Mechanisms of Ageing and Development 182 (2019) 111129

Contents lists available at ScienceDirect

Mechanisms of Ageing and Development

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

Prognostic networks for unraveling the biological mechanisms of Sarcopenia

Ana Cernea[a], Juan Luis Fernández-Martínez[a], Enrique Juan de Andrés-Galiana[a], Zulima Fernández-Muñiz[a], Juan Carlos Bermejo-Millo[b], Laura González-Blanco[b], Juan José Solano[c], Pedro Abizanda[d,e], Ana Coto-Montes[b], Beatriz Caballero[b,*]

[a] Group of Inverse Problems, Optimization and Machine Learning, Department of Mathematics, University of Oviedo, C/ Federico García Lorca, 8, 33007, Oviedo, Spain
[b] Instituto de Investigación Sanitaria del Principado de Asturias (ISPA) and Department of Morphology and Cellular Biology, Faculty of Medicine, University of Oviedo, C/ Julián Claveria 6, 33006, Oviedo, Spain
[c] Instituto de Investigación Sanitaria del Principado de Asturias (ISPA) and Geriatric Service, Monte Naranco Hospital, Av. Dolores Fernández Vega 107, 33012, Oviedo, Asturias, Spain
[d] Geriatric Service, Complejo Hospitalario Universitario de Albacete, C/Hnos. Falcó, 37, 02008, Albacete, Spain
[e] CIBERFES (CB16/10/00408), Instituto de Salud Carlos III, Madrid, Spain

ARTICLE INFO

Keywords:
Sarcopenia
Biological mechanisms
Machine learning
Prognostic networks

ABSTRACT

Sarcopenia is an age-related multifactorial process that involved several biological mechanisms, whose specific contribution and interplay is still unknown. The present study proposes prognostic networks based on machine learning approaches to unravel the interplay among those biological mechanisms mainly involved in the development of Sarcopenia. After analyzing 114 biological and clinical variables in adults older than 70 years, and using all the biological prognostic networks detected by machine learning with accuracy higher than 82%, we designed a consensus classifier based on majority vote that improve the predictive accuracy of Sarcopenia up to 91%. Additionally, we applied logistic regression analysis to propose the interplay among the most discriminative biological variables of Sarcopenia: anthropometry, body composition, functional performance of lower limbs, systemic oxidative stress, presence of depression and medication for the digestive system based on proton-pump inhibitors. Our data also demonstrate that besides a loss of muscle mass, impairments on functional performance of lower limbs are more relevant for develop Sarcopenia than those affecting the muscle strength.

1. Introduction

Sarcopenia is a burdensome geriatric syndrome characterized by a loss of muscle mass and strength as well as impaired functional performance (Cruz-Jentoft et al., 2010a, b). Nowadays, the prevalence of Sarcopenia in adults over 70 years of age is higher in women (35.3%) than men (13.1%) (Coto Montes et al., 2017). Besides female gender, other risk factors of Sarcopenia have been recently described including advanced age, low physical performance, malnutrition, depression, hypertension and pre-frailty (Coto Montes et al., 2017). Several cellular and molecular deregulations can be associated with Sarcopenia including an increased oxidative stress, a pro-inflammatory state, alterations in protein synthesis and/or degradation as well as fat infiltration in the muscle that increase its sensibility to apoptosis (Bellanti et al., 2018; Brzeszczynska et al., 2018; Coto Montes et al., 2017; Cruz-Jentoft et al., 2010a, b; Marzetti et al., 2012; Tong et al., 2011; van Dijk et al., 2018v). Given its multifactorial character, involving several different biological mechanisms, whose specific contribution and interplay is still unknown, it is really complicated a precise diagnosis of Sarcopenia as well as the development of effective preventive and therapeutic strategies. In the present study, we applied machine learning techniques and logistic regression analysis in a set of data obtained from a retrospective study to unravel interplay among those biological mechanisms mainly involved in the development of Sarcopenia.

2. Material and methods

2.1. Setting and population sample

The participants represent an urban population sample (N = 200) of older (≥70 years) people belonging to the FRADEA (Frailty and Dependence in Albacete) cohort, which has been extensively described in previous publication set (Abizanda et al., 2011; Caballero et al., 2014; Coto Montes et al., 2017). All the data used in this study were obtained through a valid signed informed consent form from each participant or participant's guardian after been informed about the

* Corresponding author.
E-mail address: caballerobeatriz@uniovi.es (B. Caballero).

https://doi.org/10.1016/j.mad.2019.111129
Received 14 June 2019; Received in revised form 26 July 2019; Accepted 18 August 2019
Available online 21 August 2019

purposes and goals of the study. This research complied with the Declaration of Helsinki concerning research with humans. The description of the variables (N = 114) studied in the population is listed in supplementary Table 1. Sarcopenia was diagnosed by applying the algorithm proposed by the European Working Group on Sarcopenia in Older People (EWGSOP) (Cruz-Jentoft et al., 2010a,b). By this way, and according with the cut-off points applied in our previous studies (Coto Montes et al., 2017), presence of Sarcopenia was detected with a low skeletal muscle index (SMI, $\leq$ 8.31 kg/m^2 for men and $\leq$ 6.68 kg/m^2 for women) together with a low gait speed ($<$ 0.8 m/s) and/or poor grip strength on the dominant hand ($\leq$ 30 kg for men and $\leq$ 20 kg for women). The methodology of the study variables has been extensively described in our previous works (Abizanda et al., 2011; Caballero et al., 2014; Coto Montes et al., 2017).

2.2. Machine learning methodologies

Machine learning analyses were applied to implement a consensus classifier based on majority voting over a set of high-discriminative variables networks in order to optimally separate the population between healthy (G1) and sarcopenic (G2) participants, exploring its uncertainty space to make decisions, as described previously (Cernea et al., 2018; De Andres-Galiana et al., 2015, 2016a,b). This methodology is summarized in Fig. 1. Briefly, this mathematical methodology is based on the estimation of the Fisher´s ratio (FR) to detect high-discriminative variables (headers) and low-discriminatory variables (helpers) in the samples classification (G1 *versus* G2). Then, we calculated the shortest list of prognostic variables with the highest predictive accuracy, based on Leave-One-Out-Cross-Validation (LOOCV), as described previously (Cernea et al., 2018). A random sampler will find the high-discriminatory prognostic networks with the highest predictive accuracy and stability, by considering both headers and helpers variables, given that high-discriminatory variables (headers) serve to span the main features of the classification, whilst variables with lowest discriminatory ratios (helpers) account for the details in the discrimination. Next, a consensus classifier based on majority vote was implemented to predict the class (G1 or G2) of new incoming samples. This algorithm used all the equivalent networks that have been sampled with a predictive LOOCV accuracy higher than 82%. Finally, a frequency classification analysis showed those most frequent variables in the samples classification.

Notably, this methodology was applied considering the following sceneries:

(A) All the prognostic variables of the study, except the Skeletal-Muscle Index (SMI) (n = 113).
(B) All the prognostic variables, excluding the SMI and those measures obtained by bioelectrical impedance analysis (BIA, n = 90).
(C) All the prognostic variables, excluding the SMI, BIA measures and functional performance variables (n = 78).

We should note that, the values of SMI are the most relevant data to detect presence of Sarcopenia in the elderly (Cruz-Jentoft et al., 2010a). Due to this, SMI data was excluded from the machine learning analysis, since that it would have the most discriminative potential, so that it could affect the real predictive value of the rest of variables included in the study. BIA measures and/or those related to the functional performance were also excluded in the sceneries B and C, respectively, in order to identify the discriminative potential of those variables that are not directly related to current criteria for diagnosing Sarcopenia at the clinical practice.

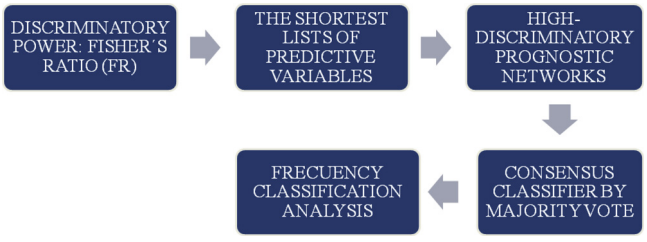


Fig. 1. Flow chart of machine learning methodology.

The stability of high-discriminative prognostic networks was examined by random hold-out experiments, using 75% of samples for training and 25% for validation, as well as a simple statistical analysis of the random hold-out results, providing the median, accuracies, the inter-quartile and the standard deviation of the predictive accuracy.

2.3. Statistical analysis

Pearson´s correlation coefficients were calculated to show correlations network among the most discriminative variables (headers) of Sarcopenia. Bivariate logistic regression was performed, including those variables most frequents in the samples classification. The calibration of this logistic equation was performed by the R^2 determination coefficient, and the area under the receiver operator characteristic (ROC) curve with its 95% confidence intervals. The model fit was verified using the Hosmer-Lemeshow goodness-of-fit test as well as its prognostic sensibility based on percentages of predictive successes total and for each class (G1 and G2). Statistical analyses were performed using SPSS, version 15.0 for Windows (SPSS Inc., Chicago, IL, USA).

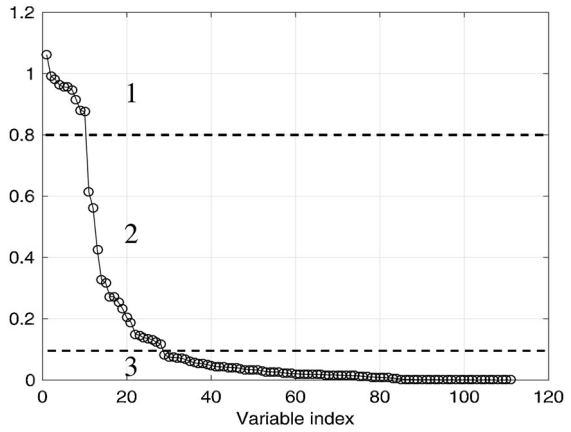
3. Results

The baseline characteristics of the population sample are detailed in supplementary Table 1. According to our previous studies (Coto Montes et al., 2017) a 26% (n = 52) of Sarcopenia prevalence was observed. Among them, a 23.1% (n = 12) of participants with decreased skeletal muscle mass and slow gait speed was observed. 28.8% (n = 15) of the participants showed decreased skeletal muscle mass and poor grip strength on the dominant hand. Notably, a total of 48.1% (n = 25) of the participants presented decreased skeletal muscle mass together with slow gait speed and deficient hand-grip strength, which might represent elderly individuals with a severe Sarcopenia, as considered by the EWGSOP (Cruz-Jentoft et al., 2010a, b).

The distribution of Fisher's ratio (FR) of all variables is shown in Fig. 2A. The curve presents 3 regions: (1) the variables with FR greater than 0.8 (n = 10). (2) The central part with FR between 0.8 and 0.1 (n = 18). (3) The tail, with FR lower than 0.1 (n = 84), where the curve gets almost horizontal. These last are helper variables and serve to expand high-frequency details in the samples classification (Saligan et al., 2014). Additionally, those variables with the highest discriminatory power (header variables, n = 28), as considered by a FR equal or higher than 0.1, are listed in Fig. 2B, thus providing those variables that optimally separate both classes (G1 *versus* G2). The variable with the highest discriminative power was the fat-free mass (FFM) with a FR of 1.06 (Fig. 2B).

The correlations network among these 28 most predictive variables for Sarcopenia is showed in Fig. 3. The fat-free mass (FFM) was the header prognostic variable with 3 sub-headers, the muscle mass on the right leg (RLMM), the fat-free mass of the trunk (TrFFM) and the muscle mass on the right arm (RAMM). These variables have strong correlation with the FFM of 0.92, 0.97 and 0.94, respectively, and delimit three main predictive branches in this correlations network. Notably, different BIA measures in legs, trunk and left arm together with anthropometry data (weight, the body muscle index and abdominal girth) correlated to RLMM. BIA measures in the trunk, epidemiological and clinical variables (gender, height, and depression) as well as functional performance data (gait speed and grip force) correlated to TrFFM. Finally, BIA measures on the right arm, the upper limb performance, weekly kilocalories consumed and hemoglobin levels correlated to

A.



B.

Header variables (n= 28)			
Variable	Fisher´s ratio	Variable	Fisher´s ratio
FFM	1.06	**GF**	0.32
RAFFM	0.99	**RAF**	0.27
RLFFM	0.98	**GS4mC**	0.27
RAMM	0.96	**TrF**	0.25
TrMM	0.96	**TF**	0.23
TrFFM	0.96	**GFC**	0.20
RLMM	0.95	**Gd**	0.19
Wg	0.91	**LAF**	0.15
LLFFM	0.88	**Sz**	0.15
LLMM	0.88	**LAMM**	0.14
WK	0.61	**LAFFM**	0.14
BMI	0.56	**GS4m**	0.13
ULP	0.42	**Hmg**	0.12
AG	0.33	**Dp**	0.12

Fig. 2. Discriminative power of the study variables (A) Fisher's ratio (FR) distribution for the Sarcopenia classification problem. (1), variables (n = 10) with a FR greater than 0.8 value. (2), variables (n = 18) with a FR between 0.8 and 0.1 values. (3), variables (n = 84) with a FR lower than 0.1 value. (B) Header variables (n = 28) with the highest discriminatory power between healthy and Sarcopenic participants, as indicated by a FR equal or higher than 0.1. Abbreviations are extensively described in supplementary Table 1.

RAMM (Fig. 3).

The shortest list of predictive variables, ordered by their individual discriminatory power (FR), together with the best predictive accuracies obtained for each of the possible sceneries considered (82% for A, 80% for B and 77% for C) is summarized in supplementary Table 2. Variables related to epidemiology (gender and height), anthropometry and body composition (weight, body mass index, and abdominal girth), and functional performance (weekly kilocalories consumed and hemoglobin levels) as well as the presence of depression were common among the three sceneries analyzed (supplementary Table 2).

Table 1 shows the high-discriminatory prognostic networks for Sarcopenia in the different sceneries considered. By this way, the predictive accuracy obtained increased to 86.5% in case A, 84.5% in case B, and 83.5% in case C. Likewise, a stability analysis was performed to validate these prognostic networks (Table 1, bottom). It can be observed that the median as well as mean accuracies are very close, with low inter-quartile range and low standard deviations. Therefore, these prognostic networks can be considered very stable.

The high-discriminatory prognostic network for the scenery A presented 13 predictive variables. It mainly included BIA measurements, specifically the fat-free mass and various variables related to body fat levels (RAF, TF, LAF, TrF, RAFR, RLFR) as well as values of anthropometry (BMI, AG), functional performance (WK, GS4mC) and medication for the digestive system (D-PPI) and the central nervous system (D-Bdz) (Table 1).

The high-discriminatory prognostic network for the scenery B presented 22 predictive variables. It mainly included anthropometry data (Wg, BMI), measurements of functional performance (GS4mC, RMR, LC, LTI), biochemistry (Gl, Lk, Try), parameters related to oxidative stress (LP, TAC) as well as clinical conditions of depression, hypertension and dyslipidemia (Dp, YI, Hyp, Dl). Medication for the following systems was also included in this scenery: digestive (D-PPi), nervous (ND-Bdz, D-Ach), respiratory (D-RS), urinary (D-AU, D-Dr) and cardiovascular (D-C) (Table 1).

The high-discriminatory prognostic network for the scenery C presented 10 predictive variables. It mainly included weight (Wg), clinical variables of depression (YI), hypertension (Hyp) and comorbidity (CHI), biochemistry (Try, Lph) and medication for the following systems: locomotor (D-LS), cardiovascular (D-H, D-C) and the central nervous system (D-Adp) (Table 1).

By using all different prognostic networks detected with a predictive accuracy higher than 82%, we have designed a consensus classifier for the Sarcopenia that increase the final predictive accuracy up to 91%. This consensus classifier works by majority voting to predict (G1 *versus* G2) the class of new incoming samples. The most frequent variables in the samples classification are listed in Table 2. The common variables among the 3 sceneries considered include the body mass index (BMI), weekly kilocalories consumed (WK), presence of hypertension (Hyp), dyslipidemia (Dl), systemic levels of lipid peroxidation (LP), neutrophils levels (Ntph), the total antioxidant activity (TAC) as well as drugs for the digestive system (D-PPI) and the central nervous system (D-Bdz, ND-Bdz) (Table 2).

Fig. 4 shows the interplay among the most frequent variables in the samples classification and with the major importance in the vote by majority consensus to predict Sarcopenia in the elderly. Specifically, this predictive signature included BIA measures (FFM), anthropometry (BMI, Wg), levels of oxidative stress (LP, TAC), functional performance (GS4mC) as well as presence of depression and medication for the digestive system (D-PPI) (Fig. 4A). These variables predict a 77.1% of the variance of Sarcopenia in our population. The Hosmer-Lemeshow ($p > 0.05$) test revealed a robust goodness of fit for this regression model. The discrimination power of this predictive signature was of 97.1% (95% CI, 0.951-0.991, $p < 0.001$) showing a total sensibility prognostic of 91.5%, with a 94.6% of successes in the prediction of the G1 class (healthy participants) in comparison to a 82.7% of successes in the prediction of the G2 class (Sarcopenia) (Fig. 4B).

4. Discussion

The present study allows to deep into the interplay of those main biological mechanisms imply in the onset and development of Sarcopenia. Machine learning methodology has been the initial step to find those biological variables mainly implicated in the diagnosis of Sarcopenia. These techniques were already used to solve other medical problems with clinical data, such as prediction of treatment response in Hodgkin lymphoma (De Andres-Galiana et al., 2015), and several decision making problems in patients with chronic lymphocytic leukemia (De Andres-Galiana et al., 2016a). Besides, it has been successfully applied into bioinformatics modeling of highly dimensional omics data (Saligan et al., 2014). This methodology revealed that the fat-free mass is the most discriminative variable for Sarcopenia. Accordingly, the best prognostic network of Sarcopenia, with a predictive accuracy of 86.5%, mainly selected those variables related to BIA measures in different body parts, what could be expected, considering that main predictors of Sarcopenia are those variables directly or indirectly related to measures of muscle mass. Notably, body fat data determined by BIA were also included in this prognostic network. These results support that not only changes on the lean mass are directly related to Sarcopenia, but also implies that body fat levels have an important role in triggering

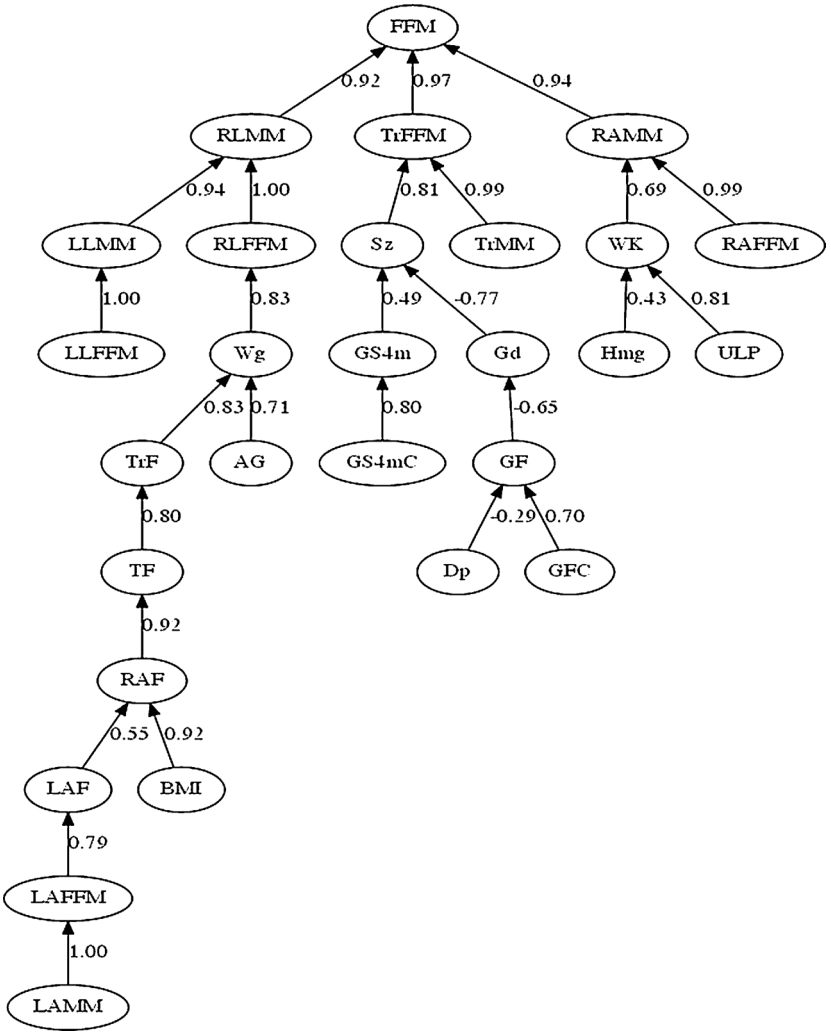


Fig. 3. Correlations network among the most predictive variables of Sarcopenia. Pearson´s correlation coefficients were calculated among the most discriminative variables (header variables) of Sarcopenia. Abbreviations are extensively described in supplementary Table 1.

Sarcopenia, according to several works that propose fat infiltration in the muscle as a relevant factor for Sarcopenia development (Auyeung et al., 2013; Bucci et al., 2013; Coto Montes et al., 2017; Choi, 2013).

Once BIA measurements were not considered, other variables achieved predictive potential, including anthropometry data, functional performance of the lower extremities, depression condition, hypertension as well as different types of medication and biochemical parameters, including those related to oxidative stress (systemic lipid peroxidation and the total antioxidant capacity). Although we observed stability in the three prognostic networks showed, we should note that the predictive combination of variables included in the scenery B was too big, since 22 predictive variables were included. Therefore, it would attract less interest for the clinical point of view. This prognostic network also revealed that impairments on functional performance, by themselves, do not have enough prognostic power for sarcopenia, unless they are combined with BIA data. This way, after excluding both BIA data as well as measures of functional performance in our mathematical modeling, we obtained a prognostic network of Sarcopenia with the lowest prognostic accuracy (83.5%). Besides these three better prognostic networks, there are several other predictive combinations that involve different networks of prognostic variables that may play a role in the binary classification (Healthy *versus* Sarcopenia) of the samples. Using all those prognostic networks with accuracy higher than 82%, we have implemented a consensus classifier by majority vote that increases the predictive accuracy of Sarcopenia to 91%.

The current diagnosis of Sarcopenia according to the EWGSOP always takes into account measurements of functional performance (gait speed) and muscle strength (grip force), besides loss of muscle mass, indicated by the SMI and assayed by BIA, among other methodologies proposed (Cruz-Jentoft et al., 2010a). In our mathematical prediction, we have always excluded the SMI, due to its relevance for the current clinical diagnosis of Sarcopenia in the elderly (Cruz-Jentoft et al., 2010a). This way, we were able to study the discriminative and/or predictive potential of other variables involved in the diagnosis of Sarcopenia, gait speed and grip force. Notably, both data were identified as header variables due to their high discriminative power, as indicated by their FR, and both were included in the shortest lists of predictive variables. We should note that, the predictive accuracy was always increased when gait speed data were considered in these predictive lists. Also, the high-discriminative prognostic networks with the highest predictive accuracy only selected gait speed as prognostic variable. Additionally, gait speed showed a higher frequency than grip force in the consensus classifier by majority vote. Given these premises, our present results strongly support the validity and reliability of the current European consensus for a clinical diagnosis of Sarcopenia, but prioritizing the component of functional performance on the lower limbs over the muscle strength component. Therefore, functional performance of the lower limbs seems to be paramount for the current diagnosis of Sarcopenia. Accordingly, this component has been also proposed as the main indicator affected for developing the condition of

Table 1
Predictive high-discriminative networks with their corresponding predictive accuracies together with their stability analysis (bottom table).

SCENERY A	SCENERY B	SCENERY C
FFM	Wg	Wg
WK	BMI	YI
BMI	GS4mC	Hyp
AG	Dp	Try
RAF	YI	Lph
GS4mC	RMR	D-LS
TF	D-PPI	D-H
LAF	Hyp	D-Adp
D-PPI	ND-Bdz	CHI
D-Bdz	LP	D-C
TrF	D-RS	
RAFR	Gl	
RLFR	Lk	
	Try	
	TAC	
	D-AU	
	D-Dr	
	D-Ach	
	LC	
	LTI	
	D-C	
	Dl	

Predictive accuracy	**86.50**	**84.50**	**83.50**
Median	84	82	80
Mean	84.33	82.83	79.77
Inter-quartile	6	6	6
Standard deviation	4.43	4.63	4.4

Table 2
Most frequent variables in the samples classification based on majority vote in the different sceneries considered. Common variables among sceneries appear on grey background. Variables in bold indicate those variables with the most prognostic importance for each one of the scenery considered (A, B and C). Abbreviations are extensively described in supplementary Table1.

SCENERY A		SCENERY B		SCENERY C	
Variable	**Frequency**	**Variable**	**Frequency**	**Variable**	**Frequency**
FFM	**2.75**	**Wg**	**5.26**	**Wg**	**6.37**
BMI	**2.66**	**BMI**	**3.65**	**BMI**	**4.46**
LP	**2.47**	**Dp**	**3.05**	**TAC**	**4.25**
D-PPI	**2.18**	TAC	2.97	D-LS	3.82
GS4mC	**2.18**	YI	2.86	BPC	3.61
Sz	1.99	D-AU	2.79	Hyp	3.61
LLFFR	1.90	LP	2.70	D-Hy	3.61
D-Ep	1.80	Hyp	2.70	YI	3.18
D-Bdz	1.80	BPC	2.67	D-AU	3.18
TrFFM	1.80	D-Hy	2.66	ND-Bdz	2.97
Cr	1.71	D-LS	2.62	LP	2.76
Ntph	1.71	Mn	2.57	D-PPI	2.76
D-DS	1.71	ND-Bdz	2.45	Dp	2.55
TrFR	1.71	AG	2.38	Y	2.34
GFC	1.71	D-C	2.38	Try	2.34
Gd	1.61	D-PPI	2.15	Lph	2.34
Hyp	1.61	Y	2.05	Dl	2.34
CdN	1.61	Dl	2.05	AG	2.12
RAF	1.61	HDL	1.98	HDL	2.12
LAFFM	1.61	D-DS	1.91	Ntph	2.12
TrF	1.61	RMR	1.89	WK	1.91
TrMM	1.61	WK	1.84	CHI	1.91
D-Apl	1.52	LC	1.83	D-C	1.91
Dl	1.42	CHI	1.76	Mn	1.70
LLP	1.42	Try	1.75	D-DS	1.70
ULP	1.42	CPR	1.75	Cr	1.49
RLFFM	1.42	D-Dr	1.72	CPR	1.49
LLMM	1.42	LTI	1.65	Lk	1.49
TAC	1.33	CdN	1.57	D-Adp	1.49
WK	1.33	Ntph	1.55	Gl	1.27

pre-fragility in the elderly (Caballero et al., 2014; Cruz-Jentoft et al., 2010a).

Finally, and based on the most frequent variables for sample classification and with the highest importance in the consensus by majority vote, we propose a predictive signature for Sarcopenia, with high-discriminative power and sensitivity, in order to go deeper into the interplay among the main biological mechanisms involved in the development of Sarcopenia in the elderly. In accord with our results, the most important variable within this predictive equation was gait speed at 4 m. Although recent works emphasize a low impact of Sarcopenia on body mass index (Benton and Silva-Smith, 2018), our results demonstrate a relevant role of changes in body mass for predicting Sarcopenia, since this variable was the second in importance in our logistic equation. The next variable of importance in the equation is medication for the digestive system, specifically proton-pump inhibitors, which are widely used to treat gastric acid-related disorders. The use of this type of medication is frequent in older adults and it is important to note that a long-time therapy with proton-pump inhibitors has been associated with several adverse effects including anemia, enteric infections, dementia, pneumonia, osteoporosis, bone fractures at multiple sites as well as ischemic heart and chronic kidney diseases (Vaezi et al., 2017; Wang et al., 2017). Likewise, prolonged use of proton-pump inhibitors may also induce myopathies (Vaezi et al., 2017) that increase the risk of muscle atrophy. Thus, our data support that negative effects of proton-pump inhibitors on the muscle quality and/or function are risk factors relevant for developing Sarcopenia in the elderly.

Recent works have demonstrated that depressive symptoms may increase the age-related deterioration of body composition (Walther et al., 2017). Accordingly, presence of depression was also included in our logistic equation. The association between Sarcopenia and depression has been observed in several recent studies, since this clinical condition is more than 2 fold prevalent in older adult with Sarcopenia as compared to healthy older adults (Coto Montes et al., 2017; Chang et al., 2017; Kilavuz et al., 2018). Other variables with predictive importance in our logistic equation were fat-free mass and the weight. It is well-known that Sarcopenia causes great impact on the lean mass index (Benton and Silva-Smith, 2018), and these values have been always the highest discriminative for Sarcopenia in our mathematical model. However, when BIA measures were excluded, the weight has achieved a notable predictive value instead of the fat-free mass.

Finally, measures of oxidative stress, based on systemic lipid peroxidation levels and the total antioxidant capacity, were also included in our logistic model. Oxidative stress has been extensively correlated to a loss of muscle mass and, therefore, it may favour the onset and development of Sarcopenia (Bellanti et al., 2018; Caballero et al., 2014; Coto Montes et al., 2017; Chung et al., 2018; Marzetti et al., 2012; Zacarias-Flores et al., 2018). In fact, lipid peroxidation has been recently described as the oxidative stress related most contributing variable to the skeletal muscle loss in Sarcopenia (Zacarias-Flores et al., 2018). Our recent works also suggest that systemic lipid peroxidation levels could be used as an early redox biomarker for Sarcopenia detection in independent older adults, before disability or functional dependence appears (Coto Montes et al., 2017). Likewise, it was observed an association between low levels of antioxidants and adverse clinical outcomes in the elderly, including deficiencies on functional physical performance of the upper body extremities (Caballero et al., 2014), the presence of frailty (Ble et al., 2006), Sarcopenia (Bellanti et al., 2018) and even an increased risk of death (Li et al., 2011; Shardell et al., 2011). Therefore, our present data support the fact that oxidative stress is a key component in the physiopathology of Sarcopenia, in accord with several works that propose a dietary control based on antioxidants as one of the best clinical intervention to prevent the onset of Sarcopenia (Chung et al., 2018; Guescini et al., 2017; Haramizu et al., 2017; Khor et al., 2014; Molino-Lova et al., 2017; van Dijk et al., 2018v).

Some limitations of this study must be considered. In order to include as many types of variables as possible, we worked with an urban population sample of just 200 participants. In this way, longitudinal studies in population sample of higher size should be performed to better document the validity and selectivity of our prognostic networks

A.

SARCOPENIA = 31.925 - 1.990*__GS4mC__ - 1.345*__BMI__ + 1.142*__D-PPI__ + 1.118*__Dp__

- 0.756*__FFM__ + 0.538*__Wg__ + 0.269*__TAC__ + 0.069*__LP__

B.

STABILITY ANALYSIS	
R^2 coefficient	0.771
H-L test	0.834
AUC (95 %CI)	0.971 (0.951-0.991)
p-value	< 0.001
Sensibility	91,5 % (94,6 % for G1 & 82,7 % for G2)

Fig. 4. Interplay among the most discriminative biological variables of Sarcopenia. (A) Predictive signature for Sarcopenia based on the most frequent variables and with the best prognostic importance in the vote by majority consensus (Frequency > 2 for A, Frequency > 3 for B, Frequency > 4 for C). (B) Stability analysis of this predictive signature of Sarcopenia. Abbreviations: AUC, the area under the receiver operator characteristic (ROC) curve; BMI, Body Mass Index; CI, confidence interval; Dp, Depression (1 = No, 2 = Yes); D-PPI, Drugs-Proton Pump Inhibitors (0 = No, 1 = Yes); FFM, Fat-Free Mass; GS4mC, categories of Gait speed at 4 m (1 = ≤ 0.8 m/s, 2 = > 0.8 m/s); H–L, Hosmer-Lemeshow test; LP, Lipid peroxidation; R^2, determination coefficient; TAC; Total Antioxidant Capacity; Wg, Weight. Sarcopenia is considered with values ≥ 0.971.

and predictive models.

5. Conclusions

On one hand, machine learning approaches had allowed us to define those biological variables with the highest predictive power for Sarcopenia. These variables include measures of body composition and anthropometry, functional performance on the lower limbs, clinical conditions such as depression, biochemical data that include systemic levels of oxidative stress as well as medication for the digestive system, in order to establish the best prognostic networks that distingue between healthy and sarcopenic older adults. These different prognostic networks can predict classification of new incoming samples by a consensus classifier based on majority voting with a high accuracy of 91%. On the other hand, in order to reveal the interplay of those main biological mechanisms implicated in the onset and development of Sarcopenia, we propose a predictive signature of high sensibility (91.5%) and discriminative power (97.1%) based on most frequent variables and with the major importance in the consensus classifier of Sarcopenia. Given our present data, in order to maintain quality and function of muscle mass, it is relevant a dietary control of fat levels, increase intake of antioxidants and appropriate physical exercises that increase functionality of the lower limbs. Likewise, it is also important to provide an emotional and social support to the elderly with Sarcopenia, as well as a strict control of medication, specifically those drugs that affect the digestive system.

Declaration of Competing Interest

None.

Acknowledgments

JC Bermejo-Millo, JJ Solano, P Abizanda, A Coto-Montes and B Caballero also belong to CIBERFES. A Coto-Montes and B Caballero are part of INEUROPA. JC Bermejo-Millo thanks his pre-doctoral fellowship from Instituto de Investigación Sanitaria del Principado de Asturias (ISPA). This work was supported by the *Instituto de Salud Carlos III* [Grant number FISS-18-PI17/02009]. This funding source has not been involved in the study design, collection, analysis and interpretation of data, writing the report or in the decision to submit the article for publication.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.mad.2019.111129.

References

Abizanda Soler, P., López-Torres Hidalgo, J., Romero Rizos, L., López Jiménez, M., Sánchez Jurado, P.M., Atienzar Núnez, P., Esquinas Requena, J.L., García Nogueras, I., Hernández Zegarra, P., Bardales Mas, Y., Campos Rosa, R., Martínez Peñalver, M., de la Osa Nieto, E., Carión González, M., Ruiz Gómez, A., Aguilar Cantos, C., Mañueco Delicado, P., Oliver Carbonell, J.L., 2011. Frailty and dependence in Albacete (FRADEA study): reasoning, design and methodology. Rev. Esp. Geriatr. Gerontol. 46, 81–88. https://doi.org/10.1016/j.regg.2010.10.004.

Auyeung, T.W., Lee, J.S., Leung, J., Kwok, T., Woo, J., 2013. Adiposity to muscle ratio predicts incident physical limitation in a cohort of 3,153 older adults–an alternative measurement of sarcopenia and sarcopenic obesity. Age (Dordr). 35, 1377–1385. https://doi.org/10.1007/s11357-012-9423-9.

Bellanti, F., Romano, A.D., Lo Buglio, A., Castriotta, V., Guglielmi, G., Greco, A., Serviddio, G., Vendemiale, G., 2018. Oxidative stress is increased in sarcopenia and associated with cardiovascular disease risk in sarcopenic obesity. Maturitas. 109, 6–12. https://doi.org/10.1016/j.maturitas.2017.12.002.

Benton, M.J., Silva-Smith, A.L., 2018. Accuracy of body mass index versus lean mass index for prediction of Sarcopenia in older women. J. Frailty Aging 7, 104–107. https://doi.org/10.14283/jfa.2018.1.

Ble, A., Cherubini, A., Volpato, S., Bartali, B., Walston, J.D., Windham, B.G., Bandinelli, S., Lauretani, F., Guralnik, J.M., Ferrucci, L., 2006. Lower plasma vitamin E levels are associated with the frailty syndrome: the InCHIANTI study. J. Gerontol. A Biol. Sci. Med. Sci. 61, 278–283. https://doi.org/10.1093/gerona/61.3.278.

Brzeszczynska, J., Meyer, A., McGregor, R., Schilb, A., Degen, S., Tadini, V., Johns, N., Langen, R., Schols, A., Glass, D.J., Roubenoff, R., Ross, J.A., Fearon, K.C.H., Greig, C.A., Jacobi, C., 2018. Alterations in the in vitro and in vivo regulation of muscle regeneration in healthy ageing and the influence of sarcopenia. J. Cachexia Sarcopenia Muscle 9, 93–105. https://doi.org/10.1002/jcsm.12252.

Bucci, L., Yani, S.L., Fabbri, C., Bijlsma, A.Y., Maier, A.B., Meskers, C.G., Narici, M.V., Jones, D.A., McPhee, J.S., Seppet, E., Gapeyeva, H., Paasuke, M., Sipila, S., Kovanen, V., Stenroth, L., Musaro, A., Hogrel, J.Y., Barnouin, Y., Butler-Browne, G., Capri, M., Franceschi, C., Salvioli, S., 2013. Circulating levels of adipokines and IGF-1 are associated with skeletal muscle strength of young and old healthy subjects. Biogerontology. 14, 261–272. https://doi.org/10.1007/s10522-013-9428-5.

Caballero, B., Rubio-Gonzalez, A., Potes, Y., Martinez-Reig, M., Sanchez-Jurado, P.M., Romero, L., Solano, J.J., Abizanda, P., Coto-Montes, A., 2014. Associations of the antioxidant capacity and hemoglobin levels with functional physical performance of the upper and lower body limbs. Age (Dordr). 36, 851–867. https://doi.org/10.1007/s11357-013-9607-y.

Cernea, A., Fernández-Martínez, J.L., deAndrés-Galiana, E.J., Fernández-Ovies, F.J., Fernández-Muñiz, Z., Alvarez-Machancoses, O., Saligan, L., Sonis, S.T., 2018. Sampling defective pathways in phenotype prediction problems via the Fisher's ratio sampler. In: Rojas, I., Ortuño, F. (Eds.), Bioinformatics and Biomedical Engineering. IWBBIO, pp. 24–32. https://doi.org/10.1007/978-3-319-78759-6_3. 2018. Springer Nature LNBI. 10814.

Coto Montes, A., Boga, J.A., Bermejo Millo, C., Rubio Gonzalez, A., Potes Ochoa, Y., Vega Naredo, I., Martinez Reig, M., Romero Rizos, L., Sanchez Jurado, P.M., Solano, J.J., Abizanda, P., Caballero, B., 2017. Potential early biomarkers of sarcopenia among independent older adults. Maturitas 104, 117–122. https://doi.org/10.1016/j.maturitas.2017.08.009.

Cruz-Jentoft, A.J., Baeyens, J.P., Bauer, J.M., Boirie, Y., Cederholm, T., Landi, F., Martin, F.C., Michel, J.P., Rolland, Y., Schneider, S.M., Topinkova, E., Vandewoude, M., Zamboni, M., 2010a. Sarcopenia: European consensus on definition and diagnosis: report of the European working group on Sarcopenia in older people. Age Ageing 39, 412–423. https://doi.org/10.1093/ageing/afq034.

Cruz-Jentoft, A.J., Landi, F., Topinkova, E., Michel, J.P., 2010b. Understanding sarcopenia as a geriatric syndrome. Curr. Opin. Clin. Nutr. Metab. Care. 13, 1–7.

Chang, K.V., Hsu, T.H., Wu, W.T., Huang, K.C., Han, D.S., 2017. Is sarcopenia associated with depression? A systematic review and meta-analysis of observational studies. Age Ageing 46, 738–746. https://doi.org/10.1093/ageing/afx094.

Choi, K.M., 2013. Sarcopenia and sarcopenic obesity. Endocr. Metab. 28, 86–89. https://doi.org/10.3803/EnM.2013.28.2.86.

Chung, E., Mo, H., Wang, S., Zu, Y., Elfakhani, M., Rios, S.R., Chyu, M.C., Yang, R.S., Shen, C.L., 2018. Potential roles of vitamin E in age-related changes in skeletal muscle health. Nutr. Res. 49, 23–36. https://doi.org/10.1016/j.nutres.2017.09.005.

De Andres-Galiana, E.J., Fernandez-Martinez, J.L., Luaces, O., Del Coz, J.J., Fernandez, R., Solano, J., Nogues, E.A., Zanabilli, Y., Alonso, J.M., Payer, A.R., Vicente, J.M., Medina, J., Taboada, F., Vargas, M., Alarcon, C., Moran, M., Gonzalez-Ordonez, A., Palicio, M.A., Ortiz, S., Chamorro, C., Gonzalez, S., Gonzalez-Rodriguez, A.P., 2015. Clin. Transl. Oncol. 17, 612–619. https://doi.org/10.1007/s12094-015-1285-z.

De Andres-Galiana, E.J., Fernandez-Martinez, J.L., Luaces, O., Del Coz, J.J., Huergo-

Zapico, L., Acebes-Huerta, A., Gonzalez, S., Gonzalez-Rodriguez, A.P., 2016a. Analysis of clinical prognostic variables for Chronic Lymphocytic Leukemia decision-making problems. J. Biomed. Inform. 60, 342–351. https://doi.org/10.1016/j.jbi.2016.02.017.

De Andres-Galiana, E.J., Fernandez-Martinez, J.L., Sonis, S.T., 2016b. Design of biomedical robots for phenotype prediction problems. J. Comput. Biol. 23, 678–692. https://doi.org/10.1089/cmb.2016.0008.

Guescini, M., Tiano, L., Genova, M.L., Polidori, E., Silvestri, S., Orlando, P., Fimognari, C., Calcabrini, C., Stocchi, V., Sestili, P., 2017. The combination of physical exercise with muscle-directed antioxidants to counteract Sarcopenia: a biomedical rationale for pleiotropic treatment with creatine and coenzyme Q10. Oxid. Med. Cell. Longev. 2017, 7083049. https://doi.org/10.1155/2017/7083049.

Haramizu, S., Asano, S., Butler, D.C., Stanton, D.A., Hajira, A., Mohamed, J.S., Alway, S.E., 2017. Dietary resveratrol confers apoptotic resistance to oxidative stress in myoblasts. J. Nutr. Biochem. 50, 103–115. https://doi.org/10.1016/j.jnutbio.2017.08.008.

Khor, S.C., Abdul Karim, N., Ngah, W.Z., Yusof, Y.A., Makpol, S., 2014. Vitamin E in sarcopenia: current evidences on its role in prevention and treatment. Oxid. Med. Cell. Longev. 2014, 914853. https://doi.org/10.1155/2014/914853.

Kilavuz, A., Meseri, R., Savas, S., Simsek, H., Sahin, S., Bicakli, D.H., Sarac, F., Uyar, M., Akcicek, F., 2018. Association of sarcopenia with depressive symptoms and functional status among ambulatory community-dwelling elderly. Arch. Gerontol. Geriatr. 76, 196–201. https://doi.org/10.1016/j.archger.2018.03.003.

Li, C., Ford, E.S., Zhao, G., Balluz, L.S., Giles, W.H., Liu, S., 2011. Serum alpha-carotene concentrations and risk of death among US adults: the third national health and nutrition examination survey follow-up study. Arch. Intern. Med. 171, 507–515. https://doi.org/10.1001/archinternmed.2010.440.

Marzetti, E., Calvani, R., Bernabei, R., Leeuwenburgh, C., 2012. Apoptosis in skeletal myocytes: a potential target for interventions against sarcopenia and physical frailty - a mini-review. Gerontology 58, 99–106. https://doi.org/10.1159/000330064.

Molino-Lova, R., Sofi, F., Pasquini, G., Vannetti, F., Del Ry, S., Vassalle, C., Clerici, M., Sorbi, S., Macchi, C., 2017. Higher uric acid serum levels are associated with better muscle function in the oldest old: results from the Mugello study. Eur. J. Intern. Med. 41, 39–43. https://doi.org/10.1016/j.ejim.2017.03.014.

Saligan, L.N., Fernandez-Martinez, J.L., De Andres-Galiana, E.J., Sonis, S., 2014. Supervised classification by filter methods and recursive feature elimination predicts risk of radiotherapy-related fatigue in patients with prostate cancer. Cancer Inform. 13, 141–152. https://doi.org/10.4137/CIN.S19745.

Shardell, M.D., Alley, D.E., Hicks, G.E., El-Kamary, S.S., Miller, R.R., Semba, R.D., Ferrucci, L., 2011. Low-serum carotenoid concentrations and carotenoid interactions predict mortality in US adults: the Third National Health and Nutrition Examination Survey. Nutr. Res. 31, 178–189. https://doi.org/10.1016/j.nutres.2011.03.003.

Tong, J., Li, W., Vidal, C., Yeo, L.S., Fatkin, D., Duque, G., 2011. Lamin A/C deficiency is associated with fat infiltration of muscle and bone. Mech. Ageing Dev. 132, 552–559. https://doi.org/10.1016/j.mad.2011.09.004.

Vaezi, M.F., Yang, Y.X., Howden, C.W., 2017. Complications of proton pump inhibitor therapy. Gastroenterology 153, 35–48. https://doi.org/10.1053/j.gastro.2017.04.047.

van Dijk, M., Dijk, F.J., Hartog, A., van Norren, K., Verlaan, S., van Helvoort, A., Jaspers, R.T., Luiking, Y., 2018v. Reduced dietary intake of micronutrients with antioxidant properties negatively impacts muscle health in aged mice. J. Cachexia Sarcopenia Muscle 9, 146–159. https://doi.org/10.1002/jcsm.12237.

Walther, A., Philipp, M., Lozza, N., Ehlert, U., 2017. Emotional support, depressive symptoms, and age-related alterations in male body composition: cross-sectional findings from the men's health 40+ study. Front. Psychol. 8, 1075. https://doi.org/10.3389/fpsyg.2017.01075.

Wang, L., Li, M., Cao, Y., Han, Z., Wang, X., Atkinson, E.J., Liu, H., Amin, S., 2017. Proton pump inhibitors and the risk for fracture at specific sites: data mining of the FDA adverse event reporting system. Sci. Rep. 7, 5527. https://doi.org/10.1038/s41598-017-05552-1.

Zacarias-Flores, M., Sanchez-Rodriguez, M.A., Garcia-Anaya, O.D., Correa-Munoz, E., Mendoza-Nunez, V.M., 2018. Relationship between oxidative stress and muscle mass loss in early postmenopause: an exploratory study. Endocrinol. Diabetes Nutr. 65 (6), 328–334. https://doi.org/10.1016/j.endinu.2018.01.009.