Identification of new biomarkers for sarcopenia and characterization of cathepsin D biomarker

Corine L'hôte[1], Benoît Cordier[1], Alain Labasse[1], Christelle Boileau[1], Bérénice Costes[1] & Yves Henrotin[1,2,3]* 🆔

[1]*Artialis Group SA, GIGA Tower, Avenue de l'hôpital, Liège, 4000, Belgium,* [2]*Musculoskeletal Innovative Research Lab (mSKIL), University of Liège, Arthropôle, Institute of Pathology, Sart-Tilman, Liège, Belgium,* [3]*Department of Physical Therapy and Rehabilitation, VIVALIA, Princess Paola Hospital, Marche-en-Famenne, Belgium*

Abstract

Background   Sarcopenia is the progressive generalized loss of skeletal muscle mass, strength, and function that occurs with aging. This study was undertaken to identify new biomarkers of sarcopenia by proteomics analysis of female sera.
Methods   A case–control study was set up, for which 19 sarcopenic subjects and 20 control subjects, according to the European Working Group on Sarcopenia Older People criteria published in 2010 (EWGSOP1), were enrolled. All the subjects were at least 65 years old and in majority female. Biomarker screening was performed by a comparative mass spectrometry analysis. Protein expression levels between the two groups were compared. One of the identified biomarkers, cathepsin D, was measured by immunoassay on the serum of the full sample set (n = 39). Its diagnostic performance was evaluated with a receiver operating characteristic curve (ROC curve).
Results   Two biomarkers were identified: fructose-biphosphate aldolase A ($P \leq 0.05$) and cathepsin D ($P \leq 0.05$). The levels of all of them were higher in sarcopenic patients. It was confirmed by immunoassay that cathepsin D levels in serum were significantly higher in the sarcopenic group of patients (P = 0.038). An inverse correlation (-0.385) was observed between cathepsin D levels in serum and gait speed. The area under the ROC curve measurement (AUC = 0.696) demonstrated that cathepsin D levels could discriminate between sarcopenic and non-sarcopenic subjects. A predictive model including cathepsin D, age, and body mass index was established to improve the diagnostic performance (AUC = 0.908).
Conclusions   Cathepsin D has been identified as a diagnostic biomarker of sarcopenia.

Keywords   Sarcopenia; Proteomics; Cathepsin D; Muscle; Diagnosis; Biomarker

Introduction

Sarcopenia is a disease of the elderly, whose physiopathological mechanisms are still poorly understood. Sarcopenia is characterized by low skeletal muscle mass and low muscle strength, which is occurring gradually with aging. Among other consequences, sarcopenia is causing a reduction of walking speed and loss of balance, affecting patient mobility and rising their risk of falls. Therefore, sarcopenia is associated with poor quality of life,[1] an increased risk of dependence, and mortality.[2] Prevalence of sarcopenia correspond to about 5–10% of persons over 65. It increases with age

and can be over 50% for persons above 80 years old.[3] Prevalence estimation depends on definition of sarcopenia considered.[4]

In the past, sarcopenia was not always considered as a disease but rather as a disorder associated to some chronic diseases affecting the elderly. Now and for few years, it has a disease status, although there is no specific treatment yet, only preventive measures such as diet supplements (proteins, collagen peptides, and vitamin D) and/or physical exercises.[1,5] In 2010, representatives (researchers and clinicians) of 13 European institutes, gathered in a Special Interest Groups (SIG) within the European Society of Clinical Nutrition

and Metabolism (ESPEN), published a consensus definition of sarcopenia.[6] Criteria were a low muscle mass, that is, a percentage of muscle mass ≥2 standard deviations below the mean measured in young adults of the same sex and ethnic background; and a low gait speed, for example, a walking speed below 0.8 m/s in the 4 m walking test or another geriatric assessment. The same year, a more detailed consensus definition of sarcopenia was published by the European Working Group on Sarcopenia in Older People (EWGSOP).[7] It describes different stages: presarcopenia, sarcopenia, and severe sarcopenia. Cut-off values by gender are reported for the main diagnostic techniques (*Table* 1). This definition was subsequently represented by the EWGSOP1 acronym and following a longitudinal study, was found in 2016 to be a good predictor of subsequent disability, hospitalization, and death.[8] Despite these results, the EWGSOP revised the 2010 definition to reach a new consensus definition in 2019, EWGSOP2,[9] in which the main diagnostic criterion is low muscle strength confirmed by imaging techniques. Physical performance that was used to being part of the diagnostic criteria in 2010 is now categorizing the severity of sarcopenia. To standardize measurement and diagnosis, muscle strength and physical performance are evaluated with several well-described physical and functional tests such as '6 min walking test (6MWT)', hand test, gait speed, or the Tinetti mobility test. This latter test was shown to be related to muscle mass and strength,[10] whereas imagery techniques are used to measure the muscle mass. Dual-energy X-ray absorptiometry (DXA) became the gold standard technique in clinical trials studying sarcopenia. The computed tomography and magnetic resonance imaging, more expensive, are other imaging techniques also used to assess the muscle mass.[11,12]

The muscle mass loss characterizing sarcopenia goes along with a modification of the muscle morphology. This includes a decrease in size and number of type II fibres (fast twitch fibres), while there is an increase in fat deposits as well as in connective tissues.[13,14] As written by Larsson *et al*.[15] and suggested previously by others,[16,17] these muscle transformations could be the consequence of the progressive loss of motoneurons, which was indeed previously observed.[18] Proteins involved in this muscle morphology changes, which would end up in the blood general circulation or in urine, could be potential biomarkers for sarcopenia.

Biomarkers are useful in several ways. They can be assayed either to perform a diagnosis, avoiding more costly examination, to predict disease progression or to indicate a subgroup of patients sensitive to a specific treatment, or to follow-up on a treatment efficiency. A combination of biomarkers could be required for diagnosis, and this combination could be different from early stage to late stage of the disease. For sarcopenia, circulating biomarkers would be particularly useful for early stage as there is currently no blood test for this disease.

Likewise, other skeletal muscle-specific proteins entering in the blood circulation following a tissue damage,[19] α-actin, which is released in serum after skeletal muscle damage,[20] could be a potential biomarker. Procollagen-3 NH2 terminal peptide (P3NP), is an example of biomarkers associated with muscle anabolism to hormonal response.[21,22] However, these muscle proteins might not be specific enough to differentiate sarcopenia from other muscle diseases. A 2015 review reports some circulating biomarkers for sarcopenia[23] including well-known proteins of the inflammatory response such as the C-reactive protein, interleukin 6, and tumour necrosis factor-α. Furthermore, in 2018, the creatinine (Cr)/cystatin C (CysC) ratio has been evaluated among a group of 677 rural elderly Japanese without severe renal deficiency and proposed as a predictor of sarcopenia.[24] Correlation studies were performed, but performance of these markers for diagnosis was not checked. The same year, a study, on 21 potential serum biomarkers,[25] reported that the levels of four proteins (interleukin 6, secreted protein acidic and rich in cysteine, macrophage migration inhibitory factor, and insulin-like growth factor 1), measured using immunoassays, were statistically significantly different between normal subjects and patients with moderate sarcopenia (according to EWGSOP1). Diagnostic performance evaluated with receiver operating characteristic (ROC) curves showed area under the curve (AUC) below 0.7. Combination of these four markers in a logistic regression model resulted in an AUC of 0.763. This result indicates that the performance of the method is good but may be not enough for diagnostic purposes. Therefore, new biomarkers have to be identified. In this work, we compared the proteome of female serum coming from sarcopenic and non-sarcopenic subjects to identify new discriminating biochemical markers.

Table 1 Sarcopenia diagnostic criteria published in 2010 (EWGSOP1)

Stage	Muscle mass		Muscle strength		Performance
Normal (S1)					
Presarcopenia (S2)	↓				
Sarcopenia (S3)	↓	AND	↓	OR	↓
Severe sarcopenia (S4)	↓	AND	↓	AND	↓
Main technique	DXA (muscle mass index)		Handgrip strength		Gait speed
Cut-off	Men ≤7.25 kg/m^2		Men <30 kg		<0.8 m/s
	Women ≤5.67 kg/m^2		Women <20 kg		

The subgroups considered for our study are S1: normal subjects with all criteria normal; S2: presarcopenic with lower muscle mass only; S3: sarcopenic patients; and S4: severe sarcopenic patients. DXA, dual-energy X-ray absorptiometry.

Materials and methods

Subjects and groups

A prospective and multi-centre study was set up to enrol subjects aged 65 or older, in two main groups (controls and sarcopenic). According to the definition of the European Working Group on Sarcopenia in Older People published in 2010 (EWGSOP1)[7] (*Table* 1), four subgroups were possible: control (S1), presarcopenia (S2), sarcopenia (S3), and severe sarcopenia (S3) (*Table* 2). However, as no fundamental clinical difference between groups was observed, subgroups were pooled to end up with 20 subjects considered as controls (S1 and S2) and 19 considered as sarcopenic patients (S3 and S4). Their muscle mass was evaluated by measuring the Skeletal Mass Index (SMI) corresponding to appendicular lean mass using the DXA technique, while their muscular function (strength and performance) was estimated using validated methods (handgrip strength test and 6 m walk test) and according to standardized procedures.[26–28] A blood sample and a urine sample were collected from enrolled patients. Clinical data were extracted from medical records, and results of routine laboratory tests were collected during study visits. Other data were collected via patient questionnaires: Nutritional Status Assessment, evaluation of protein consumption, lifestyle, history of falls and fractures, level of physical activity, assessment of cognitive function with the Mini-Mental State Examination (MMSE) score, and other functional evaluations. Data were encoded in paper Case Report Forms.

Cachexia was part of exclusion criteria [Mini Nutritional Assessment (MNA) score $< 17/30$; body mass index (BMI) < 17 kg/m^2]. Subjects whose muscle strength and/or walking speed were lower than normal but whose DXA showed a muscle mass above the sarcopenia threshold did not belong to either the control group or the sarcopenia group. Provided that they had not presented an exclusion criterion, the samples of these subjects were kept, but results were not reported in this article.

Proteomic analyses

Serum samples from 10 female control subjects and 10 female sarcopenic subjects were analysed by mass spectrometry (MS) as described in the succeeding text. Protein content of serum samples was determined using the 'RC DC™' kit from Bio-Rad (Brussels, Belgium); 1800 μg proteins were treated using the ProteoPrep immunodepletion plasma kit (Sigma, St. Louis, USA) according the protocol of the manufacturer. The final flow through was aliquoted and stored at −20°C until further analysis. One-third of the sample was used for protein assay (RC DC kit from Bio-Rad); 3 μg of each depleted sample were reduced, alkylated, and reduced again. Removal of impurities incompatible with MS analysis was performed using a 2D Clean-Up Kit (GE) according the manufacturer's recommendations. Protein pellets were solubilized in bicarbonate ammonium of 50 mM. Protein samples were then digested with trypsin (16 h at 37°C) with trypsin/total proteins ratio (W/W) of 1/10 and 3 h at 37°C with a ratio of 1/20 in 80% acetonitrile. The reaction was stopped by addition of trifluoroacetic acid. The samples were dried in a speed vacuum. Samples were dissolved in water 0.1% formic acid and then aliquoted; 0.7 g of protein digest was purified using a Zip-Tip C18 high capacity according to the manufacturer's recommendations. Samples were dried in a speed vacuum and solubilized in 100 mM ammonium formate (pH 10) at 0.067 μg/μL; 9 μL of sample was spiked with an internal control corresponding to a MassPREP™ Protein Digestion Standard Mixtures (MPDSMIX) from Waters (Milford, USA), containing four digested proteins, at a quantity of 150 fmol of Alcohol DeHydrogenase (ADH) digest per injection. More precisely, the MassPREP™ Protein Digestion Standard Mixtures (MPDSMIX) from Waters consists of two standard mixtures (MPDS Mix 1 and MPDS Mix 2). Mix 1 and Mix 2 contain the same standard proteins but in different molecular ratio. Mix 2 contains twice the amount of Mix 1 for enolase, which allow to have a positive control for difference in quantities. Mix 1 was spiked into control samples and Mix 2 in sarcopenic sample. The liquid chromatography method was a 2D liquid chromatography method with three steps of 180 min at high pH with increasing percentage of acetonitrile. The eluted peptides were loaded on the low pH column [5 min gradient from 99% of A (0.1% formic acid) to 93% of B (acetonitrile) followed by a 135 min gradient from 93% of A to 65% of B]. The mass spectrometer method is a TopN-MSMS method where *N* was set to 12, meaning that the spectrometer acquired one full MS spectrum, selected the 12 most intense peaks in this spectrum (singly charged precursors excluded), and made a full MS2 spectrum of each of these 12 compounds. The parameters for MS spectrum acquisition were mass range from 400 to 1750 m/z, resolution of 70 000, AGC target of 1e6, or maximum injection time of 200 ms. The parameters for MS2 spectrum acquisition were isolation window of 1.6 m/z, collision energy (NVE) of 25, resolution of 17 000, AGC target of 1e5, or maximum injection time of 50 ms. The database searches and quantitative analyses were performed with MaxQuant software version 1.5.2.8.[29] Protein identifications were considered significant if proteins were identified with at least two peptides per protein taking into account only a false discovery rate < 0.01.

Table 2 Subgroup description

Subgroup	Median age	Female count	Male count
S1	77.5	7	1
S2	77	10	2
S3	83	7	2
S4	77.5	9	1

Age and sex distribution among the subgroups.

Cathepsin D measurement in serum

Cathepsin D in serum was assayed by sandwich enzyme-linked immunosorbent assay (ELISA) with the Human Cathepsin D ELISA Kit from Abcam, Cambridge, USA (# ab213470).

Statistical analysis

Perseus software version 1.6.2.3 was used to do a statistical comparative analysis of the proteomic data of the two groups of samples. Student's t-tests are performed with 250 permutations to adjust the P value ($\alpha = 0.05$). Visualization of the t-tests results was performed with a volcano plot.[30]

For the other data analyses and graphical representations, R software version 3.5.1 (2018-07-02) was used. Statistical significance was also set to a P value < 0.05 ($\alpha = 0.05$). The non-parametric Mann–Whitney–Wilcoxon test was applied when comparing two groups. The R Package 'pROC' was used to plot the ROC curve and calculate the AUC. A multivariate logistic regression was performed to establish a predictive model with the logit function. The Akaike information criterion (AIC) and the analysis of deviance were used to compare nested models.

Results

Subject and subgroup characteristics

Sixty-two subjects were enrolled. Among them, 4 showed screening failures and 19 had a decrease of muscle strength without a decrease of muscle mass, which was not a sarcopenia condition as defined by the EWGSOP1 (*Table* 1). Therefore, data from the remaining 39 patients were studied.

EWGSOP1 definition implies that sarcopenic patients (S3, *Table* 1) had a low handgrip strength ($<$30 kg for men and $<$20 kg for women) or a decreased walking speed ($<$0.8 m/s) combined to a low skeletal mass index (SMI; $\leq$7.25 kg/m^2 for men and $\leq$5.67 kg/m^2 for women). Regarding severe sarcopenic patients (S4, *Table* 1), measurements of these three criteria were lower than normal.

Under these conditions, our sarcopenic S3 subgroup included nine patients with handgrip strength or walking speed lower than the EWGSOP1 cut-off (*Table* 2). In addition, there were 10 severe sarcopenic patients with handgrip strength and walking speed below the EWGSOP1 cut-off (S4 subgroup, *Table* 2). All 19 patients (S3 plus S4) were also having low SMI. Moreover, eight subjects were normal for all characteristics considered (S1 subgroup, *Table* 2). The remaining 12 patients with low muscle mass but no decrease in handgrip strength and gait speed were classified as part of the S2 subgroup. All subgroups were composed mainly of women (*Table* 2).

For all patients, some other clinical measurements were collected, such as the MNA, the MMSE, the Tinetti mobility, and balance score or the Dijon Physical Activity Score. Subgroups were compared statistically on clinical criteria (*Table* 3) with $\alpha = 0.05$. Based on these results, it was not possible to conclude about Mental State association with sarcopenia. The most different subgroups seemed to be S1 and S4. Indeed, it was for the comparison S1 vs. S4 that most criteria were statistically different. Not as many clinical criteria were shown to be different between S3 and S4 subgroups as only the performance of the 6MWT was statistically different. S1 and S2 differ by definition on the SMI and consequently on the BMI (*Table* 3). This indicated that S1 and S2 subgroups were clinically close together. Likewise, S3 and S4 subgroups were also clinically close. Because numbers in each subgroup were small and because we could observe clinically related subgroups, it was decided to pool S3 patients with S4 patients to constitute our sarcopenic group of 19 patients for further studies. These patients were compared with a control group

Table 3 Subgroup comparison

	P value					
	S1 vs. S2	S1 vs. S3	S1 vs. S4	S2 vs. S33	S2 vs. S4	S3 vs. S4
BMI	**0.0171**	**0.0171**	**0.0171**	0.5329	0.8898	0.9048
SMI	**0.0065**	**0.0187**	**0.0033**	0.8078	0.5474	0.3642
Left hand strength test	0.7855	**0.0271**	**0.0022**	**0.0204**	**0.0013**	0.3679
Right hand strength test	0.7063	**0.0241**	**0.0058**	**0.0058**	**0.0058**	0.7746
6 min walking test	0.5890	0.1377	**0.0009**	0.5065	**0.0005**	**0.0009**
Tinetti score	0.5722	0.4816	**0.0314**	0.4816	**0.0062**	0.1849
Dijon score	0.3142	0.6290	0.3142	0.1900	0.0543	0.3142
MNA score	0.2390	0.1059	**0.0048**	0.2390	**0.0119**	0.2247
MMSE score	0.5183	0.5183	0.5183	1.0000	0.5183	0.5183

Results of index measurements, tests, and scores for the different factors considered are compared between subgroups. The P values are calculated by pairwise comparisons between group levels with corrections for multiple testing using R software and pairwise Wilcoxon rank sum tests. BMI, body mass index; MMSE, Mini-Mental State Examination; MNA, Mini Nutritional Assessment; SMI, skeletal muscle index. Bold corresponds significant P value.

of 20 control subjects corresponding to S1 subjects pooled with S2 ones.

Comparison of clinical characteristics of the control group vs. the sarcopenic group

Subject characteristics for some clinical variables were summarized in *Figure* 1 by box plots. The variables shown correspond to two physical characteristics of the individuals (age and BMI) and seven characteristics more specific to sarcopenia (skeletal mass index, hand strength, mobility or physical performances, and nutrition). In most cases, data were not normally distributed, so the control group was compared with the sarcopenic group using the Mann–Whitney–Wilcoxon non-parametric test (*Figure* 1). It could be observed that age distribution was not the same between control and sarcopenic patients. Regarding strength (right or left hand), mobility (Tinetti test), physical performance (Dijon score), or gait speed (6MWT), they were significantly lower in the sarcopenic group compared with the control group (all the *P* values being below 0.05). This was also the case of the BMI, the skeletal mass index (SMI), and the MNA score. However, no significant difference was observed for the MMSE score between the control and sarcopenic groups (not shown). To summarize the overall variations, as age increased, variables related to strength, mobility, and physical performance, characterizing sarcopenia status, decreased significantly as expected. In addition, sarcopenic patients have a lower BMI as well as a lower skeletal mass index. The control group had a satisfactory nutritional status while some sarcopenic patients showed a risk of malnutrition without being undernourished (too low BMI was an exclusion criterion).

Proteomic analysis

Serum proteins were identified and quantified by MaxQuant analysis of the MS data. Then a comparison of the levels of serum proteins between the two groups of samples (sarcopenic vs. control) was performed with Perseus software by applying *t*-tests with corrections for multiple hypotheses. Statistical significance in protein levels between sarcopenic and controls was plotted in a volcano plot (*Figure* 2) to visualize results. Among the six significant proteins, three had unidentified protein and gene names (Perseus). Nevertheless, the Fasta header allows the following identifications: P02769 protein ID was from bovine serum albumin, which is part of the MassPREP Digestion Standard Mixtures; CON__Q3SZR3 protein ID was identified as alpha-1-acid glycoprotein precursor from Taurus, a potential contaminant; and P00489 protein ID corresponds to RABIT Glycogen phosphorylase also part of the Standard Mixtures. Then the three other significant proteins identified with their protein names

were Enolase 1 from *Saccharomyces cerevisiae*, which is an MS positive control; then the two sample-specific proteins with significant but modest different levels between the two groups were fructose-bisphosphate aldolase A and CTSD (cathepsin D). Main proteomic information regarding these two likely biomarkers was summarized in *Table* 4. For information in the MassPREP Digestion Standard Mixtures, there is a fourth protein (ADH from *S. cerevisiae*), which level is not different between the two mixes and was identified as a protein with no significant difference in its level between sarcopenic and control samples.

Cathepsin D levels measured by immunoassay in serum

The level of cathepsin D (ng/mL) has been assayed by ELISA in all the serum samples. Like for the other variables, a box-plot representation was performed to compare the data distribution between the non-sarcopenic and sarcopenic groups (*Figure* 3). The median level of cathepsin D was higher in the sarcopenic group (364.0 ng/mL) than in the control group (314.7 ng/mL). The calculated *P* value being 0.038.

To better understand the link between cathepsin D and sarcopenia, correlation coefficients between cathepsin D and other variables (6MWT, Tinetti test, strength hand test, MNA score, MMSE score, SMI, and Dijon score) specific of sarcopenia were calculated. A *t*-test on the slope was computed to test for slopes different from zero. The highest correlation and also the only significant one was with the result of the 6MWT (Pearson coefficient = −0.385; *P* = 0.0155) as shown in *Figure* 4A. This weak inverse correlation implies that as gait speed decreases, the level of cathepsin D increases. Another way to represent this correlation was with a box plot (*Figure* 4B). One can observe that the highest CTSD levels were associated with the lowest gait speed patients (*P* = 0.038).

To evaluate the performance of cathepsin D to discriminate between sarcopenic and non-sarcopenic patients, a ROC curve was plotted based on the cathepsin D results. Specificity and sensitivity were displayed for the optimal threshold calculated by the software (*Figure* 5A). At the optimal threshold (338 ng/mL), the sensitivity was 63.2% and the specificity was 75%. The AUC was about 0.7. To improve the method, a model including cathepsin D and other variables was investigated. Logistic regression fit well diagnostic purposes, and a logit model was chosen. To start with, a logit model with only cathepsin D was looked at (not shown). In that case, the coefficient for cathepsin D was not significant (*P* = 0.0571), and the AIC for the model was 52.505. Another more complex model was evaluated, which included cathepsin D, age, BMI, and age by BMI interactions (Table 5). The AIC of this more complex model is 40.39 and was lower than a model with cathepsin D only. Therefore, this latter model

Figure 1 Sample data characteristics. Data distribution of nine numeric variables by group (control vs. sarcopenic) following EWGSOP1 diagnostic criteria. Significant difference between the two data samples was evaluated using the non-parametric Mann–Whitney–Wilcoxon test with R software. The *P* value is displayed above each graph. The control group is composed of 20 subjects (8 normal + 12 presarcopenic), and the sarcopenic group is composed of 19 subjects (9 sarcopenic + 10 severe sarcopenic).
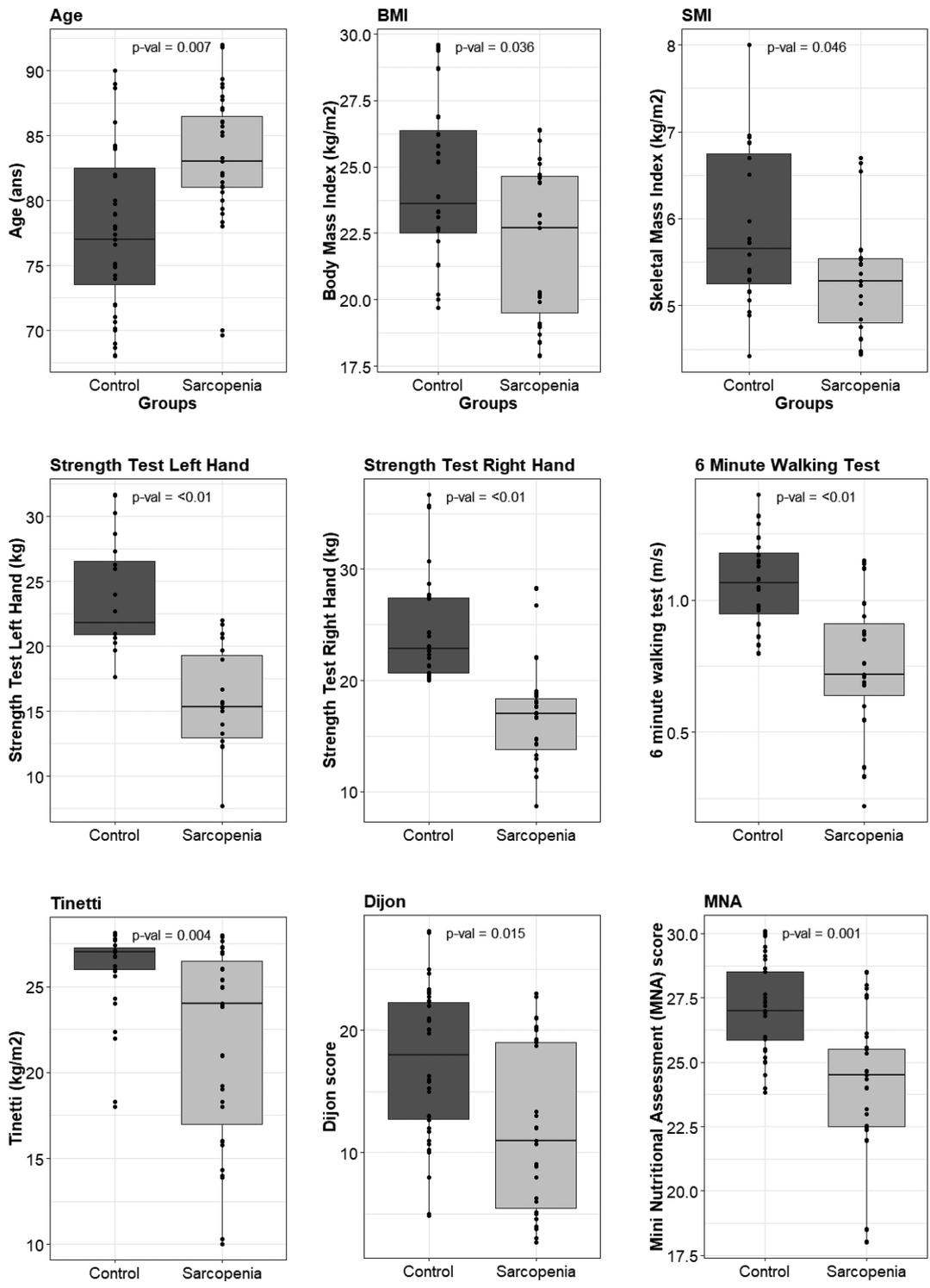
Figure 2 Proteomic analysis of serum proteins showing significant differences in protein levels between sarcopenic patients and control patients. Results were displayed in a volcano plot: significance ($-\log(P)$) vs. protein quantity difference. To illustrate the table of results, the significant proteins with unidentified protein or gene names are in black. Fasta sequence information indicated that P02769 protein ID was bovine serum albumin and P00489 protein ID was RABIT Glycogen phosphorylase. Both were peptides from the MassPREP Digestion Standard Mixtures. The third protein ID, CON__Q3SZR3, was identified as alpha-1-acid glycoprotein precursor from Taurus, a potential contaminant. Identified gene names are indicated in red: ENO1 is the yeast Enolase 1 (another positive control), CTSD is cathepsin D, and ALDOA is fructose-bisphosphate aldolase A. These two latest proteins are proteins with significant higher levels in sarcopenic samples.
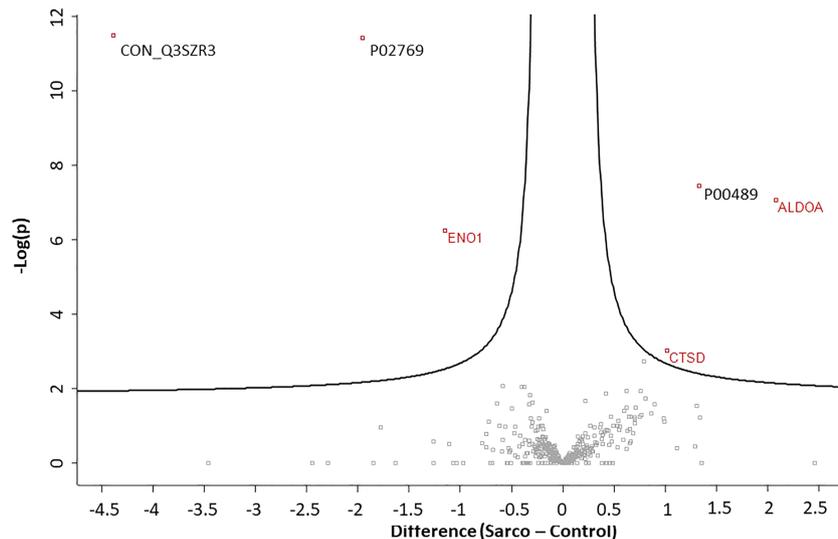
seemed to be better. The odds ratio for cathepsin D would be just above 1 (1.01). The values predicted by this model resulted in a ROC curve with an AUC of 0.908 (*Figure* 5B), which was better than the one coming from cathepsin D biomarker on its own. It also indicated a good ability of the model to discriminate between sarcopenic and non-sarcopenic subjects. At the optimal threshold (0.559), the specificity was 85% and the sensitivity was 89.5%. The importance of cathepsin D contribution to the chosen model was checked by removing CTSD from this model. The new model was therefore calculated with age, BMI, and age by BMI interactions (data not shown). The analysis of variance comparing the two models (with or without CTSD) showed a statistically significant difference (P = 0.009). Moreover, the AIC of the model without CTSD increased to 45.14 indicating a relatively worse model. As a result, the AUC of the ROC curve was smaller (0.858), the sensitivity being 73.7%.

Discussion

The main aim of our study was to uncover new potential biomarkers of sarcopenia out of proteomic analyses. To start with, patients were enrolled in a case–control study and allocated to the control group or to the sarcopenic group, depending on their characteristics (clinical evaluation of several tests) and accordingly to the main sarcopenia definition published in 2010 by a group of experts in the field. New criteria were

proposed in 2019, which could not be applied to our population sample, even in a post hoc analysis, as recruitment started in 2015. Therefore, we cannot anticipate that similar results would be obtained by using the new 2019 criteria. Regardless of sarcopenia clinical definition, included sarcopenic patients were different to control subjects. Indeed, sarcopenic patients were weaker, showing lower physical performance. This was associated to a lower BMI, lower skeletal mass index, and a higher risk of being undernourished.

A proteomic method with MS technology, which is a high-throughput way, was used for screening for new biomarkers in serum samples. The level of only two proteins was found significantly higher in sarcopenic patients than in the control group: fructose-bisphosphate aldolase A and modestly cathepsin D. However, our sample size might not be large enough to identify more proteins. Moreover, despite that the proteomic analysis was performed on only female samples, the biomarkers uncovered are not necessarily specific to women. Mainly female subjects were enrolled. There were not enough men to study by stratification the effect of gender on biomarkers either in the proteomic analysis or in the immunoassay analysis. Sex was not a variable part of the ones we studied. The enrolled male subjects were kept in the immunoassay analysis.

Regarding the identified proteins, fructose-bisphosphate aldolase A is a glycolytic enzyme. It is mainly found in heart, skeletal muscles, and erythrocytes and can be associated with skeletal muscle disorders.[31] Aldolase A is released in the

Table 4 Proteomic result summary: proteins with significant level differences between sarcopenic and control samples

−Log(protein P value)	−Log(tthreshold P value)	Statistical significance	Level difference	Majority protein IDs	One protein name	Gene name	Peptides	Unique peptides	Sequence coverage (%)	Unique sequence coverage (%)	Few keywords
7.066	2.144	+	2.080	P04075	Fructose-bisphosphate aldolase A	ALDOA	12	10	45.3	37.6	Actin cytoskeleton organization; glycolysis
3.029	2.675	+	1.015	P07339	Cathepsin D	CTSD	8	8	29.4	29.4	Autophagy; aspartic-type endopeptidase activity

Control proteins such as yeast enolase and contaminant proteins like keratins were removed. Remaining proteins, which levels were significantly different between the sarcopenic and control groups, are indicated in this table. Numbers of sequenced peptides and the corresponding protein sequence coverage are indicated. Keywords were selected from UniProt biological process, molecular function, and cellular component. This table summarizes the results displayed on the volcano plot. The statistical significance at $\alpha = 0.05$ is indicated by '+'; it corresponds to a protein P value (uncorrected) being below the adjusted threshold P value or on the volcano plot [−log(protein P value)] being above the black line representing the adjusted threshold [−log(threshold P value)]. CTSD, cathepsin D.

Figure 3 Cathepsin D levels (ng/mL) distribution per group (control vs. sarcopenic) according to EWGSOP1 criteria. Significant difference between the two data samples was evaluated using the non-parametric Mann–Whitney–Wilcoxon test with R software. The *P* value is displayed above the graph. The control group is composed of 20 subjects (8 normal + 12 presarcopenic), and the sarcopenic group is composed of 19 subjects (9 sarcopenic + 10 severe sarcopenic).
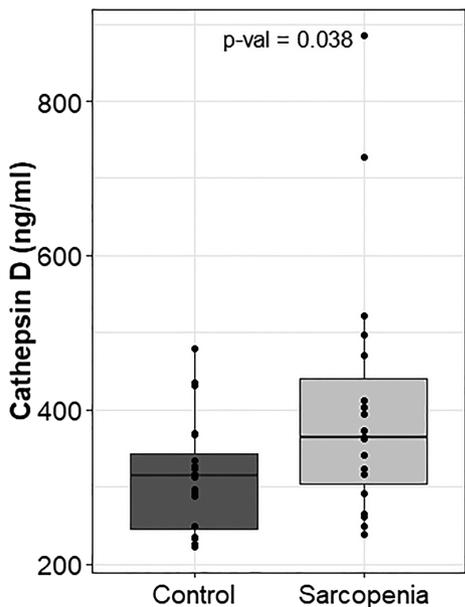
bloodstream upon muscle damage. Thus, high level of Aldolase A in serum can indicate muscle disorders such as Duchenne muscular dystrophy. Aldolase A is also a potential

biomarker for osteoarthritis.[32] Cathepsin D is a lysosomal enzyme. High levels of this enzyme can be associated with a state of growth or with enhanced autophagy. Cathepsins in general are involved in lysosomal death pathways, autophagy, and aging mostly in neurodegenerative disorders.[33] Cathepsin B, for example, is involved in myoblast differentiation,[34] and cathepsin D is present in atrophied muscles and apoptotic myofibers.[35] In summary, from what is already known, these two proteins could be involved in muscle energy metabolism and apoptosis of myoblasts.

Our biological data analyses have confirmed that cathepsin D is a promising biomarker for sarcopenia when considering EWGSOP1 criteria, which is in part likely due to some inverse correlation with gait speed. Although to get a good diagnostic performance, this biomarker needs to be associated with age and BMI. Age is an important variable as it is one of the main characteristics of sarcopenia although a possible confounding variable. Like age, it is also probable that BMI brings complementary information to cathepsin D on sarcopenia making the prediction with the logit model stronger.

In conclusion, we have set up a case–control study for sarcopenia according to the definition published in 2010 by the EWGSOP group of experts. We have identified two potential biomarkers by proteomic analysis. One of them is cathepsin D, whose levels in serum were slightly but significantly higher in sarcopenic patients than in control subjects, confirming its biomarker status. In addition, cathepsin D levels were showing some inverse correlation with gait speed. Finally, we have established a good predictive model for sarcopenia when cathepsin D is combined to age and BMI.

Figure 4 Inverse correlation between cathepsin D levels and speed gait. (*A*) Plot representation of the inverse correlation between cathepsin D levels and the results of the 6 min walking test (6MWT). Data points from sarcopenic patients are represented by 'S' letter, and data points from control patients are represented by 'C' letter. The Pearson correlation coefficient is −0.385, and the *t*-test on the slope indicates that it is significantly different from zero (*P* value = 0.0155). (*B*) Discrimination between normal and low gait speed patients with cathepsin D. Significant difference between the two data samples was evaluated using the non-parametric Mann–Whitney–Wilcoxon test with R software. The *P* value (0.038) is displayed above the graph.
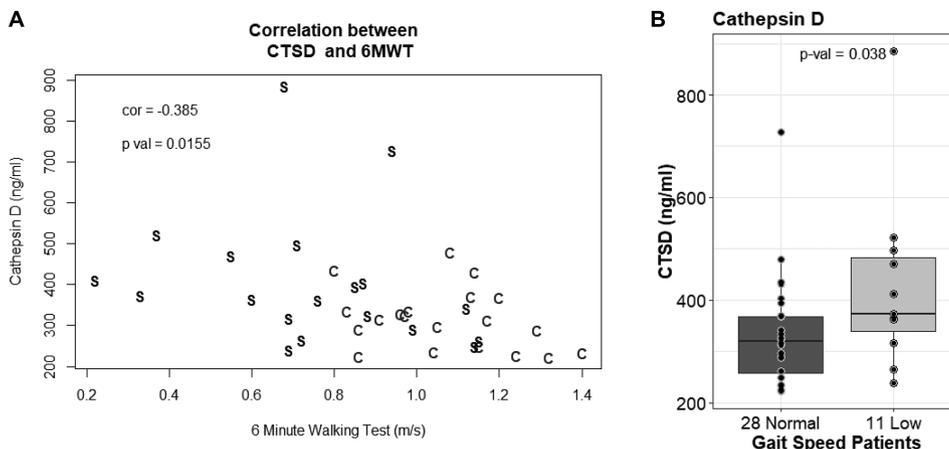
Figure 5 Performance evaluation, with a receiver operating characteristic curve (ROC curve) of cathepsin D as a biomarker for sarcopenia. (*A*) Cathepsin D on its own, showing 75% specificity, 63.2% sensitivity for a threshold of 338.2 ng/mL of cathepsin D, corresponding to an area under the curve (AUC) of 0.696. (*B*) Cathepsin D in combination with other variables [age, body mass index (BMI), and age–BMI interaction], showing 85% specificity, 89.5% sensitivity for a threshold of 0.559 (predicted model value), corresponding to an AUC of 0.908.
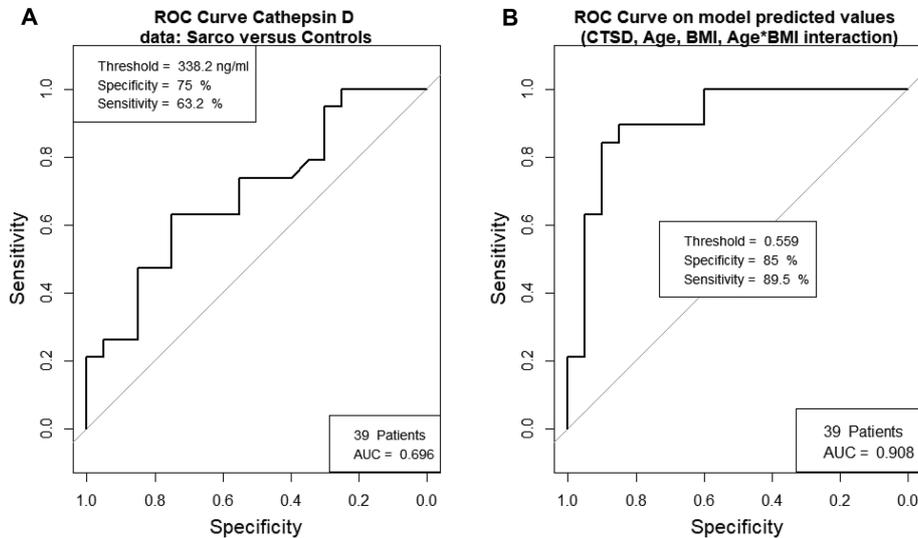
Table 5 Logistic regression model summary

Variables	Coefficient estimate	Standard error	Probability
Intercept	−113.9	79.13	0.1499
CTSD	0.01035	0.005012	0.0388*
Age	1.474	0.9698	0.1287
BMI	3.764	3.191	0.2382
Age * BMI	−0.0513	0.03957	0.1948

Akaike information criterion = 40.39. BMI, body mass index; CTSD, cathepsin D.
*Significant at $\alpha = 0.05$.

Acknowledgements

The authors of this manuscript certify that they comply with the ethical guidelines for authorship and publishing in the

Conflict of interest

C.L., B.C., A.L., and B.C. were employees of ARTIALIS SA. C.B. received consulting fees from ARTIALIS SA. Y.H. was the founder and chairman of ARTIALIS SA.

References

1. Santilli V, Bernetti A, Mangone M, Paoloni M. Clinical definition of sarcopenia. *Clin Cases Miner Bone Metab* 2014;**11**:177–180.
2. Sipers WMWH, de Blois W, Schols JMGA, van Loon LJC, Verdijk LB, Sarcopenia IS. Related to mortality in the acutely hospitalized geriatric patient. *J Nutr Health Aging* 2019;**23**:128–137.
3. Berger MJ, Doherty TJ. Sarcopenia: prevalence, mechanisms, and functional consequences. *Interdiscip Top Gerontol* 2010;**37**:94–114.
4. Dam T-T, Peters KW, Fragala M, Cawthon PM, Harris TB, McLean R, et al. An evidence-based comparison of operational criteria for the presence of sarcopenia. *J Gerontol A Biol Sci Med Sci* 2014;**69**:584–590.
5. Zdzieblik D, Oesser S, Baumstark MW, Gollhofer A, König D. Collagen peptide supplementation in combination with resistance training improves body composition and increases muscle strength in elderly sarcopenic men: a randomised controlled trial. *Br J Nutr* 2015;**114**:1237–1245.
6. Consensus definition of sarcopenia, cachexia and pre-cachexia: joint document elaborated by Special Interest Groups (SIG) cachexia-anorexia in chr… - PubMed - NCBI. https://www.ncbi.nlm.nih.gov/pubmed/20060626. Accessed 30 April 2019.
7. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, et al. Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on Sarcopenia in Older People. *Age Ageing* 2010;**39**:412–423.
8. Bianchi L, Ferrucci L, Cherubini A, Maggio M, Bandinelli S, Savino E, et al. The predictive value of the EWGSOP definition of sarcopenia: results from the InCHIANTI

study. *J Gerontol A Biol Sci Med Sci* 2016;**71**:259–264.

9. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing* 2019;**48**:16–31.

10. Curcio F, Basile C, Liguori I, Della-Morte D, Gargiulo G, Galizia G, et al. Tinetti mobility test is related to muscle mass and strength in non-institutionalized elderly people. *Age (Dordr)* 2016;**38**:525–533.

11. Vellas B, Pahor M, Manini T, Rooks D, Guralnik JM, Morley J, et al. Designing pharmaceutical trials for sarcopenia in frail older adults: EU/US Task Force recommendations. *J Nutr Health Aging* 2013;**17**:612–618.

12. Dardevet D, Savary-Auzeloux I, Remond D, Mosoni L, Marzetti E, Buford TW, et al. Commentaries on Viewpoint: muscle atrophy is not always sarcopenia. *J Appl Physiol* 2012;**113**:680–684.

13. Miljkovic N, Lim J-Y, Miljkovic I, Frontera WR. Aging of skeletal muscle fibers. *Ann Rehabil Med* 2015;**39**:155–162.

14. Narici MV, Maffulli N. Sarcopenia: characteristics, mechanisms and functional significance. *Br Med Bull* 2010;**95**:139–159.

15. Larsson L, Degens H, Li M, Salviati L, Lee YI, Thompson W, et al. Sarcopenia: aging-related loss of muscle mass and function. *Physiol Rev* 2019;**99**:427–511.

16. Kwan P. Sarcopenia, a neurogenic syndrome? *J Aging Res* 2013;**2013**:791679.

17. Kwon YN, Yoon SS. Sarcopenia: neurological point of view. *J Bone Metab* 2017;**24**:83–89.

18. Drey M, Krieger B, Sieber CC, Bauer JM, Hettwer S, Bertsch T, et al. Motoneuron loss is associated with sarcopenia. *J Am Med Dir Assoc* 2014;**15**:435–439.

19. Nedergaard A, Karsdal MA, Sun S, Henriksen K. Serological muscle loss biomarkers: an overview of current concepts and future possibilities. *J Cachexia Sarcopenia Muscle* 2013;**4**:1–17.

20. Martinez-Amat A, Boulaiz H, Prados J, Marchal J, Padial P, Caba O, et al. Release of α-actin into serum after skeletal muscle damage. *Br J Sports Med* 2005;**39**:830–834.

21. Berry SD, Ramachandran VS, Cawthon PM, Gona P, McLean RR, Cupples LA, et al. Procollagen type III N-terminal peptide (P3NP) and lean mass: a cross-sectional study. *J Fraility Aging* 2013;**2**:129–134.

22. Bhasin S, He EJ, Kawakubo M, Schroeder ET, Yarasheski K, Opiteck GJ, et al. N-terminal propeptide of type III procollagen as a biomarker of anabolic response to recombinant human GH and testosterone. *J Clin Endocrinol Metab* 2009;**94**:4224–4233.

23. Calvani R, Marini F, Cesari M, Tosato M, Anker SD, von Haehling S, et al. Biomarkers for physical frailty and sarcopenia: state of the science and future developments. *J Cachexia Sarcopenia Muscle* 2015;**6**:278–286.

24. Kusunoki H, Tsuji S, Wada Y, Fukai M, Nagai K, Itoh M, et al. Relationship between sarcopenia and the serum creatinine/cystatin C ratio in Japanese rural community-dwelling older adults. *JCSM Clin Rep* 2018;**3**.

25. Kwak JY, Hwang H, Kim S-K, Choi JY, Lee S-M, Bang H, et al. Prediction of sarcopenia using a combination of multiple serum biomarkers. *Sci Rep* 2018;**8**:1–7.

26. Sousa-Santos AR, Amaral TF. Differences in handgrip strength protocols to identify sarcopenia and frailty—a systematic review. *BMC Geriatr* 2017;**17**:238.

27. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002;**166**:111–117.

28. Pahor M. Mobility and functional outcomes for sarcopenia trials. *J Frailty Aging* 2015;**4**:123–124.

29. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008;**26**:1367–1372.

30. Tyanova S, Cox J. Perseus: a bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods Mol Biol* 2018;**1711**:133–148.

31. Tarnopolsky MA. Myopathies related to glycogen metabolism disorders. *Neurotherapeutics* 2018;**15**:915–927.

32. Ganguly A. Levels of C-reactive protein, creatine kinase-muscle and aldolase A are suitable biomarkers to detect the risk factors for osteoarthritic disorders: a novel diagnostic protocol. *Caspian J Intern Med* 2019;**10**:25–35.

33. Stoka V, Turk V, Turk B. Lysosomal cathepsins and their regulation in aging and neurodegeneration. *Ageing Res Rev* 2016;**32**:22–37.

34. Jane DT, DaSilva L, Koblinski J, Horwitz M, Sloane BF, Dufresne MJ. Evidence for the involvement of cathepsin B in skeletal myoblast differentiation. *J Cell Biochem* 2002;**84**:520–531.

35. Nagano K. Alteration of cathepsin-D expression in atrophied muscles and apoptotic myofibers by hindlimb unloading in a low-temperature environment. *J Phys Ther Sci* 2015;**27**:3585–3591.