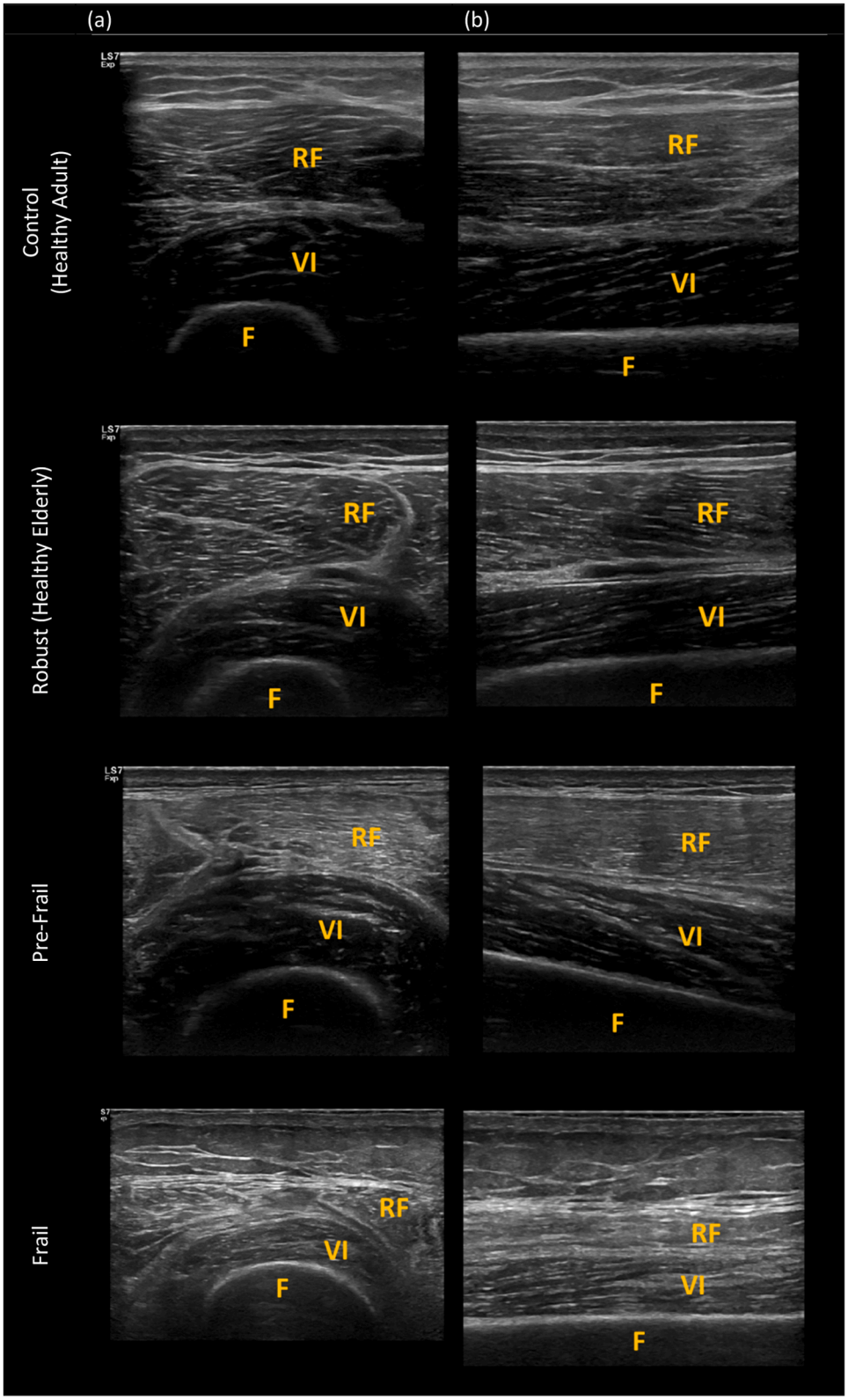


Fig. 3. Examples of ultrasound images corresponding to control, robust, pre-frail and frail participants: (a) axial and (b) sagittal images of the anterior mid-thigh. For the frailty condition, a progressive increase of the echo intensity (increased whiteness) in the rectus femoris and vastus intermedius muscles can be observed; as well as a change in muscle structure, changing from a well-defined striated muscle to a less defined, blurry structure. With the advance of frailty, a progressive decrease in muscle thickness can be also observed. Abbreviations: RF: rectus femoris; VI: vastus intermedius; F: femur.

muscle thickness and subcutaneous fat thickness. The ICC values were 0969 (95%CI: 0957 - 0992) for muscle thickness and 0992 (95%CI: 0984 - 0998) for subcutaneous fat thickness.

2.4. ROI definition (segmentation)

The *rectus femoris* and *vastus intermedius* muscles were manually segmented using an in-house software tool developed in MATLAB (R2017b, MathWorks Inc., Natick, NA, USA). To perform the segmentation of each muscle ROI, a 2D transverse slice was selected and

segmented by a radiologist with three year-experience in muscle ultrasound (R.M.M). The segmented ROI excluded bone, fatty tissue, and muscle fascia (Fig. 1).

Before feature extraction, it was verified that ROI sizes were consistent between groups, since some texture features may be influenced by the region size in the texture analysis process (Chitalia and Kontos, 2019; Sikio et al., 2015). For this purpose and subsequent analysis, the frailty phenotype was dichotomized into non-frail (control and robust groups) and frail (pre-frail and frail groups) (Miron Mombiela et al., 2021). The Mann-Whitney-Wilcoxon test for independent samples was used to determine the *p*-value between the ROI areas of both groups.

Finally, ROI normalization was performed to reduce the influence of image brightness and contrast variation using the $\mu \pm 3\sigma$ method (μ: mean of the gray levels, σ: standard deviation)[30].

2.5. Feature extraction

Texture features were computed using the Radiomics MATLAB package (Vallières et al., 2015). A total of forty-three features were calculated for each ROI based on five statistical methods (Supplemental Table S1): intensity histogram (3 features), gray-level co-occurrence matrix (9 features), gray-level run-length matrix (13 features), gray-level size-zone matrix (13 features) and neighborhood gray-tone difference matrix (5 features).

In matrix-based methods, only one matrix per lesion was computed to achieve rotation-invariant features. For this purpose, the neighboring properties at a one-pixel distance in four directions (0º, 45º, 90º, and 135º) were averaged equally (Vallières et al., 2015).

Finally, all features were standardized to zero mean and unit variance to avoid the differences in the feature scales could affect the model-building process (Kuhn and Johnson (2013) [32].

2.6. Clinical data extraction and outcome measurements

Epidemiological data and details of each patient's medical history were obtained (age, gender, weight, height, body mass index (BMI), frailty criteria, and quality of life). The presence of comorbidities and risk factors was recorded at baseline and followed up with its incidence for two years. The primary outcome measure was death. Medical records were reviewed two years after the ultrasound examination to determine if the patient was still alive. The secondary outcome was comorbidity, and it was calculated according to the Charlson Comorbidity index (Charlson et al., 1987).

2.7. Feature selection and classification: Data integration and model building

Feature selection and classification were then applied to the data extracted from the ultrasound images to choose the models that more accurately classify frail and non-frail samples.

The feature selection step was implemented within the model-building process to avoid overfitting (Ambroise and McLachlan, 2022) and two different methods were evaluated. First, a filter method based on the *p*-value provided by the Mann-Whitney-Wilcoxon test for independent groups of samples was tested. Then, a wrapper method known as support vector machine-recursive feature elimination (SVM-RFE) was also analyzed (Guyon et al., 2002). Using these techniques, different rankings of features ordered according to their discriminative power were obtained and then used by the models to select the optimal number of features.

Six predictive models corresponding to different families of classifiers were studied: naïve Bayes (NB), k-nearest neighbors (KNN), multilayer perceptron (MLP), random forest (RF), and support vector machine (SVM) with the linear and radial kernel.

Each model was evaluated using a leave-group-out test within a nested cross-validation structure (Fig. 4), so datasets were randomly split into training (75%) and testing (25%) a total of 100 times, forming 100 different groups. The training set of each group was used to obtain the ranking of features and these features were progressively added one by one obtaining in this way different subsets of features. An inner 10-fold cross-validation was used to execute the hyperparameter tuning process for each subset of features. For each classifier, the following hyperparameters were evaluated:

- NB: it was selected the Gaussian kernel for estimating the probability density function.
- KNN: the number of neighbors was selected from {1, 3, 5, 7, 9, 11, 13, 15, 17} and it employed the Euclidean distance.
- RF: the number of trees was fixed to 500 and the number of variables considered at each split was selected from {2, 4, 6, 8, 10, 12, 14, 16}.
- MLP: a hidden layer with several nodes selected from {3, 5, 7, 9, 11, 13, 15} was employed.
- SVM: an SVM with linear kernel was evaluated, selecting the cost parameter from $\{2^{-4}.2^{4}\}$. An SVM with Gaussian kernel was also analyzed, selecting the cost from $\{2^{-4}.2^{4}\}$ and the kernel spread from $\{10^{-4}.10^{4}\}$.

Final results were provided by averaging the area under the receiver operating characteristic (ROC) curve over all groups (mean ± CI95%), considering all possible feature subsets. The whole process was developed using R language, version 3.2.5 (R Development Core Team, Vienna, Austria).

2.8. Statistical analysis

Descriptive data are presented with mean ± standard deviation (SD) with the distribution of the data verified by the Kolmogorov-Smirnov normality test. To assess the differences in physical characteristics and the study variables according to sex, t-Student was used for parametric variables and Mann-Whitney for the non-parametric variables. The evaluation of the different variables of the study according to frailty phenotype and control group was determined using ANOVA for the parametric variables (followed by Bonferroni posthoc test) and for the non-parametric Kruskal-Wallis test were used.

The classification performance of the developed models was evaluated using the area under the ROC curve. The Frailty phenotype was dichotomized into non-frail (controls and robust group) and at risk of frailty or frail (pre-frail and frail groups) for the analysis. To quantify the effects of each model on the likelihood of having a frailty phenotype, after adjusting for ultrasound characteristics (subcutaneous fat thickness and muscle thickness) and physical characteristics (age, sex, BMI, gait speed, and muscle strength), a stepwise multiple logistic regression analysis was applied. A multiple logistic model was computed for the outcome measurements. The false discovery rate method by Benjamini & Hochberg was applied as multiple comparisons were performed during the study (Benjamini and Hochberg, 1995).

A p-value < 0.05 was regarded as statistically significant. All statistical analyses were performed with SPSS version 24.0 for Windows (IBM SPSS, Inc., Chicago, IL).

3. Results

3.1. Patient characteristics and clinical indicators

The sample was composed of 101 subjects, 46 females and 55 males, and it was homogeneous regarding age, BMI, gait speed, muscle strength, frailty phenotype, and quality of life according to sex. Baseline characteristics according to sex are summarized in Table 1. The study sample is homogeneous regarding age, BMI, gait speed, muscle strength, and the values of quality of life according to sex. Weight, height, subcutaneous fat tissue, and muscle thickness had statistically significant

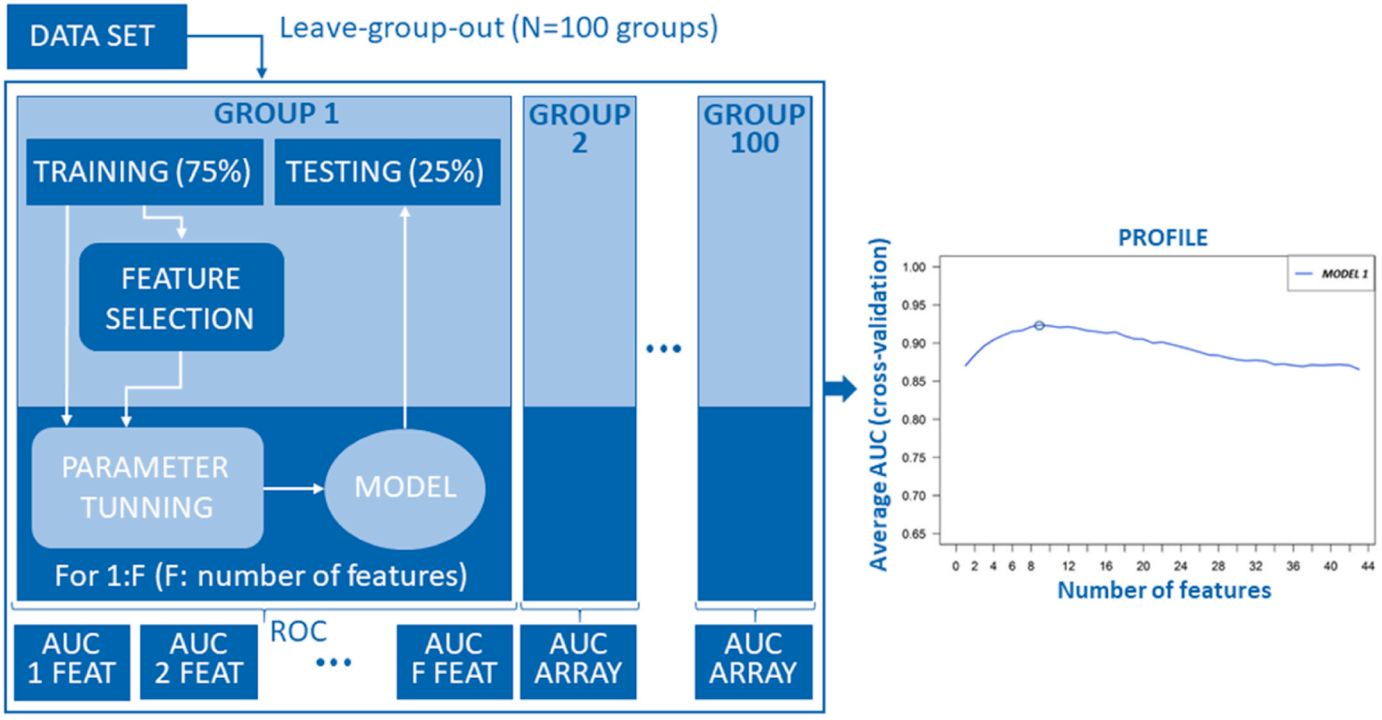


Fig. 4. Nested cross-validation structure developed to evaluate the models. Each dataset of features is randomly split into training (75%) and testing (25%), and this process is repeated 100 times. The models are evaluated with the AUC, considering the different subsets of features obtained when adding progressively one by one the features of the ranking. Abbreviations: AUC: area under the receiver operating characteristic (ROC) curve.

differences between men and women. The distribution of subjects according to the frailty phenotype was 24 controls, 22 robust, 30 pre-frail, and 25 frail patients. Baseline characteristics corresponding to the different groups are shown in Table 2. The data showed that there were statistically significant differences regarding age, height, BMI, gait speed, muscle thickness, and quality of life among the different groups. The data shows that there are statistically significant differences between age, height, BMI, gait speed, muscle thickness, and quality of life, depending on frailty phenotype. The weight and subcutaneous fat thickness were distributed homogeneously within the groups. Moreover, females and males were homogeneously distributed among the groups.

The frequency of comorbidities at baseline and two years later according to sex and frailty phenotype was also determined (Supplemental Tables S2 and S3). The analysis showed that the average number of comorbidities per subject and the estimated 10-year survival were homogenous according to sex at baseline and in follow-up. The average number of comorbidities for each group was different ($p = <0.001$) and increased in frequency at two years follow-ups according to frailty phenotype with an associated reduced estimated 10-year survival for the subjects of 20%. The mortality of the sample after a 2-year follow-up was around 10% and occurred only in the pre-frail and frail groups (Tables 1 and 2).

3.2. Influence of the ROI size

There were no statistically significant differences between non-frail (control and robust) and frail (pre-frail and frail) groups when evaluating the ROI areas selected in the *rectus femoris* muscle (non-frail ROI (mean ± SD): 285.76 ± 88 mm^2; frail ROI (mean ± SD): 256.79 ± 105.94 mm^2; $p = 0.06$). There were also no significant differences in the ROI areas selected in the *vastus intermedius* muscle (non-frail ROI (mean ± SD): 240.03 ± 67.88 mm^2; frail ROI (mean ± SD): 288.03 ± 81.40 mm^2; $p = 0.08$). The ROI sizes were consistent between groups so they should not influence the texture analysis.

3.3. Flailty classification

All features extracted from the *rectus femoris* and the *vastus intermedius* muscles were analyzed with the six classifiers and both feature selection methods. The best classification results in terms of AUC are shown in Table 3.

Regarding the *rectus femoris* muscle, the highest AUC value was obtained using the *p*-value feature selection method and the MLP model, achieving an AUC of 0.69 ± 0.20 (AUC ± 95%CI) with the top 22 features of the ranking. When using the SVM-RFE feature selection method, the maximum AUC values were very similar for all the models, achieving the highest value with the RF model and 29 features (AUC= 0.67 ± 0.19). Although the results were similar there were statistically significant differences between these results and those obtained with the *p*-value method ($p < 0.001$).

Concerning the *vastus intermedius* muscle, all models achieved an AUC > 0.7 using the *p*-value feature method. The best performance was also obtained with the MLP model (AUC = 0.79 ± 0.17), using only the first two features of the ranking. In this case, the best result in terms of AUC using the wrapper method was achieved with the SVM model (AUC = 0.75 ± 0.18), without statistically significant differences ($p = 0.82$) but using all the features of the ranking. Fig. 5 shows the ROC curves obtained for the six models under analysis when using features extracted from both muscle ROIs and both feature selection methods.

Then, the models that achieved the highest AUC in each of the four analyses were selected for the stepwise multiple regression, adjusted for the ultrasound and physical characteristics. In summary, the four-tested predictive models with stepwise multiple logistic regression analysis were statistically significant ($p < 0.001$) for the discrimination of frail vs. non-frail subjects. The models explained between 23% and 100% (Nagelkerke R^2) of the variance in the frailty phenotype and correctly classified from 70% to 87% of cases. They were independent of age, sex, BMI, subcutaneous fat thickness, and muscle strength, with one exception for age. Decreased gait speed was associated with a very high likelihood of exhibiting a frailty phenotype, while increased muscle

Table 1
Characteristics of the Sample According to Sex.

Group	Female (n = 46)	Male (n = 55)	Total (n = 101)	Statistical Test
Variable	Mean ± SD	Mean ± SD	Mean ± SD	p-value
Physical Characteristics				
Age (years)	66 ± 16	64 ± 15	65 ± 15	0.67*
Weight (Kg)	71 ± 14	78 ± 13	74 ± 14	**0.01***
Height (m)	1.58 ± 0.1	1.69 ± 0.1	1.6 ± 0.1	**< 0.001**
BMI (kg/m^2)	28.4 ± 5.1	27.3 ± 4.1	27.8 ± 4.6	0.24
Gait Speed (s)	3.7 ± 1.0	3.5 ± 1.4	3.6 ± 1.2	0.08*
MS (Kg)	27.9 ± 11.8	26.0 ± 11.6	26.9 ± 11.7	0.38
Ultrasound Measurements				
SFT (cm)	1.53 ± 0.50	0.71 ± 0·23	1.08 ± 0.55	**< 0.001***
MT (cm)	2.36 ± 0.73	2.68 ± 0.67	2.53 ± 0.71	**0.03**
Quality of Life (A.U.)	119 ± 6	119 ± 7	119 ± 7	0.83
Frailty Phenotype	***N (%)***	***N (%)***	***N (%)***	0.22*
Control	11 (24)	13 (24)	24 (24)	
Robust	6 (13)	16 (29)	22 (22)	
Pre-frail	15 (33)	15 (27)	30 (30)	
Frail	14 (30)	11 (24)	25 (25)	
Outcomes				
CCI – at baseline	3 (7)	3 (5)	3 (3)	0.72*
Estimated 10-year survival – at baseline	61%	66%	64%	0.69*
CCI – at two-year follow-up	5 (11)	4 (7)	4 (4)	0.49*
Estimated 10-year survival - at two year follow-up	41%	47%	45%	0.54*
Death	5(11)	7 (12)	12 (12)	0.78*

Abbreviations: SD: standard deviation; BMI: body mass index; MS: muscle strength; MT: muscle thickness; SFT: subcutaneous fat thickness; AU: arbitrary units; CCI: Charlson comorbidity index, table shows average of the number of comorbidities presented by each subject and their estimated 10-year survival.
* Variables without normal distribution, nonparametric tests used for the analysis.

thickness decreased that likelihood. The full description of the models can be found in the supplemental materials with their corresponding tables (Tables S4, S5, S6, and S7), and a summary of all the models can be found in Table 4.

To quantify the effects of each model on the likelihood of mortality and comorbidity, a multiple logistic regression analysis was applied. The regression analysis on comorbidity (Table 5a) was statistically significant in 2 out of 4 models ($p = 0.180$ to < 0.001). The models explained between 12% and 64% (Nagelkerke R^2) of the variance of comorbidity. Although the data showed a correct prediction of the outcome from 66% to 82% of the cases, the regression analysis plots revealed that these were associations and not predictions. The regression analysis on mortality (Table 5b) was statistically significant for all models ($p = 0.016$ to < 0.001). The models explained between 39% and 100% (Nagelkerke R^2) of the variance of mortality and correctly predicted 89–100% of cases.

Finally, as we tested the significance of a considerable number of variables, which creates concerns over the accumulated Type 1 error (H. J.W.L. Aerts et al., 2014; Hugo J.W.L. HugoJ.W.L. Aerts et al., 2014), the false discovery rate was used to give reasonable guidance on the validity of the presented results (Fernandez-Lozano et al., 2015). The q-value in this study was 10.8%, which means that approximately 11% of significant results are false positives.

4. Discussion

Frailty is an age-dependent condition in the elderly population and is associated with multiple adverse clinical outcomes, such as falls, disability, or increased morbidity/mortality. Currently, available biomarkers are not good at representing the multifactorial aspect of the condition and are weakly associated with clinically meaningful outcomes. The focus of this study was to evaluate the potential of muscle ultrasound image analysis for the assessment and risk prediction of frailty phenotype. The results obtained indicate that frailty (pre-frail vs. frail) and risk prediction (morbidity/mortality) can be assessed by employing a set of texture features extracted from muscle ultrasound images combined with a machine-learning approach. This was demonstrated in three ways. First, we developed different predictive models able to categorize frailty phenotype as present or absent with moderate to good accuracy ($0.67 \leq AUC \leq 0.79$). Second, the stepwise multiple logistic regression analysis, demonstrated that the developed models correctly classified 70–87% of the cases. Models' accuracy improved and classified correctly 87–100% of the cases when muscle thickness or gait speed was inputted into the models. Third, the models were also associated with increased comorbidity and were highly predictive of death for pre-frail and frail subjects.

To obtain the best subset of features used by the predictive models to achieve the highest AUC values, two feature selection methods were employed, a filter method based on the p-value and a wrapper (SVM-RFE). Although similar outcomes were obtained using both methods, it was observed that the number of features used by the models was lower when using the p-value method than when using the SVM-RFE. Models with fewer features are less complex and better candidates to be generalizable. Therefore, in this case, we consider the feature selection method based on the filter a better option for future analysis.

For this feature selection method, the top 10 features used by all the models were derived mainly from the gray-level run-length matrix (GLRLM) and the gray-level run-length matrix (GLSZM) feature extraction methods. Concretely, 8 out of the 10 in the case of the *rectus femoris* and 9 of the 10 in the case of the *vastus intermedius* muscle. The other 3 features included in the top 10 are derived from the gray-level co-occurrence matrix (GLCM) method. The GLRLM and GLSZM methods measure the distribution of groups of connected pixels with the same gray-level value, providing regional information and defining the heterogeneity properties of these regions. The GLCM method describes the spatial interrelationship between adjacent pixels with different or identical gray-level values, providing information on local gray-level variations. According to these results, it can be interpreted that heterogeneity differences are detected between non-frail and frail subjects and that these differences depend mainly on regional variations of gray levels.

Regarding both muscles analyzed, it is noteworthy that, in general, the highest AUC values have been achieved using features extracted from the *vastus intermedius*, with $AUC > 0.7$ for all models. However, when including the clinical indicators, the percentage of correctly classified cases was higher in the models developed with features extracted from the *rectus femoris*. Despite these differences, the results show that with features extracted from both muscles, it is possible to identify muscle changes between non-frail and frail subjects.

Although a wide range of functional, anthropometric, and biochemical markers are available, they have shown limited clinical applicability (Mitnitski et al., 2015; Calvani et al., 2015). This highlights the idea that there might not be one single biological marker that can reliably track frailty Calvani et al., 2015, Rodríguez-Mañas and Fried, 2015, Erusalimskiy et al., 2016). The presented work is not the first attempt to address this issue. Howlett et al. examined the discriminative ability of the standard frailty index (FI-CSHA) constructed from data obtained during the clinical evaluation and a second frailty index from laboratory data plus systolic and diastolic blood pressure measurements (FI-LAB) which achieved an AUC of 0.72 for FI-LAB and 0.73 for FI-CSHA, respectively. Mitnitski et al. (Mitnitski et al., 2015) examined the effect of a biomarker-based frailty index (FI-B) combining 40 biomarkers of cellular aging, inflammation, hematology, and immunosenescence in predicting mortality. They concluded that the FI-B was more

Table 2
Characteristics of the Sample According to Frailty Phenotype.

Group	Controls (n = 24)	Robust (n = 22)	Pre-frail (n = 30)	Frail (n = 25)	Total (n = 101)	Statistical Test
Variable	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD	p-value
Physical Characteristics						
Age (years)	43 ± 12	68 ± 6	73 ± 7	74 ± 8	65 ± 15	**< 0.001***
Weight (Kg)	73 ± 14	74 ± 11	76 ± 16	75 ± 16	74 ± 14	0.83*
Height (m)	1.67 ± 0.08	1.66 ± 0.08	1.64 ± 0.09	1.58 ± 0.09	1.64 ± 0.09	**0.002**
BMI (kg/m^2)	26.1 ± 3.8	26.8 ± 3.5	28.1 ± 4.2	29.8 ± 5.7	27.8 ± 4.6	**0.02**
Gait Speed (s)	2.7 ± 0.6	2.9 ± 0.5	3.7 ± 0.7	5.0 ± 1.2	3.6 ± 1.2	**< 0.001***
MS (Kg)	30.5 ± 12.5	29.1 ± 13.8	25.2 ± 10.5	23.3 ± 9.3	26.9 ± 11.7	0.20*
Ultrasound Measurements						
SFT (cm)	1.2 ± 0.5	0.9 ± 0.5	1.1 ± 0.5	1.1 ± 0.7	1.1 ± 0.6	0.37*
MT (cm)	3.25 ± 0.59	2.44 ± 0.54	2.41 ± 0.64	2.08 ± 0.54	2.53 ± 0.71	**< 0.001**
Quality of Life (A.U.)	119 ± 5	122 ± 7	120 ± 5	115 ± 8	119 ± 7	**0.001**
Sex	***N (%)***	***N (%)***	***N (%)***	***N (%)***	***N (%)***	0.23*
Female	11 (46)	6 (27)	15 (50)	14 (56)	46 (46)	
Male	13 (54)	16 (73)	15 (50)	11 (44)	55 (55)	
Frailty Criteria						**< 0.001***
No positive criterion	17 (70)	22 (100)	0	0	39 (39)	
1 positive criterion	5 (21)	0	13 (43)	0	18 (18)	
2 positive criteria	2 (8)	0	17 (57)	0	19 (19)	
3 positive criteria	0	0	0	11(44)	11 (11)	
4 positive criteria	0	0	0	7 (28)	7 (7)	
5 positive criteria	0	0	0	7 (28)	7 (7)	
Outcomes						
CCI – at baseline	1 (4)	3 (14)	4 (13,3)	5 (20)	3 (3)	**< 0.001***
Estimated 10-year survival – at baseline	95%	70%	57%	36%	64%	**< 0.001***
CCI – at two-year follow-up	1 (4)	5 (23)	5 (16·7)	7 (28)	4 (4)	**< 0.001***
Estimated 10-year survival - at two-year follow-up	91%	43%	33%	16%	45%	**< 0.001***
Death	0	0	3 (10)	9 (45)	12 (12)	**< 0.001***

Abbreviations: SD: standard deviation; BMI: body mass index; MS: muscle strength; MT: muscle thickness; SFT: subcutaneous fat thickness; AU: arbitrary units; CCI: Charlson comorbidity index, table shows average of the number of comorbidities presented by each subject and their estimated 10-year survival.
* Variables without normal distribution, nonparametric tests used for the analysis.

Table 3
Best classification results obtained for all models using features extracted from the *rectus femoris* and *vastus intermedius* muscles (n = 101).

Feature Selection Method	P-value		SVM - RFE	
Muscles	Rectus Femoris	Vastus Intermedius	Rectus Femoris	Vastus Intermedius
Model				
NB	0.66 ± 0.19	0.77 ± 0.18	0.64 ± 0.20	0.66 ± 0.18
k-NN	0.65 ± 0.20	0.76 ± 0.18	0.60 ± 0.21	0.73 ± 0.20
RF	0.68 ± 0.22	0.74 ± 0.16	**0.67 ± 0.19***	0.72 ± 0.17
MLP	**0.69 ± 0.20***	**0.79 ± 0.17***	0.65 ± 0.21	0.74 ± 0.20
SVM_L	0.68 ± 0.20	0.78 ± 0.17	0.66 ± 0.20	0.74 ± 0.21
SVM_R	0.65 ± 0.26	0.75 ± 0.17	0.64 ± 0.20	**0.75 ± 0.18***

Abbreviations: SVM-RFE: support vector machine - recursive feature elimination; NB: naïve Bayes; k-NN: k-nearest neighbors; RF: random forests; MLP: multilayer perceptron; SVM_L: support vector machine with linear kernel; SVM_R: support vector machine with radial kernel; AUC: area under the curve; CI: confidence interval.
All data are presented as AUC ± 95%CI. The AUC values were classified as poor for values ≤ 0.60, fair for 0.61–0.70, moderate for 0.71–0.80, good for 0.81–0.90 and very good for 0.91–1.00.
* Represents the best model in terms of AUC for each muscle and feature selection method.

powerful for mortality prediction than considering any individual biomarker, with an AUC for FI-B = 0.66. Higher accuracy has been achieved with the approach presented in this work (0.67 ≤ AUC ≤ 0.79), and the developed models also showed a high potential for mortality prediction (89–100% of the cases correctly predicted). In a recent study performed by Wilkinson et al.(Wilkinson et al., 2021), 5 features extracted from the *rectus femoris* using the GLCM method for feature extraction were studied and related to muscle function according to image texture homogeneity changes. They concluded that texture analysis may indicate changes in muscle microstructure and that better muscle quality is associated with values representing greater image texture homogeneity. Our findings are in agreement with these results and, in addition, we have proved that other higher-order statistical methods for feature extraction, such as GLRLM and the GLSZM, play also an important role in detecting muscle changes in the rectus femoris, as well as the *vastus intermedius* muscle.

Therefore, we propose the use of these predictive models as quantitative imaging biomarkers (QIBs) for frailty for several reasons: as increased heterogeneity of the echo intensity texture features reflects frailty-related muscle dysfunction, these anatomical-physiological characteristics are what makes it a quality biomarker (Sullivan et al., 2015). This study can have a clinical impact as imaging is routinely used in clinical practice, providing an opportunity to improve decision-support in frailty treatment and prevention. Imaging can play an important role in the evaluation of geriatric patients as a clinical evaluation of these patients becomes increasingly difficult because of overall frailty, comorbidities, and medication effects (Sadro et al., 2015). As these patients have an increased number of illnesses and use the healthcare system more frequently (Cesari et al., 2016), doctors will inevitably ask for more imaging studies, a trend that is also on the rise. The developed methodology provides a non-invasive, low-cost, and accurate way of assessing frailty, being able to detect nuances that humans cannot.

The use of machine learning and artificial intelligence is also beginning to show promising results in medical data pre-processing for predictive modeling and risk factor discovery for frailty phenotype (Hassler et al., 2019; Kruse et al., 2018). There is no question that the application of advanced image processing and analysis procedures (e.g.

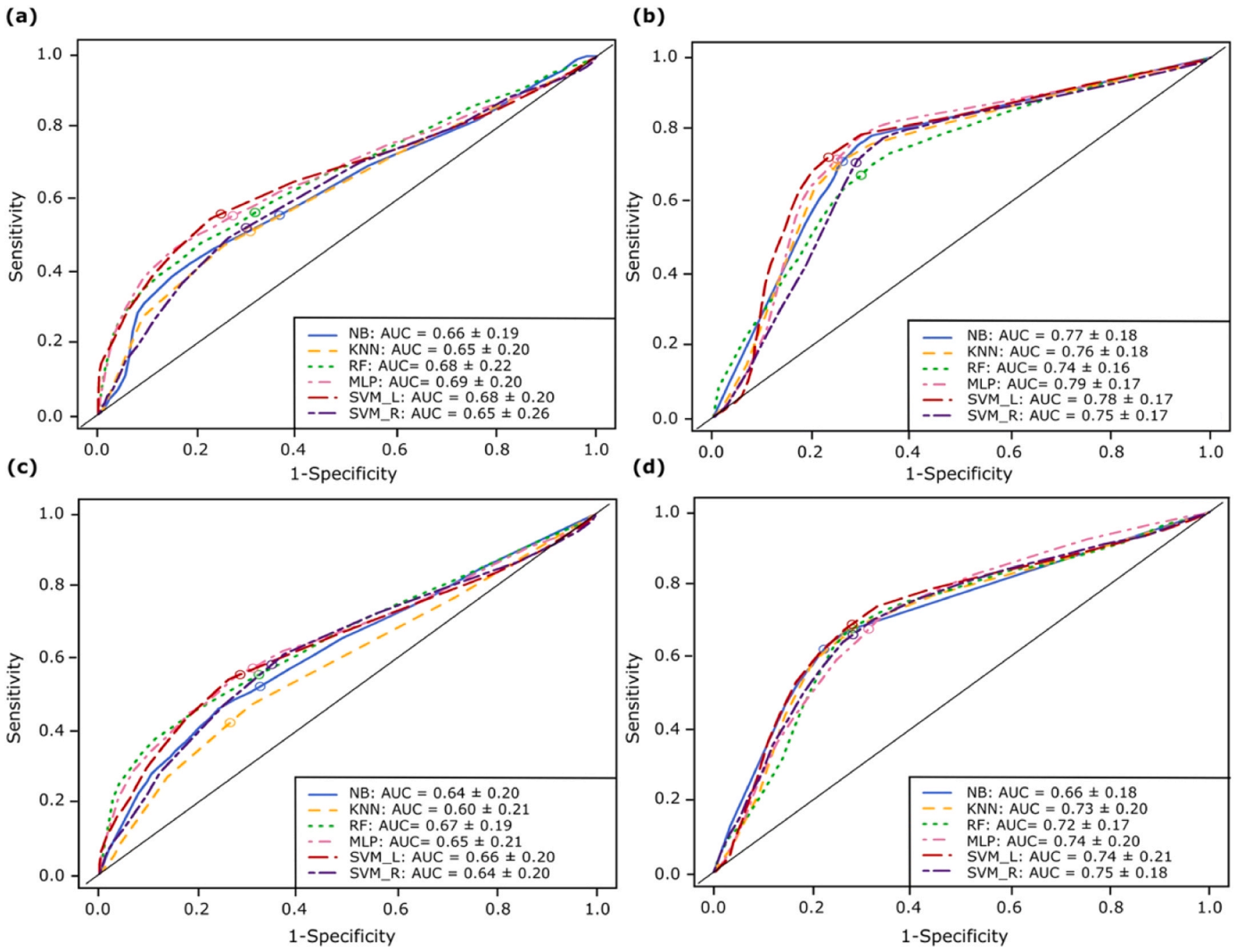


Fig. 5. Average ROC curves for the best subsets of features obtained with the six models under analysis when using: (a) features extracted from the rectus femoris muscle and the p-value feature selection method or (c) the SVM-RFE method; (b) features extracted from the vastus intermedius muscle and the p-value feature selection method or (d) the SVM-RFE method. The highlighted points on the curves indicate the optimal cut-off points that weigh sensitivity and specificity equally.

semi-automated or automated segmentation and registration) provide for a more robust and efficient incorporation of QIBs into clinical and preclinical decision-making (Aerts, Hugo J W L et al., 2014; Gillies et al., 2016; Kumar et al., 2012; Lambin et al., 2012). This includes the development of apps, artificial intelligence (AI) software, and/or automatization of QIBs that can be added directly to the pipeline and workflow of imaging studies already performed in the Clinical Setting. For example, the automatic inclusion of our frailty models for muscle quality in radiology reports for patients over 60 years old that get an ultrasound in our emergency department, or at the ultrasound outpatient clinic, could be considered. This will help determine if the patients have an increased risk of frailty and adverse outcome, which in turn could be useful for patient prognosis, and influence any medical or surgical decision-making. This approach may promote the early detection of otherwise subclinical or unsuspected frailty, otherwise known as opportunistic screening; diagnostic assessment of clinically manifest frailty, and risk stratification of subjects with a suspected or confirmed diagnosis.

According to the World Health Organization, the purpose of screening is to identify people in an apparently healthy population who are at higher risk of a health problem or condition so that early treatment or intervention can be offered and thereby reduce the incidence and/or mortality of the health problem within the population. Within the realm of imaging, opportunistic screening has more recently been used to describe the practice of systematically leveraging imaging data that are incidental to the clinical indication for obtaining the study (Pickhardt et al., 2023). Opportunistic screening that delivers on the promise of predictive data for better risk stratification, like the QIBs presented in this study, may help to rebalance the value equation in favor of patients (Pickhardt et al., 2023). Assuming that the value added by opportunistic ultrasound-based screening is firmly established, it begs the question of whether this puts "intentional" screening back in play. The emergence of fully automated artificial intelligence solutions should allow for efficient and objective US-based assessment, as well as broad application to large, diverse patient populations, which will further refine appropriate risk stratification. Demonstrating measurable improvements in population health outcomes, associated with reduced costs through disease prevention, should be attractive to both patients and healthcare systems (Pickhardt et al., 2022).

The use of ultrasound was recently expanded into daily clinical workups to assess both muscle quantity and quality. The European Geriatric Medicine Society recently proposed a consensus protocol for using ultrasound for muscle mass assessments (Cruz-Jentoft et al., 2019). Although it remains inconsistently diagnosed, once frailty sets in, the therapeutic approach is typically conservative; management is primarily dietary supplements and physical activity. This scenario makes the case

Table 4
Stepwise Multiple Logistic Regression Frailty Models.

Feature Selection Method	Models	Steps	Model Chi-Square [df]	% Correct Predictions	Hosmer y Lemeshow Test	Nagelkerke-R^2
P-Value	*Frailty Model 1*	Step 1: Echo Intensity Texture Analysis	51.6 [22], p = < 0.001	80	0.96	0.54
		Step 2: Adjustment with Ultrasound Characteristics	62.8 [24], p = < 0.001	85	0.07	0.62
		Step 3: Adjustment with Physical Characteristics	106.3 [29], p = < 0.001	96	0.24	0.87
	Frailty Model 2	Step 1: Echo Intensity Texture Analysis	19.1 [2], p = < 0.001	70	0.03	0.23
		Step 2: Adjustment with Ultrasound Characteristics	23.6 [4], p = < 0·001	71	0.62	0.28
		Step 3: Adjustment with Physical Characteristics	81.4 [9], p = < 0.001	87	0.334	0.740
SVM-RFE	*Frailty Model 3*	Step 1: Echo Intensity Texture Analysis	73.07 [29], p = < 0.001	87	0.32	0.69
		Step 2: Adjustment with Ultrasound Characteristics	92.6 [31], p = < 0·001	91	0.43	0.80
		Step 3: Adjustment with Physical Characteristics	139·2 [36], p = 0·002	100	1.00	1.00
	Frailty Model 4 *	Step 1: Echo Intensity Texture Analysis	93.3 [43], p = < 0·001	0.81	0.98	87
		Step 2: Adjustment with Ultrasound Characteristics	NA	NA	NA	NA
		Step 3: Adjustment with Physical Characteristics	NA	NA	NA	NA

Abbreviations: SVM - RFE: support vector machine – recursive feature elimination.
*Steps 2 and 3 of the stepwise multivariate logistic analysis could not run in the SPSS program due to the high number of variables for this model (43 texture features), for a relatively small number of subjects (n = 101).

Table 5a
Multiple Logit Regression Analysis of Frailty Models for Comorbidity Outcome.

Feature Selection Method	Models	Model Chi-Square [df]	% Correct Predictions	Hosmer y Lemeshow Test	Nagelkerke-R^2	Plot
P-Value	Frailty Model 1	27.9 [22], $p = 0.18$	78	0.22	0.33	Association
	Frailty Model 2	**9.4 [2], $p = 0.01$**	66	0.70	0.12	Association
SVM-RFE	Frailty Model 3	39.4 [29], $p = 0.09$	73	0.54	0.44	Association
	Frailty Model 4	**66.2 [43], $p = 0.02$**	83	0.90	0.64	Predictive

Table 5b
Multiple Logit Regression Analysis of Frailty Models for Mortality Outcome.

Feature Selection Method	Frailty Textures Analysis Models	Model Chi-Square [df]	% Correct Predictions	Hosmer y Lemeshow Test	Nagelkerke-R^2	Plot
P-Value	Frailty Model 1	**38.6 [22], $p = 0.02$**	91.1	0.90	0.61	Predictive
	Frailty Model 2	**22.4 [2], $p = <0.001$**	89.1	0.83	0.39	Predictive
SVM-RFE	Frailty Model 3	**73.6 [29], $p = <0.001$**	100	1.00	1.00	Predictive
	Frailty Model 4	**73.6 [43], $p = 0.002$**	100	1.00	1.00	Predictive

Abbreviations: SVM - RFE: support vector machine – recursive feature elimination.

for a more proactive approach to intervention because if frailty is diagnosed early, it can be managed inexpensively and successfully, but if left undiagnosed; its impact can be progressively debilitating (Miron Mombiela et al., 2020). In the long term, it might be possible to track frailty over time with ultrasound imaging, select an appropriate therapeutic intervention, and monitor response to treatment, including exercise and/or physical therapy (Miron Mombiela et al., 2020). However, we believe using the developed frailty models in the clinical setting would need additional validation. We would propose enhancing the models by using a larger dataset of subjects during the training process. Additionally, we suggest validating the algorithms using an independent dataset to ensure they perform well and are reliable in different real-world clinical scenarios. We consider this approach necessary for establishing the credibility and effectiveness of the frailty models before implementing them in clinical practice.

A result that requires further explanation or discussion is why decreased gait speed was associated with a very high likelihood of exhibiting frailty phenotype, while increased MT decreased that likelihood within the context of the developed frailty models. In the stepwise multiple analysis regression, we included muscle thickness in the second step and gait speed in the third step, so we could evaluate the degree of effect each variable had over the overall predictive models. It was found that muscle thickness, although it was a significant variable, its overall effect was smaller in comparison with gait speed. Although both variables add important information to the predictive models and improve their discriminative capacity, gait speed played a much greater role than muscle thickness (Table 4 and Table S4-S7, in the supplemental material). We believe this is because gait speed represents the physical performance of the subjects, which refers to their capacity to function and the muscle quality it poses to perform those functions. Muscle characteristics beyond size are known to affect muscle function and strength and contribute to mobility limitation Pillen and van Alfen (2011)).

We must also underline that another explanation for this result is that of the Fried items, gait speed appears to be the strongest predictor over the other four items (unintentional weight loss, self-reported

exhaustion, weakness, and low physical activity) (Afilalo, 2011; Jung et al., 2018; Silva et al., 2016). Gait speed can predict life expectancy, functional dependency, and institutionalization in older adults, and it has been used as a marker for physical fitness (Jung et al., 2018; Sanchis et al., 2014). A study performed among elderly Brazilians evaluated the contribution of each item to determine the frailty phenotype. They concluded that when gait speed was positive, subjects were more likely to develop frailty (OR = 10.50, $p < 0.001$); the same for muscle strength (OR = 7.31, $p < 0.001$), but to a lesser degree (Silva et al., 2016). Muscle strength in our study was not significantly associated with the predictive models, a discrepancy when compared to other studies. Not only because muscle weakness is another Fried item, but also because previously published ultrasound studies on muscle echogenicity have found increased echo intensity to be inversely related to grip strength and general measures of muscle performance (Akima et al., 2017; Cadore et al., 2012; Fukumoto et al., 2012; Nishihara et al., 2014; Rech et al., 2014; 5Watanabe et al., 2013). We believe the reason for this was that muscle strength measured by hand dynamometer indicated the strength of the arm and not the femoral quadriceps that was measured by ultrasound. This may also explain why in this study gait speed had such a high correlation with the predictive models, one that was lacking with muscle strength. Lower limbs are more relevant for physical function than upper limbs, even when handgrip strength has been widely used and is well correlated with the most relevant outcomes (Chan et al., 2014).

In summary, the application of machine learning and artificial intelligence for advanced image processing and analysis provides a more robust and efficient incorporation of QIBs into clinical and preclinical decision-making (Sogawa et al., 2017; Arts et al., 2007, Aerts et al., 1995, Lambin et al., 2012). This will help determine if the patients have an increased risk of frailty and adverse outcomes, which in turn could be useful for patient prognosis, and influence medical or surgical decision-making (Bentov et al., 2019). This approach may promote opportunistic screening of otherwise subclinical or unsuspected frailty; diagnostic assessment of clinically manifested frailty, and risk stratification of subjects with a suspected or confirmed diagnosis.

4.1. Limitations

We still need a better understanding of the clinical implications of the use of this technology and findings must be interpreted considering several limitations. First, larger longitudinal studies in diverse populations are needed to determine whether changes over time in echo intensity are meaningful and whether interventions that improve muscle dysfunction can be followed up by ultrasound. Second, it was a single sonographer, and the same ultrasound machine was used throughout the study. Therefore, the capability of using echo intensity texture features from ultrasound images, and the influence of a plurality of devices remains unknown. Third, the development of textural analysis signatures with the use of machine learning due to the high number of variables, also known as high dimensionality and small n-to-p data bias, raises concern for its generalizability and repeatability as sources of variation can exist in each step of the workflow. In addition, unrecognized confounding variables in the database used are a concern even when we performed a robust statistical validation to avoid bias and over-fitting of the predictive models. Hence, the present study must be validated against a completely independent data set, preferably from another institution[20]. Fourth, the current study only used Fried frailty criteria to identify the subjects, future ultrasound studies would benefit from comparing several frailty tools, including those that rely less on physical attributes. And lastly, although rigorous inclusion and exclusion criteria were applied to the patient selection, bias or systematic error could have been introduced as patients were recruited from the ultrasound section of our Radiology Department, in comparison to patients attending the general practitioner office or community-dwelling seniors.

5. Conclusion

In conclusion, the presented results show that 2D features extracted from muscle ultrasound images can be used as a novel indicator of muscle quality, assessing with good accuracy risk prediction of frailty phenotype. These promising results denote that, soon, the combination of texture-based muscle ultrasound image analysis and clinical data in a machine-learning approach could be used for both opportunistic assessments of frailty and risk stratification in the elderly population.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Funding

This work was supported by the following grants: Grant PID2020-113839RB-I00 funded by MCIN/AEI/10.13039/501100011033 to C.B. DM acknowledges financial support from the Conselleria d'Educació, Investigació, Cultura i Esport, Generalitat Valenciana (grants AEST/2018/021 and AEST/2019/037).

CRediT authorship contribution statement

Rebeca Mirón Mombiela: Conceptualization, Investigation, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Silvia Ruiz-España**: Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. **David Moratal**: Methodology, Validation, Supervision, Writing – review & editing. **Consuelo Borras:** Conceptualization, Validation, Resources, Writing – review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors do not have any conflict of interest to disclose.

Data availability

The data that has been used is confidential.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.mad.2023.111860.

References

Aerts, H.J.W.L., Velazquez, E.R., Leijenaar, R.T.H., et al., 2014. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat. Commun. 5, 4006. https://doi.org/10.1038/ncomms5006.

Aerts, Hugo J.W.L., Velazquez, E.R., Leijenaar, R.T.H., Parmar, C., Grossmann, P., Carvalho, S., Lambin, P., 2014. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat. Commun. 5, 4006. https://doi.org/10.1038/ncomms5006.

Afilalo, J., 2011. Frailty in patients with cardiovascular disease: why, when, and how to measure. Curr. Cardiovasc. Risk Rep. 5 (5), 467–472. https://doi.org/10.1007/s12170-011-0186-0.

Akima, H., Yoshiko, A., Tomita, A., et al., 2017. Relationship between quadriceps echo intensity and functional and morphological characteristics in older men and women. Arch. Gerontol. Geriatr. 70, 105–111. https://doi.org/10.1016/j.archger.2017.01.014.

Ambroise, C., McLachlan, G.J., 2002. Selection bias in gene extraction on the basis of microarray gene-expression data. Proc. Natl. Acad. Sci. 99, 6562–6566. https://doi.org/10.1073/pnas.102102699.

Arts, I.M.P., Pillen, S., Overeem, S., et al., 2007. Rise and fall of skeletal muscle size over the entire life span. J. Am. Geriatr. Soc. 55, 1150–1152. https://doi.org/10.1111/j.1532-5415.2007.01228.x.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bentov, I., Kaplan, S.J., Pham, T.N., et al., 2019. Frailty assessment: from clinical to radiological tools. Br. J. Anaesth. 123, 37–50. https://doi.org/10.1016/j.bja.2019.03.034.

Cadore, E.L., Izquierdo, M., Conceicao, M., Radaelli, R., Pinto, R.S., Baroni, B.M., Kruel, L.F., 2012. Echo intensity is associated with skeletal muscle power and cardiovascular performance in elderly men. Exp. Gerontol. 47 (6), 473–478. https://doi.org/10.1016/j.exger.2012.04.002.

Calvani, R., Marini, F., Cesari, M., et al., 2015. Biomarkers for physical frailty and sarcopenia: state of the science and future developments. J. Cachex. Sarcopenia Muscle 6, 278–286. https://doi.org/10.1002/jcsm.12051.

Cesari, M., Prince, M., Thiyagarajan, J.A., et al., 2016. Frailty: an emerging public health priority. J. Am. Med Dir. Assoc. 17, 188–192. https://doi.org/10.1016/j.jamda.2015.12.016.

Chan, O.Y.A., van Houwelingen, A.H., Gussekloo, J., Blom, J.W., den Elzen, Wendy P.J., 2014. Comparison of quadriceps strength and handgrip strength in their association with health outcomes in older adults in primary care. Age 36 (5), 9714. https://doi.org/10.1007/s11357-014-9714-4.

Charlson, M.E., Pompei, P., Ales, K.L., et al., 1987. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J. Chronic Dis. 40, 373–383. https://doi.org/10.1016/0021-9681(87)90171-8.

Chitalia, R.D., Kontos, D., 2019. Role of texture analysis in breast MRI as a cancer biomarker: a review. J. Magn. Reson Imaging 49, 927–938. https://doi.org/10.1002/jmri.26556.

Cruz-Jentoft, A.J., Bahat, G., Bauer, J., Boirie, Y., Bruyère, O., Cederholm, T., Cooper, C., Landi, F., Rolland, Y., Sayer, A.A., Schneider, S.M., Sieber, C.C., Topinkova, E., Vandewoude, M., Visser, M., Zamboni, M., Writing Group for the European Working Group on Sarcopenia in Older People 2 (EWGSOP2), and the Extended Group for EWGSOP2, 2019. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 48 (1), 16–31. https://doi.org/10.1093/ageing/afy169.

Fernandez-Lozano, C., Seoane, J.A., Gestal, M., et al., 2015. Texture classification using feature selection and kernel-based techniques. Soft Comput. 19, 2469–2480. https://doi.org/10.1007/s00500-014-1573-5.

Fried, L.P., Tangen, C.M., Walston, J., et al., 2001. Frailty in older adults: evidence for a phenotype. J. Gerontol. A Biol. Sci. Med. Sci. 56, M146–M156. https://doi.org/10.1093/gerona/56.3.m146.

Fritz, B., Müller, D.A., Sutter, R., et al., 2018. Magnetic resonance imaging–based grading of cartilaginous bone tumors. Investig. Radio. 53 (11), 663–672. https://doi.org/10.1097/RLI.0000000000000486.

Fukumoto, Y., Ikezoe, T., Yamada, Y., et al., 2012. Skeletal muscle quality assessed from echo intensity is associated with muscle strength of middle-aged and elderly persons. Eur. J. Appl. Physiol. 112, 1519–1525. https://doi.org/10.1007/s00421-011-2099-5.

Gillies, R.J., Kinahan, P.E., Hricak, H., 2016. Radiomics: images are more than pictures, they are data. Radiology 278 (2), 563–577. https://doi.org/10.1148/radiol.2015151169.

Gruther, W., Benesch, T., Zorn, C., Paternostro-Sluga, T., Quittan, M., Fialka-Moser, V., Crevenna, R., 2008. Muscle wasting in intensive care patients: ultrasound observation of the M. quadriceps femoris muscle layer. J. Rehabil. Med. 40 (3), 185–189. https://doi.org/10.2340/16501977-0139.

Guyon, I., Weston, J., Barnhill, S., et al., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn 46, 389–422. https://doi.org/10.1023/A:1012487302797.

Hassler, A.P., Menasalvas, E., García-García, F.J., et al., 2019. Importance of medical data preprocessing in predictive modeling and risk factor discovery for the frailty syndrome. BMC Med. Inf. Decis. Mak. 19, 33. https://doi.org/10.1186/s12911-019-0747-6.

Jung, H., Kang, M., Choi, J., Yoon, S., Kim, S., Kim, K., Kim, C., 2018. Simple method of screening for frailty in older adults using a chronometer and tape measure in clinic. J. Am. Geriatr. Soc. 66 (1), 157–160. https://doi.org/10.1111/jgs.15204.

Kawai, H., Kera, T., Hirayama, R., Hirano, H., Fujiwara, Y., Ihara, K., Obuchi, S., 2018. Morphological and qualitative characteristics of the quadriceps muscle of community-dwelling older adults based on ultrasound imaging: Classification using latent class analysis. Aging Clin. Exp. Res. 30 (4), 283–291. https://doi.org/10.1007/s40520-017-0781-0.

Kruse, C., Goemaere, S., De Buyser, S., et al., 2018. Predicting mortality and incident immobility in older Belgian men by characteristics related to sarcopenia and frailty. Osteoporos. Int 29, 1437–1445. https://doi.org/10.1007/s00198-018-4467-z.

Kuhn, M., Johnson, K., 2013. Data Pre-processing. Applied Predictive Modeling. Springer,, New York, NY, pp. 27–59. https://doi.org/10.1007/978-1-4614-6849-3.

Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S.A., Schabath, M.B., Gillies, R.J., 2012. Radiomics: the process and the challenges. Magn. Reson. Imaging 30 (9), 1234–1248. https://doi.org/10.1016/j.mri.2012.06.010.

Lambin, P., Rios-Velazquez, E., Leijenaar, R., et al., 2012. Radiomics: extracting more information from medical images using advanced feature analysis. Eur. J. Cancer 48, 441–446. https://doi.org/10.1016/j.ejca.2011.11.036.

Larroza, A., Bodí, V., Moratal, D., 2016. Texture Analysis in Magnetic Resonance Imaging: Review and Considerations for Future Applications. In: Assessment of cellular and organ function and dysfunction using direct and derived MRI methodologies, 75. InTech,, Rijeka, Croatia, p. 106. https://doi.org/10.1093/ehjci/jeaa028.

M. Sikio K.K. Holli-Helenius P. Ryymin et al. The effect of region of interest size on textural parameters 9th Int. Symp. Image Signal Process. Anal. (ISPA) 2015 149 153 doi: 10.1109/ISPA.2015.7306049.

Materka, A., Strzelecki, M., 1998. Texture Analysis Methods-A Review. Tech Univ lodz, Inst. Electron. COST B11 report,, Brussels, pp. 9–11.

Mirón Mombiela, R., Facal de Castro, F., Moreno, P., et al., 2017. Ultrasonic echo intensity as a new noninvasive in vivo biomarker of frailty. J. Am. Geriatr. Soc. 65, 2685–2690.

Mirón Mombiela, R., Vucetic, J., Rossi, F., et al., 2020. Ultrasound biomarkers for Sarcopenia: what can we tell so far. Semin Musculoskelet. Radio. 24, 181–193. https://doi.org/10.1055/s-0039-3402745.

Mirón Mombiela, R., Vucetic, J., Monllor, P., Cárdenas-Herrán, J.S., Taltavull de La Paz, P., Borrás, C., 2021. Diagnostic performance of muscle echo intensity and fractal dimension for the detection of frailty phenotype. Ultrason Imaging Nov. 43 (6), 337–352. https://doi.org/10.1177/01617346211029656.

Mitnitski, A., Collerton, J., Martin-Ruiz, C., et al., 2015. Age-related frailty and its association with biological markers of ageing. BMC Med. 13, 161. https://doi.org/10.1186/s12916-015-0400-x.

Mueller, N., Murthy, S., Tainter, C., Lee, J., Riddell, K., Fintelmann, F., Eikermann, M., 2016. Can sarcopenia quantified by ultrasound of the rectus femoris muscle predict adverse outcome of surgical intensive care unit patients as well as frailty? A prospective, observational cohort study. Ann. Surg. 264 (6), 1116–1124. https://doi.org/10.1097/SLA.0000000000001546.

Nishihara, K., Kawai, H., Hayashi, H., Naruse, H., Kimura, A., Gomi, T., Hoshi, F., 2014. Frequency analysis of ultrasonic echo intensities of the skeletal muscle in elderly and young individuals. Clin. Interv. Aging 9, 1471–1478. https://doi.org/10.2147/CIA.S67820.

Nodera, H., Sogawa, K., Takamatsu, N., et al., 2018. Age-dependent texture features in skeletal muscle ultrasonography. J. Med. Investig. 65, 274–279. https://doi.org/10.2152/jmi.65.274.

Nodera, H., Sogawa, K., Takamatsu, N., et al., 2019. Texture analysis of sonographic muscle images can distinguish myopathic conditions. J. Med. Investig. 66, 237–247. https://doi.org/10.2152/jmi.66.237.

Ortiz-Ramón, R., Larroza, A., Ruiz-España, S., et al., 2018. Classifying brain metastases by their primary site of origin using a radiomics approach based on texture analysis: a feasibility study. Eur. Radio. 28, 4514–4523. https://doi.org/10.1007/s00330-018-5463-6.

Pickhardt, P.J., Summers, R.M., Garrett, J.W., Krishnaraj, A., Agarwal, S., Dreyer, K.J., Nicola, G.N., 2023. Opportunistic screening: radiology scientific expert panel (Jun). Radiology 307 (5), e222044. https://doi.org/10.1148/radiol.222044.

Pillen, S., van Alfen, N., 2011. Skeletal muscle ultrasound. Neurol. Res. 33 (10), 1016–1024. https://doi.org/10.1179/1743132811Y.0000000010.

Rech, A., Radaelli, R., Goltz, F.R., da Rosa, L.H., Schneider, C.D., Pinto, R.S., 2014. Echo intensity is negatively associated with functional capacity in older women. Epub 2014 Aug 29 Age 36 (5), 2. https://doi.org/10.1007/s11357-014-9708-2.

Ríos-Díaz, J., del Baño-Aledo, M.E., Tembl-Ferrairó, J.I., et al., 2019. Quantitative neuromuscular ultrasound analysis as biomarkers in amyotrophic lateral sclerosis. Eur. Radio. 29, 4266–4275. https://doi.org/10.1007/s00330-018-5943-8.

Rodriguez-Mañas, L., Fried, L.P., 2015. Frailty in the clinical scenario. Lancet 385, e7–e9. https://doi.org/10.1016/S0140-6736(14)61595-6.

Rodríguez-Mañas, L., 2015. Use of biomarkers. J. Frailty Aging 4 (3), 125–128. https://doi.org/10.14283/jfa.2015.46.

Rodríguez-Mañas, L., Féart, C., Mann, G., et al., 2013. Searching for an operational definition of frailty: a delphi method based consensus statement. The frailty operative definition-consensus conference project. J. Gerontol. A Biol. Sci. Med. 68, 62–67. https://doi.org/10.1093/gerona/gls119.

Sadro, C.T., Sandstrom, C.K., Verma, N., et al., 2015. Geriatric trauma: a radiologist's guide to imaging trauma patients aged 65 years and older. Radiographics 35, 1263–1285. https://doi.org/10.1148/rg.2015140130.

Sanchis, J., Bonanad, C., Ruiz, V., Fernández, J., García-Blas, S., Mainar, L., Núñez, J., 2014. Frailty and other geriatric conditions for risk stratification of older patients with acute coronary syndrome. Am. Heart J. 168 (5), 784–791. https://doi.org/10.1016/j.ahj.2014.07.022.

Silva, Silvia Lanziotti Azevedo da, Neri, A.L., Ferrioli, E., Lourenço, R.A., Dias, R.C., 2016. Phenotype of frailty: The influence of each item in determining frailty in community-dwelling elderly - the fibra study. Cienc. Saude Coletiva 21 (11), 3483–3492. https://doi.org/10.1590/1413-812320152111.23292015.

Sogawa, K., Nodera, H., Takamatsu, N., et al., 2017. Neurogenic and myogenic diseases: quantitative texture analysis of muscle US data for differentiation. Radiology 283, 492–498. https://doi.org/10.1148/radiol.2016160826.

Sullivan, D.C., Obuchowski, N.A., Kessler, L.G., et al., 2015. Metrology standards for quantitative imaging biomarkers. Radiology 277, 813–825. https://doi.org/10.1148/radiol.2015142202.

Ticinesi, A., Meschi, T., Narici, M.V., Lauretani, F., Maggio, M., 2017. Muscle ultrasound and sarcopenia in older individuals: a clinical perspective. J. Am. Med. Dir. Assoc. 18 (4), 290–300. https://doi.org/10.1016/j.jamda.2016.11.013.

Vallières, M., Freeman, C.R., Skamene, S.R., et al., 2015. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys. Med. Biol. 60, 5471–5496. https://doi.org/10.1088/0031-9155/60/14/5471.

Wang, S., Summers, R.M., 2012. Machine learning and radiology. Med. Image Anal. 16, 933–951. https://doi.org/10.1016/j.media.2012.02.005.

Watanabe, Y., Yamada, Y., Fukumoto, Y., et al., 2013. Echo intensity obtained from ultrasonography images reflecting muscle strength in elderly men. Clin. Inter. Aging 8, 993–998. https://doi.org/10.2147/CIA.S47263.

Wilkinson, T.J., Ashman, J., Baker, L.A., et al., 2021. Quantitative muscle ultrasonography using 2D textural analysis: a novel approach to assess skeletal muscle structure and quality in chronic kidney disease. Ultrason Imaging 43, 139–148. https://doi.org/10.1177/01617346211009788.

Yamada, M., Kimura, Y., Ishiyama, D., Nishio, N., Abe, Y., Kakehi, T., Arai, H., 2017. Differential characteristics of skeletal muscle in community-dwelling older adults. J. Am. Med. Dir. Assoc. 18 (9), 807.e16. S1525-8610(17)30278-5.