

## Article

# A Machine-Learning Approach to Target Clinical and Biological Features Associated with Sarcopenia: Findings from Northern and Southern Italian Aging Populations

Roberta Zupo <sup>1,†</sup> , Alessia Moroni <sup>2,†</sup> , Fabio Castellana <sup>3,\*</sup> , Clara Gasparri <sup>2</sup> , Feliciano Catino <sup>4</sup>,  
Luisa Lampignano <sup>3</sup> , Simone Perna <sup>5</sup> , Maria Lisa Clodoveo <sup>1</sup> , Rodolfo Sardone <sup>6</sup>   
and Mariangela Rondanelli <sup>7,8</sup> 

- <sup>1</sup> Department of Interdisciplinary Medicine, University “Aldo Moro”, Piazza Giulio Cesare 11, 70100 Bari, Italy; roberta.zupo@uniba.it (R.Z.); marialisa.clodoveo@uniba.it (M.L.C.)
  - <sup>2</sup> Endocrinology and Nutrition Unit, Azienda di Servizi alla Persona “Istituto Santa Margherita”, University of Pavia, 27100 Pavia, Italy; alessia.moroni02@universitadipavia.it (A.M.); clara.gasparri01@universitadipavia.it (C.G.)
  - <sup>3</sup> Unit of Data Sciences and Technology Innovation for Population Health, National Institute of Gastroenterology IRCCS “Saverio de Bellis”, Research Hospital, Castellana Grotte, 70013 Bari, Italy; luisa.lampignano@irccsdebellis.it
  - <sup>4</sup> Department of Innovation and Smart City, Municipality of Taranto, 74121 Taranto, Italy; feliciano.catino@comune.taranto.it
  - <sup>5</sup> Department of Food, Environmental and Nutritional Sciences, Division of Human Nutrition, University of Milan, 20133 Milan, Italy; simone.perna@unimi.it
  - <sup>6</sup> Local Healthcare Authority of Taranto, 74121 Taranto, Italy; rodolfo.sardone@asl.taranto.it
  - <sup>7</sup> Department of Public Health, Experimental and Forensic Medicine, University of Pavia, 27100 Pavia, Italy; mariangela.rondanelli@unipv.it
  - <sup>8</sup> IRCCS Mondino Foundation, 27100 Pavia, Italy
- \* Correspondence: fabio.castellana@irccsdebellis.it  
† These authors contributed equally to this work.



**Citation:** Zupo, R.; Moroni, A.; Castellana, F.; Gasparri, C.; Catino, F.; Lampignano, L.; Perna, S.; Clodoveo, M.L.; Sardone, R.; Rondanelli, M. A Machine-Learning Approach to Target Clinical and Biological Features Associated with Sarcopenia: Findings from Northern and Southern Italian Aging Populations. *Metabolites* **2023**, *13*, 565. <https://doi.org/10.3390/metabo13040565>

Academic Editor: Cholsoon Jang

Received: 1 March 2023

Revised: 12 April 2023

Accepted: 13 April 2023

Published: 17 April 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Epidemiological and public health resonance of sarcopenia in late life requires further research to identify better clinical markers useful for seeking proper care strategies in preventive medicine settings. Using a machine-learning approach, a search for clinical and fluid markers most associated with sarcopenia was carried out across older populations from northern and southern Italy. A dataset of adults > 65 years of age ( $n = 1971$ ) made up of clinical records and fluid markers from either a clinical-based subset from northern Italy (Pavia) and a population-based subset from southern Italy (Apulia) was employed ( $n = 1312$  and  $n = 659$ , respectively). Body composition data obtained by dual-energy X-ray absorptiometry (DXA) were used for the diagnosis of sarcopenia, given by the presence of either low muscle mass (i.e., an SMI <7.0 kg/m<sup>2</sup> for males or <5.5 kg/m<sup>2</sup> for females) and of low muscle strength (i.e., an HGS < 27 kg for males or <16 kg for females) or low physical performance (i.e., an SPPB ≤ 8), according to the EWGSOP2 panel guidelines. A machine-learning feature-selection approach, the random forest (RF), was used to identify the most predictive features of sarcopenia in the whole dataset, considering every possible interaction among variables and taking into account nonlinear relationships that classical models could not evaluate. Then, a logistic regression was performed for comparative purposes. Leading variables of association to sarcopenia overlapped in the two population subsets and included SMI, HGS, FFM of legs and arms, and sex. Using parametric and nonparametric whole-sample analysis to investigate the clinical variables and biological markers most associated with sarcopenia, we found that albumin, CRP, folate, and age ranked high according to RF selection, while sex, folate, and vitamin D were the most relevant according to logistics. Albumin, CRP, vitamin D, and serum folate should not be neglected in screening for sarcopenia in the aging population. Better preventive medicine settings in geriatrics are urgently needed to lessen the impact of sarcopenia on the general health, quality of life, and medical care delivery of the aging population.

**Keywords:** sarcopenia; nutrition; body composition; machine learning; artificial intelligence; elderly; older adults; aging; Salus in Apulia; Italy

## 1. Introduction

As the burden of population aging increases [1], a multidisciplinary research effort is needed to fill the knowledge gap around risk biopathways and to foster preventive strategies against sarcopenia, a multifactorial syndrome characterized by a progressive and generalized loss of skeletal muscle mass as well as poor physical endurance, which is often combined with subclinical systemic inflammation [2]. Such a decline in skeletal muscle function is a nearly inevitable part of the aging process and has a considerable impact on health care costs and quality of life, since it raises the chance of negative outcomes including falls, fractures, physical impairments, and mortality [3].

Sarcopenia is a geriatric giant triggered by adverse muscle changes commonly experienced late in life. According to the Revised European Consensus on Definition and Diagnosis (EWGSOP2) [4], the dimensions that best define sarcopenia are low levels of three parameters: muscle strength, muscle quantity/quality, and physical performance, of which the latter is an indicator of severity. Hence, poor muscle quantity and quality confirms the presence of sarcopenia, whereas low physical performance clearly rates its severity. Moreover, to detect this condition early in clinical settings, the consensus group focused on a novel algorithm path, so called FACS (Find—Assess—Confirm—Severity), that also took into account the probability of sarcopenia, with reference to the SARC-F questionnaire or the calf circumference. As for the assessment of the muscle mass, a large number of tools are detailed in the literature, but only a few are effectively applicable in the clinical setting and therefore considered in the consensus. Of these, MRI and computed tomography (CT) are the gold standard, but they cannot really be applied in a context beyond research and thus are poorly understood in the clinical setting [5]. Dual-energy X-ray absorptiometry (DXA) is the most available, reliable, and feasible approach to directly assess body composition variables [6], while bioelectrical impedance analysis (BIA) is an alternative indirect method—low-cost compared to DXA—that can be used to screen a much larger population, as DXA is not handheld and cannot be applied in specific populations (e.g., pregnant women, bedridden patients, etc.).

Understanding multimodal indicators that can successfully anticipate the onset of sarcopenia and avert the deleterious cascade of late-life multimorbidity remains an open issue from a preventative standpoint. Given sarcopenia as a multifactorial condition, a single fluid or clinical marker cannot be easily pinpointed or be helpful, and thus focus turns to the implementation of a panel that includes multidomain markers. Further, ideal markers of sarcopenia should be valid, replicable, reliable, specific, affordable, and easily available [7]. In the field of nutrition, the scientific community openly acknowledges haemoglobin, albumin, leptin, uric acid, iron, and vitamin D, amongst others, in predicting the risk of sarcopenia [8–10].

Of note, the heterogeneous biological, clinical, and social complexities involving the individual, as well as the different characteristics of the personnel and place where such assessments are conducted, inevitably play a role in decisions about risk variables, their cut-points, and their ranking [11,12]. For these reasons, there could be arguments that a single gold standard variable as well as the best cut-points could translate poorly from the epidemiology field and computational modelling to the real clinical practice. To address this challenge and add to literature surrounding the biopathways of sarcopenia, this research sought to generate scientific evidence by testing a convenience sample composed of two subsets, i.e., one clinical-based and one population-based, both aged over 65 years. Machine-learning-based techniques were implemented and compared in order to evaluate the clinical and fluid markers most associated with the sarcopenia condition across the two population settings.

## 2. Materials and Methods

### 2.1. Study Population

#### 2.1.1. Northern-Italy Population Subset (Santa Margherita Institute, Pavia)

We enrolled male and female subjects aged  $\geq 65$  years, with a body mass index (BMI) between 20 and 30 kg/m<sup>2</sup> [13], who are outpatients at the metabolic rehabilitation unit of the Santa Margherita Institute, Department of Public Health, University of Pavia. Subjects with the following conditions were excluded from the study: severe kidney disease (glomerular filtration rate < 30 mL/min), moderate-to-severe hepatic failure (Child–Pugh Class of B or C), endocrine diseases associated with disorders of calcium metabolism (with the exception of osteoporosis), psychiatric disorders, and cancer (in the previous 5 years). The recruitment period was between January 2020 and January 2022. Informed consent was obtained from all subjects involved in the study. The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of IRCCS S. De Bellis (protocol code n. 68, 9 April 2019).

#### 2.1.2. Southern-Italy Population Subset (the Salus in Apulia Study)

Participants of the Salus in Apulia population-based study were recruited from the electoral rolls of Castellana Grotte (Bari, Apulia, Southern Italy). The recruiting and evaluation centre was the National Institute of Gastroenterology IRCCS “S. De Bellis” Research Hospital, and the initiative was supported by the Italian Ministry of Health and the Apulia Regional Government. The Salus is an ongoing longitudinal population-based study, activated in 2014, of a representative population of residents in Castellana Grotte (Apulia, southern Italy) who were 65 years of age or older at the time of initial recruitment. While the minimum age of 65 was required for enrolment in the Salus, conversely, the exclusion criteria were lack of mental capacity to express consent, having digestive tract cancers or other malignancies, including dementia and motoneuron diseases, or being under major therapies, which could affect nutritional/physical status. The study design and data collection method are detailed elsewhere [14,15]. Briefly, the entire sampling frame consisted of the 4021 elderly residents in the health registry of the Apulia Region as of 31 December 2014. The study was born as multidisciplinary, including the assessments of the cognitive, sensory, physical, and nutritional domains, as illustrated in some of our previous work [16], and aimed to search for new biological and phenotypic determinants to predict and prevent risky trajectories of aging. Specifically, the data used for the present study came from a subset of the Salus, which included 479 elders who had undergone all examinations required for the purposes of this study. The IRB approved the study of the lead institution, the National Institute of Gastroenterology and Research Hospital “Saverio de Bellis”, and all subjects completed informed consent forms before their evaluation. The study met the principles of the Helsinki Declaration and adhered to the “Standards for Reporting Diagnostic Accuracy Studies” (STARD) guidelines (<http://www.stard-statement.org/>, accessed on 12 January 2023) and the “Strengthening the Reporting of Observational Studies in Epidemiology” (STROBE) guidelines.

### 2.2. Fluid Biomarker Assessment

A blood sample was collected in the morning after overnight fasting to measure the levels of fasting blood glucose (FBG), glycated haemoglobin (HbA1c), total cholesterol, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, and triglycerides, using standard automated enzymatic colorimetric methods (AutoMate 2550, Beckmann Coulter, Brea, CA, USA) under strict quality control. LDL cholesterol was calculated using the Friedewald equation. Plasma glucose was determined using the glucose oxidase method (Sclavus, Siena, Italy). Blood cell count was determined by a Coulter haematology analyser (Beckman–Coulter, Brea, CA, USA). Serum FT3, FT4, and TSH were measured using a competitive photometric method based on the solid-phase antigen-linked technique (LIASON FT3, LIASON FT4, LIASON TSH, Dia-Sorin, Saluggia,

Italy). Serum high-sensitivity C-reactive protein (CRP) was assayed using a latex particle-enhanced immunoturbidimetric assay (Kamiya Biomedical Company, Seattle, WA, USA) (reference range: 0–5.5 mg/L; inter-assay coefficient of variation: 4.5%). Serum 25(OH)D was quantified by a chemiluminescence method (Diasorin Inc., Stillwater, MN, USA), and all samples were analysed in duplicate.

### 2.3. Clinical and Physical Assessment

Height was measured to the nearest 0.5 cm using a wall-mounted stadiometer (Seca 711; Seca, Hamburg, Germany). Body weight was determined at the time of DXA to the nearest 0.1 kg using a calibrated balance beam scale (Seca 711; Seca, Hamburg, Germany). BMI was calculated as weight in kilograms divided by height in metres squared ( $\text{kg}/\text{m}^2$ ). Low physical performance was assessed using the Short Physical Performance Battery (SPPB), an objective tool for measuring the physical performance status of the lower extremities [17]. The SPPB is based on three timed tasks: standing balance, walking speed, and chair sit-to-stand tests [12]. The timed results of each subtest were rescaled according to the predefined cutoff points, obtaining a score ranging from 0 (worst performance) to 12 (best performance). A cutoff value of 8 in the SPPB score was considered to indicate low physical performance, in accordance with both EWGSOP panels [4,18]. Handgrip strength (HGS) was assessed using the Jamar Plus Digital Hand Dynamometer (Patterson Medical, Cedarburg, WI, USA). Seated with arms 90 degrees to the sides, 2 trials were taken per arm in an alternating fashion with 30 s of rest between trials. The highest reading was recorded [4].

Bone mineral density (BMD) and whole-body lean mass were measured using DXA (Discovery WI, Hologic, Inc., Marlborough, MA, USA). The skeletal muscle mass index (SMI) was defined as the sum of the muscle masses of the four limbs as appendicular skeletal muscle mass divided by squared height.

Whole-body lean mass (kg) was taken as the sum of the fat-free, bone-free mass of the arms and legs as lean mass. According to the operational definition by the European Working Group on Sarcopenia in Older People (EWGSOP2) [4], the diagnosis of sarcopenia was given by the presence of both low muscle mass (that is, an  $\text{SMI} < 7.0 \text{ kg}/\text{m}^2$  for males or  $< 5.5 \text{ kg}/\text{m}^2$  for females) [4], and low muscle strength, as defined by a low handgrip strength (HGS), that is, an  $\text{HGS} < 27 \text{ kg}$  for males or  $< 16 \text{ kg}$  for females, or low a physical performance (that is, an  $\text{SPPB} \leq 8$ ).

### 2.4. Statistical Analysis

The entire sample was first divided according to the population setting, i.e., subjects from the Salus in Apulia population-based study and those from the clinical setting of the Santa Margherita Hospital, in order to assess the overlap of the samples (Table 1). Then, the overall population was further subdivided according to the outcome variable, that is, the sarcopenia condition (presence/absence), and groups were compared to describe the clinical and functional differences in terms of frequency and associations (Table 2).

Normal distributions of quantitative variables were tested using the Kolmogorov–Smirnov test. Data are reported as mean  $\pm$  standard deviation ( $M \pm SD$ ) for continuous measures and frequency and percentages (%) for all categorical variables. In order to focus on the practical differences between the groups in terms of effect size (ES) [19], differences between continuous variables, between the groups, were calculated using Wilcoxon's effect size difference between ranks, and their 95% confidence intervals (CI) to assess the magnitude of ES [20]. Prevalence differences were calculated to assess differences between categorical variables.

A machine-learning feature-selection approach, the random forest (RF), was employed to identify the most predictive features in the dataset for the sarcopenic condition considering every possible interaction between them, considering also the non-linear relationships that classical models could not assess (Figure 1). Three RF regression models were built. The first model also included the variables used for the detection of sarcopenia in order

to assess which variables were most important in the algorithm of detection using the Mean Decrease Gini. The variables SMI and SPPB were used within an ensemble learning approach in order to highlight possible relationships even in light of variables known to be associated with the condition of sarcopenia. Such an inclusion is therefore useful in order to highlight possible interactions of a nonparametric nature between predictive factors in the classification of sarcopenia. The second RF regression model was performed subdividing according to the clinical study centre in order to assess if there were differences in the association ranking due to the clinical setting. A third RF regression was performed using only socio-demographic and haematochemical variables in order to assess which variables were most associated with the sarcopenia condition using the Mean Decrease Gini (Figure 2).

**Table 1.** Description of the whole sample according to population setting, i.e., northern (clinical setting) versus southern (population-based setting) Italy.

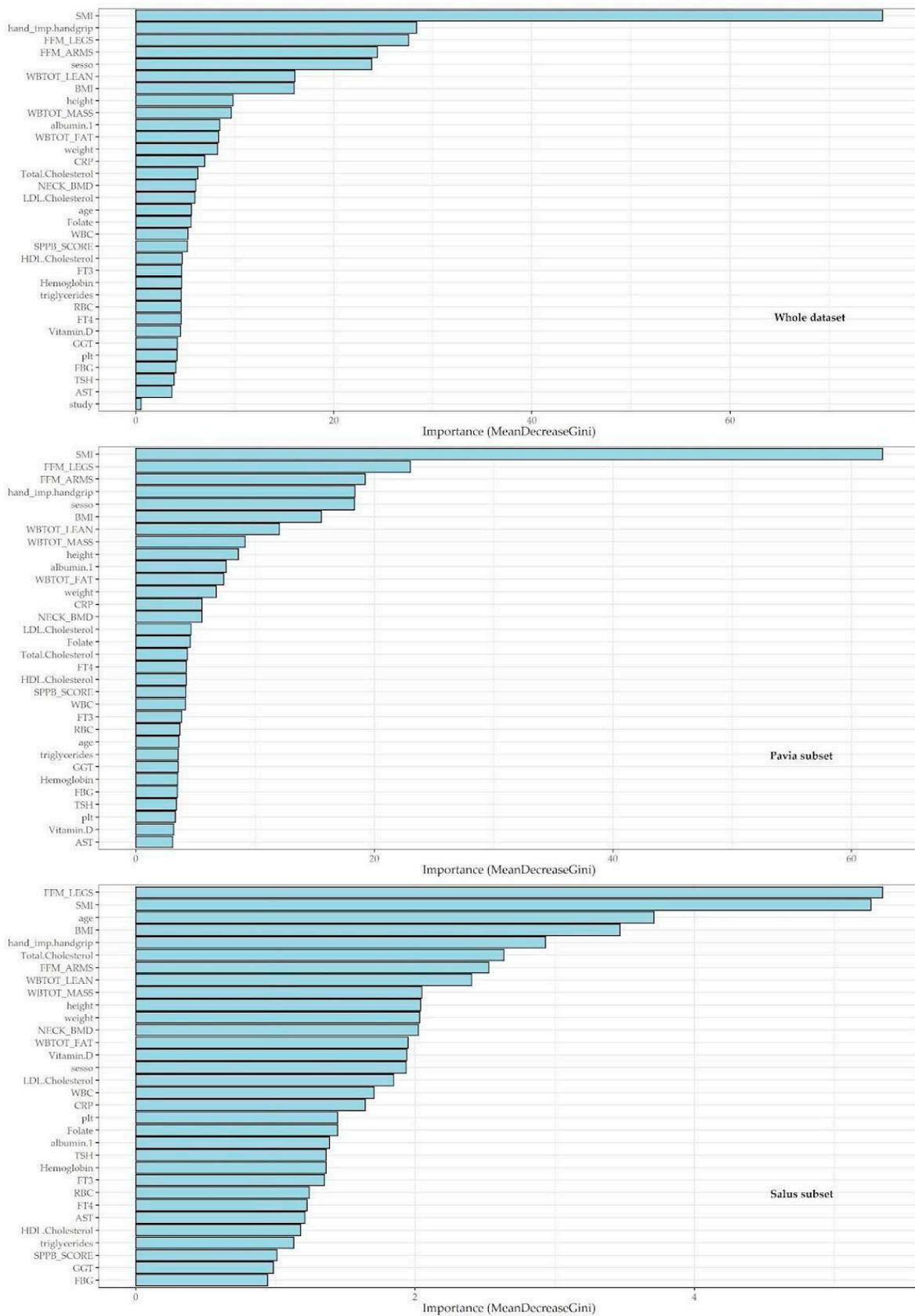
	Clinical-Based Subset (Pavia)		Population-Based Subset (Apulia)		Wilcoxon's ES
	Mean $\pm$ SD	Median (IQR)	Mean $\pm$ SD	Median (IQR)	
Prop. (%)	1312 (73.30)		479 (26.70)		
Age (years)	79.79 $\pm$ 7.18	80 (10)	74.81 $\pm$ 5.67	74.09 (7.58)	<b>0.32 (0.28 to 0.36)</b>
Sex					
Female	949 (72.30)		255 (53.20)		<b>19.10 (14.01 to 24.18)</b>
Male	363 (27.70)		224 (46.80)		
Sarcopenia	165 (12.60)		35 (7.30)		<b>−5.27 (−8.21 to −2.33)</b>
Albumin (g/dL)	3.67 $\pm$ 0.51	3.74 (0.67)	4.08 $\pm$ 0.42	4.05 (0.4)	<b>0.37 (0.34 to 0.41)</b>
AST (U/L)	18.65 $\pm$ 17.7	14 (10)	20.39 $\pm$ 9.33	19 (8)	<b>0.26 (0.22 to 0.30)</b>
BMI (kg/m <sup>2</sup> )	25.57 $\pm$ 5.88	24.7 (6.7)	29.75 $\pm$ 4.94	29.32 (6.08)	<b>0.26 (0.22 to 0.30)</b>
CRP (mg/dL)	1.25 $\pm$ 2.63	0.35 (0.84)	0.43 $\pm$ 0.62	0.43 (0.33)	0.02 (−0.01 to 0.05)
FBG (mg/dL)	108.19 $\pm$ 41.06	96 (33)	102.68 $\pm$ 25.31	96 (21)	0.01 (−0.03 to 0.03)
FFM arms (kg)	4.12 $\pm$ 1.38	3.86 (1.70)	4.71 $\pm$ 1.354	4.51 (2.15)	<b>0.20 (0.16 to 0.24)</b>
FFM legs (kg)	13.419 $\pm$ 3.21	12.98 (4.33)	13.24 $\pm$ 3.02	13.14 (4.71)	0.02 (−0.03 to 0.04)
Folate (ng/mL)	9.13 $\pm$ 9.19	5.9 (7.82)	8.98 $\pm$ 5.95	7.6 (5.3)	<b>0.11 (0.08 to 0.16)</b>
FT3 (pmol/L)	2.32 $\pm$ 0.53	2.3 (0.64)	3.32 $\pm$ 0.39	3.32 (0.46)	<b>0.67 (0.65 to 0.70)</b>
FT4 (pmol/L)	5.76 $\pm$ 5.74	1.46 (9.78)	0.97 $\pm$ 0.66	0.9 (0.19)	<b>0.54 (0.51 to 0.58)</b>
GGT (U/L)	32.98 $\pm$ 42.09	20 (18)	19.44 $\pm$ 16.61	15 (9)	<b>0.21 (0.17 to 0.26)</b>
HGS (kg)	18.68 $\pm$ 7.74	18 (9.33)	23.13 $\pm$ 8.28	22 (11.67)	<b>0.24 (0.20 to 0.29)</b>
HDL Cholesterol (mg/dL)	47.96 $\pm$ 17.65	47 (21)	52.41 $\pm$ 14.02	52 (17)	<b>0.12 (0.08 to 0.17)</b>
Height (cm)	156.79 $\pm$ 9.62	155 (13)	157.45 $\pm$ 9.03	157 (13)	0.04 (−0.01 to 0.09)
Haemoglobin (g/dL)	12.35 $\pm$ 1.72	12.4 (2.2)	13.68 $\pm$ 1.44	13.7 (1.8)	<b>0.35 (0.31 to 0.39)</b>
LDL Cholesterol (mg/dL)	111.76 $\pm$ 37.32	107.9 (48.85)	104.9 $\pm$ 32.66	104 (49)	<b>0.07 (0.20 to 0.12)</b>
Neck BMD	0.74 $\pm$ 0.18	0.73 (0.2)	0.77 $\pm$ 0.58	0.7 (0.17)	<b>0.04 (0.01 to 0.09)</b>
Platelets (103 cells/mm <sup>3</sup> )	247.02 $\pm$ 103.99	238.6 (103.72)	230.55 $\pm$ 71.2	226 (75.5)	<b>0.07 (0.03 to 0.12)</b>
RBC (106 cells/mm <sup>3</sup> )	4.18 $\pm$ 0.65	4.17 (0.69)	4.66 $\pm$ 0.55	4.66 (0.62)	<b>0.38 (0.34 to 0.42)</b>
SMI (kg/m <sup>2</sup> )	7.08 $\pm$ 1.37	7 (1.72)	7.17 $\pm$ 1.18	7.1 (1.75)	0.03 (−0.01 to 0.08)
SPPB score	6.26 $\pm$ 3.04	6 (5)	8.27 $\pm$ 2.86	9 (5)	<b>0.28 (0.25 to 0.33)</b>
Total Cholesterol (mg/dL)	185.68 $\pm$ 41.34	184.5 (58)	180.14 $\pm$ 37.31	179 (50.5)	<b>0.05 (0.01 to 0.10)</b>
Triglycerides (mg/dL)	132.2 $\pm$ 80.08	111 (75.25)	113.56 $\pm$ 75.43	93 (63.5)	<b>0.13 (0.09 to 0.18)</b>
TSH ( $\mu$ U/mL)	2.13 $\pm$ 2.28	1.5 (1.61)	1.87 $\pm$ 1.76	1.55 (1.29)	0.01 (−0.03 to 0.03)
Vitamin D (nmol/L)	13.29 $\pm$ 10.77	10.2 (10.2)	27.98 $\pm$ 12.69	26.4 (12.7)	<b>0.54 (0.51 to 0.58)</b>
WBC (103 cells/mm <sup>3</sup> )	7.01 $\pm$ 3.14	6.54 (2.41)	6.36 $\pm$ 1.81	6.09 (2.13)	<b>0.10 (0.06 to 0.15)</b>
Whole Body Fat (kg)	22.60 $\pm$ 10.46	21.30 (13.26)	30.04 $\pm$ 8.943	28.68 (11.84)	<b>0.34 (0.30 to 0.38)</b>
Whole Body Lean Mass (kg)	38.84 $\pm$ 7.928	37.29 (10.52)	42.77 $\pm$ 8.503	41.61 (13.56)	<b>0.20 (0.16 to 0.25)</b>
Whole Body Mass (kg)	63.43 $\pm$ 15.59	61.70 (20.01)	72.82 $\pm$ 13.93	72.34 (18.93)	<b>0.28 (0.25 to 0.33)</b>
Weight (kg)	62.85 $\pm$ 15.53	60.55 (19.2)	73.86 $\pm$ 14.31	72.6 (18.7)	<b>0.33 (0.30 to 0.37)</b>

*n* = 1791. All data are shown as mean  $\pm$  SD, median (min to max) for continuous variables and as *n* (%) for proportions. Statistically significant data are indicated in bold type. Abbreviations: AST: aspartate aminotransferase; BMI: body mass index; CRP: C-reactive protein; FBG: fasting blood glucose; FFM: fat-free mass; GGT: gamma-glutamyl transferase; HGS: handgrip strength; BMD: bone mineral density; RBC: red blood cells; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery; TSH: thyroid-stimulating hormone.

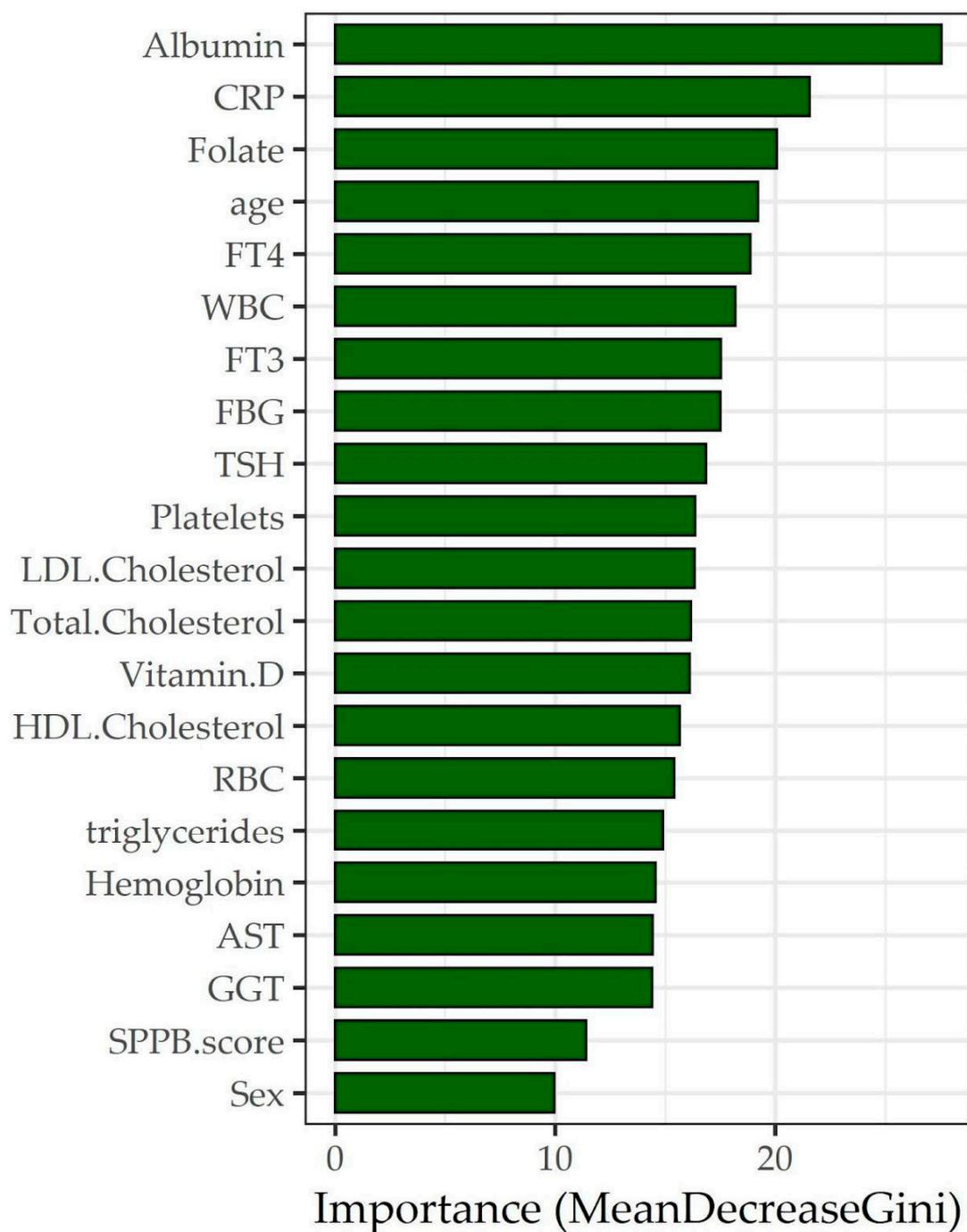
**Table 2.** Description of the whole sample by sarcopenia condition (presence/absence).

	Without Sarcopenia		With Sarcopenia		Wilcoxon's Effect Size
	Mean $\pm$ SD	Median (IQR)	Mean $\pm$ SD	Median (IQR)	
Prop. (%)	1591 (88.80)		200 (11.20)		
Age (years)	78.04 $\pm$ 7.1	78 (10.67)	81.79 $\pm$ 6.76	82.1 (10)	<b>0.16 (0.12 to 0.21)</b>
Sex					
Female	1113 (70.00)		91 (45.50)		<b>24.46 (17.20 to 31.72)</b>
Male	478 (30.00)		109 (54.50)		
Population setting					
Pavia subset	1147 (72.10)		165 (82.50)		−10.41 (−16.12 to −4.70)
Apulia subset	444 (27.90)		35 (17.50)		
Albumin (g/dL)	3.81 $\pm$ 0.5	3.87 (0.57)	3.53 $\pm$ 0.62	3.56 (0.9)	<b>0.15 (0.12 to 0.22)</b>
AST (U/L)	18.98 $\pm$ 15.17	16 (11)	20.15 $\pm$ 20.96	14 (11)	0.03 (−0.01 to 0.08)
BMI (kg/m <sup>2</sup> )	27.35 $\pm$ 5.83	26.64 (7)	21.47 $\pm$ 3.92	21.3 (5.12)	<b>0.33 (0.30 to 0.37)</b>
CRP (mg/dL)	0.9 $\pm$ 2.02	0.34 (0.44)	2.1 $\pm$ 3.76	0.5 (1.92)	<b>0.11 (0.06 to 0.16)</b>
FBG (mg/dL)	106.47 $\pm$ 36.6	96 (29)	108.74 $\pm$ 44.6	93.5 (35)	0.02 (−0.02 to 0.05)
FFM arms (kg)	4.38 $\pm$ 1.38	4.05 (1.91)	3.45 $\pm$ 1.24	3.292 (2.00)	<b>0.19 (0.15 to 0.25)</b>
FFM legs (kg)	13.70 $\pm$ 3.08	13.31 (4.45)	10.73 $\pm$ 2.45	10.52 (3.53)	<b>0.29 (0.25 to 0.33)</b>
Folate (ng/mL)	8.86 $\pm$ 7.9	6.4 (6.8)	10.91 $\pm$ 11.75	6.85 (10.85)	0.03 (−0.01 to 0.08)
FT3 (pmol/L)	2.62 $\pm$ 0.66	2.55 (0.89)	2.32 $\pm$ 0.7	2.32 (0.74)	<b>0.13 (0.09 to 0.18)</b>
FT4 (pmol/L)	4.64 $\pm$ 5.42	1.21 (9.27)	3.14 $\pm$ 4.61	1.18 (0.52)	0.01 (−0.02 to 0.04)
GGT (U/L)	28.72 $\pm$ 35.73	18 (15)	34.5 $\pm$ 49.29	21 (17)	0.04 (−0.01 to 0.09)
HGS (kg)	20.45 $\pm$ 8.19	19 (10.17)	15.26 $\pm$ 5.92	14 (9)	<b>0.20 (0.16 to 0.24)</b>
HDL Cholesterol (mg/dL)	49.66 $\pm$ 16.71	49 (21)	45.11 $\pm$ 17.6	44 (21)	<b>0.08 (0.04 to 0.13)</b>
Height (cm)	156.79 $\pm$ 9.38	156 (13)	158.33 $\pm$ 10.05	158 (15)	0.04 (−0.01 to 0.08)
Haemoglobin (g/dL)	12.76 $\pm$ 1.74	12.9 (2.3)	12.28 $\pm$ 1.8	12.3 (2.35)	<b>0.09 (0.05 to 0.14)</b>
LDL Cholesterol (mg/dL)	110.64 $\pm$ 36.16	107.4 (48.3)	104.2 $\pm$ 36.59	100.2 (52.5)	<b>0.05 (0.01 to 0.10)</b>
Neck BMD	0.76 $\pm$ 0.35	0.72 (0.19)	0.71 $\pm$ 0.22	0.69 (0.24)	<b>0.06 (0.02 to 0.11)</b>
Platelets (103 cells/mm <sup>3</sup> )	242.83 $\pm$ 95.19	237 (91.45)	240.9 $\pm$ 107.23	225.9 (103.33)	0.01 (−0.03 to 0.04)
RBC (106 cells/mm <sup>3</sup> )	4.32 $\pm$ 0.66	4.32 (0.75)	4.17 $\pm$ 0.63	4.18 (0.73)	<b>0.07 (0.03 to 0.12)</b>
SMI (kg/m <sup>2</sup> )	7.29 $\pm$ 1.24	7.19 (1.73)	5.59 $\pm$ 0.86	5.44 (1.32)	<b>0.40 (0.37 to 0.44)</b>
SPPB score	6.94 $\pm$ 3.1	7 (4)	5.62 $\pm$ 3.03	6 (5)	<b>0.13 (0.09 to 0.18)</b>
Total Cholesterol (mg/dL)	185.45 $\pm$ 39.86	184 (55)	174.24 $\pm$ 43.02	169 (64)	<b>0.09 (0.05 to 0.14)</b>
Triglycerides (mg/dL)	127.36 $\pm$ 79.02	106 (74)	126.07 $\pm$ 81.52	100.5 (63)	0.13 (−0.03 to 0.04)
TSH ( $\mu$ U/mL)	2.1 $\pm$ 2.23	1.54 (1.55)	1.74 $\pm$ 1.37	1.42 (1.48)	<b>0.04 (0.01 to 0.08)</b>
Vitamin D (nmol/L)	17.36 $\pm$ 13.01	13.9 (16.85)	16.13 $\pm$ 13.31	10.9 (12.43)	<b>0.04 (0.01 to 0.09)</b>
WBC (103 cells/mm <sup>3</sup> )	6.76 $\pm$ 2.9	6.36 (2.3)	7.46 $\pm$ 2.41	6.98 (2.83)	<b>0.10 (0.06 to 0.15)</b>
Whole Body Fat (kg)	25.55 $\pm$ 10.50	24.49 (14.07)	16.94 $\pm$ 7.91	16.58 (10.77)	<b>0.26 (0.23 to 0.31)</b>
Whole Body Lean Mass (kg)	40.47 $\pm$ 82.44	38.75 (12.35)	35.28 $\pm$ 6.91	34.64 (11.30)	<b>0.19 (0.15 to 0.24)</b>
Whole Body Mass (kg)	67.47 $\pm$ 15.46	66.28 (20.93)	53.78 $\pm$ 12.05	53.64 (16.62)	<b>0.27 (0.24 to 0.32)</b>
Weight (kg)	67.25 $\pm$ 15.83	65.4 (21)	54.23 $\pm$ 11.87	54.3 (17.23)	<b>0.26 (0.22 to 0.30)</b>

$n = 1791$ . All data are shown as mean  $\pm$  SD, median (min to max) for continuous variables and as  $n$  (%) for proportions. Statistically significant data are indicated in bold type. Abbreviations: AST: aspartate aminotransferase; BMI: body mass index; CRP: C-reactive protein; FBG: fasting blood glucose; FFM: fat-free mass; GGT: gamma-glutamyl transferase; HGS: handgrip strength; BMD: bone mineral density; RBC: red blood cells; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery; TSH: thyroid-stimulating hormone.



**Figure 1.** Variable-importance plots for random forest regression models with sarcopenia status as the dependent variable and other collected variables as regressors.



**Figure 2.** Variable-importance plot for the random forest regression model with sarcopenia status as the dependent variable and socio-demographic and haematochemical variables as regressors.

The accuracy of the logistic regression model was calculated using a first confusion matrix, as shown in Table 3. A logistic regression model was performed on the sarcopenia condition as a dependent variable and sociodemographic and blood chemistry parameters as regressors (Table 4) in order to assess differences between the parametric and non-parametric approach in the prediction of sarcopenia. The accuracy of the third RF regression model was calculated using a further confusion matrix, as shown in Table 5.

**Table 3.** Confusion matrix (Salus dataset as test dataset) on sarcopenia status.

		Prediction	
		Without	With
Test dataset (Salus)	Without	444 (94.50)	–
	With	26 (5.50)	9 (100.00)
<b>Accuracy (CI 95%)</b>			
94.57 (92.15 to 96.42)			
<b>Sensitivity</b>			
1.00			
<b>Specificity</b>			
0.94			

**Table 4.** Logistic regression model with sarcopenic condition as the dependent variable and all haematochemical features as regressors.

	OR	Stand. Err.	CI 95%	p-Value
<b>(Intercept)</b>	<b>0.013</b>	<b>1.61</b>	<b>0.001 to 0.301</b>	<b>&lt;0.01</b>
<b>Age (years)</b>	<b>1.054</b>	<b>0.01</b>	<b>1.024 to 1.085</b>	<b>&lt;0.01</b>
<b>Sex (Male)</b>	<b>4.384</b>	<b>0.18</b>	<b>3.027 to 6.351</b>	<b>&lt;0.01</b>
<b>SPPB score</b>	<b>0.906</b>	<b>0.03</b>	<b>0.847 to 0.969</b>	<b>&lt;0.01</b>
WBC (103 cells/mm <sup>3</sup> )	1.051	0.01	0.998 to 1.108	0.07
RBC (106 cells/mm <sup>3</sup> )	1.200	0.16	0.860 to 1.674	0.28
Platelets (103 cells/mm <sup>3</sup> )	1.001	0.01	0.999 to 1.002	0.76
FBG (mg/dL)	0.999	0.01	0.995 to 1.003	0.57
Triglycerides (mg/dL)	1.001	0.01	0.989 to 1.013	0.88
Total Cholesterol (mg/dL)	1.004	0.01	0.950 to 1.062	0.87
HDL Cholesterol (mg/dL)	1.001	0.01	0.943 to 1.062	0.98
LDL Cholesterol (mg/dL)	0.997	0.01	0.943 to 1.053	0.90
<b>TSH (μU/mL)</b>	<b>0.871</b>	<b>0.01</b>	<b>0.784 to 0.969</b>	<b>0.01</b>
<b>FT3 (pmol/L)</b>	<b>0.555</b>	<b>0.17</b>	<b>0.393 to 0.783</b>	<b>&lt;0.01</b>
<b>FT4 (pmol/L)</b>	<b>0.921</b>	<b>0.01</b>	<b>0.887 to 0.956</b>	<b>&lt;0.01</b>
CRP (mg/dL)	1.065	0.01	0.999 to 1.135	0.06
<b>Folate (ng/mL)</b>	<b>1.022</b>	<b>0.01</b>	<b>1.005 to 1.038</b>	<b>&lt;0.01</b>
<b>Vitamin D (nmol/L)</b>	<b>1.015</b>	<b>0.01</b>	<b>1.002 to 1.028</b>	<b>0.01</b>
Haemoglobin (g/dL)	1.007	0.01	0.873 to 1.16	0.92
GGT (U/L)	0.999	0.01	0.995 to 1.004	0.73
AST (U/L)	0.999	0.01	0.989 to 1.011	0.99
<b>Albumin (g/dL)</b>	<b>0.562</b>	<b>0.19</b>	<b>0.386 to 0.818</b>	<b>&lt;0.01</b>

Abbreviations: AST: aspartate aminotransferase; BMI: body mass index; CRP: C-reactive protein; FBG: fasting blood glucose; FFM: fat-free mass; GGT: gamma-glutamyl transferase; HGS: handgrip strength; RBC: red blood cells; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery; TSH: thyroid-stimulating hormone.

**Table 5.** Logistic regression model confusion matrix.

		Reference	
		No	Yes
Prediction	No	1581	171
	Yes	10	29
Accuracy	89.89		
Sensitivity	14.50		
Specificity	99.37		

### 3. Results

A total of 1791 subjects made up the entire sample; of these, the majority ( $n = 1312$ , 73.3%) came from the clinical setting of northern Italy. Age over 65 years was a common feature of both samples, and the mean age ( $\pm$ standard deviation, SD) was  $79.79 \pm 7.18$  years and  $74.81 \pm 5.67$  years for the northern and southern Italian populations, respectively.

Table 1 shows a description of the entire sample according to the population setting, i.e., northern (clinical setting) versus southern (population-based setting) Italy. Here, when analysing the between-group practical differences in terms of effect size (ES), as defined by Wilcoxon's effect size difference and respective 95% confidence intervals (CI), the age (ES: 0.32, 95%CI 0.28–0.36) and sex (ES: 19.10, 95%CI 14.01–24.18) showed significant differences, meaning that the northern Italian sample was older and included more females than the southern counterpart. There was a significantly higher prevalence of sarcopenia in the northern sample (12.6% vs. 7.3%) compared to the southern sample (ES:  $-5.27$ , 95%CI  $-8.21$  to  $-2.33$ ). Along these lines, moving toward DXA-derived body composition variables, significant differences showed up for arm free-fat mass (FFM), whole-body lean mass, and whole-body fat mass. That is, the southern sample showed better values for FFM (ES: 0.20, 95%CI 0.16 to 0.24), whole-body lean mass (ES: 0.20, 95%CI 0.16 to 0.25), and fat mass (ES: 0.34, 95%CI 0.30 to 0.38) than the northern counterpart. Moreover, body weight (ES: 0.33, 95%CI 0.30 to 0.37) and BMI (ES: 0.26, 95%CI 0.22 to 0.30) were also significantly higher in the Apulian sample. Physical function variables were overall better in the population-based southern sample, as indicated by the handgrip strength (ES: 0.24, 95%CI 0.20 to 0.29) and SPPB score (ES: 0.28, 95%CI 0.25 to 0.33). Also, femoral neck BMD followed the trend between the two samples (ES: 0.04, 95%CI 0.01 to 0.09). As for the fluid biomarker pattern, significantly better levels of albumin (ES: 0.37, 95%CI 0.34 to 0.41), folate (ES: 0.11, 95%CI 0.08 to 0.16), vitamin D (ES: 0.54, 95%CI 0.51 to 0.58), haemoglobin (ES: 0.35, 95%CI 0.31 to 0.39), RBC (ES: 0.38, 95%CI 0.34 to 0.42), triglycerides (ES: 0.13, 95%CI 0.09 to 0.18), and HDL cholesterol (ES: 0.12, 95%CI 0.08 to 0.17) emerged for the southern population compared with the northern counterpart.

Following the same analytical approach, Table 2 shows a description of the whole sample by sarcopenia condition (presence/absence). To substantiate the internal validity of the findings, sarcopenic subjects were predominantly older (ES: 0.16, 95%CI 0.12 to 0.21) and male (ES: 24.46, 95%CI 17.20 to 31.72). Our findings indicated significantly lower values of albumin (ES: 0.15, 95%CI 0.12 to 0.22), BMI (ES: 0.33, 95%CI 0.30 to 0.37), FFM of the arms (ES: 0.19, 95%CI 0.15 to 0.25), FFM of the legs (ES: 0.29, 95%CI 0.25 to 0.33), FT3 (ES: 0.13, 95%CI 0.09 to 0.18), handgrip strength (ES: 0.20, 95%CI 0.16 to 0.24), HDL cholesterol (ES: 0.08, 95%CI 0.04 to 0.13), haemoglobin (ES: 0.09, 95%CI 0.05 to 0.14), LDL cholesterol (ES: 0.05, 95%CI 0.01 to 0.10), femoral neck BMD (ES: 0.06, 95%CI 0.02 to 0.11), RBC (ES: 0.07, 95%CI 0.03 to 0.12), SMI (ES: 0.40, 95%CI 0.37 to 0.44), TSH (ES: 0.04, 95%CI 0.01 to 0.08), vitamin D (ES: 0.04, 95%CI 0.01 to 0.09), whole-body fat mass (ES: 0.26, 95%CI 0.23 to 0.31), whole-body lean mass (ES: 0.19, 95%CI 0.15 to 0.24), and body weight (ES: 0.26, 95%CI 0.22 to 0.30) for the sarcopenic population compared with the counterpart. Conversely, the

sarcopenic sample showed significantly higher values of CRP (ES: 0.11, 95%CI 0.06 to 0.16) and WBC (ES: 0.10, 95%CI (0.06 to 0.15) than the non-sarcopenic counterpart.

Figure 1 shows a plot of important variables from the RF regression model with sarcopenia condition as the dependent variable and the other variables as regressors. The following rationale guided the selection procedure. The first model was run including both the fluid markers and the functional domains of sarcopenia to evaluate which variables were the most important to be included in the algorithm of sarcopenia detection by using the Mean Decrease Gini. The graph showed that SMI (Mean Decrease Gini greater than 60), followed by handgrip strength, FFM (arms and legs), sex (all showing a 20 to 40 Mean Decrease Gini), BMI, and total lean body mass (Mean Decrease Gini over 15) were the most relevant. The second RF regression model was run separately after dividing the sample into two subsets according to the population setting (northern vs. southern Italy, i.e., clinical vs. population-based setting) to assess whether there were differences in the importance ranking due to the different settings. Here, the ranking showed almost overlapping findings, that is, SMI, handgrip strength, FFM (legs and arms), and sex being the top variables of importance in both subsets. A third RF regression was performed using only the socio-demographic and haematochemical variables to assess which variables were most associated with the condition of sarcopenia in the total sample. Here, albumin, CRP, and folate were shown to be top-ranked (Mean Decrease Gini: 27.5, 21.56, and 20.08, respectively) (Figure 2). The reliability of the third RF regression model was evaluated using a confusion matrix (Table 3). The findings showed an accuracy of 94.57 (95%CI 92.15 to 96.42), a sensitivity of 1.00, and a specificity of 0.94.

To explore the differences between the parametric and nonparametric approaches in predicting sarcopenia, a logistic regression model was run on sarcopenia condition as the dependent variable and sociodemographic and fluid biomarkers variables as regressors (Table 4). Male gender showed the strongest association with sarcopenia (OR: 4.384, 95%CI 3.027 to 6.351). Slightly at risk were those subjects showing lower SPPB scores (OR: 0.906, 95%CI 0.847 to 0.969), serum folate (OR: 1.022, 95%CI 1.005 to 1.038), and vitamin D (OR: 1.015, 95%CI 1.002 to 1.028). Again, the reliability of the logistic regression model was evaluated using a confounding matrix (Table 5). The findings showed an accuracy of 89.89 (95%CI 87.70 to 90.62), a sensitivity of 14.50, and a specificity of 99.37.

#### 4. Discussion

This study aimed to investigate clinical and fluid markers most associated with the condition of sarcopenia by implementing a machine-learning-based approach across two subsets (clinical and population-based) of adults over 65 years of age to provide evidence on how to better set up preventive strategies in sarcopenia settings. The research was carried out involving populations from northern and southern Italy, respectively from the Santa Margherita Clinic (Pavia, northern Italy) and the Salus in Apulia population-based study (Apulia, southern Italy). The top-ranked association variables for sarcopenia in RF selection overlapped across the two population subsets, and included SMI, HGS, FFM of legs and arms, and sex. Given these similarities in terms of clinical features, we implemented parametric and nonparametric analysis on the whole sample to investigate those variables most associated with sarcopenia, and found albumin, CRP, folate, and age to rank high in RF selection, while sex, folate, and vitamin D were most relevant in logistic.

First, to substantiate the internal validity of our findings, it is worth noting that statistical analyses by sarcopenia condition (presence/absence) showed a higher proportion of males as well as older age, and lower levels of albumin, haemoglobin, vitamin D, and BMI in the sarcopenic population. In terms of comparison, our clustered analyses by population setting showed meaningful differences between subsets in the prevalence of sarcopenia (higher in the clinical setting population, i.e., 12% compared to 7% of the population-based counterpart) and, in turn, fluid biological markers and functional proxies. The clinical setting population showed lower (worse) values in FFM legs and arms, HGS, SPPB, and BMI, as well as lower serum levels of albumin, folate, vitamin D, haemoglobin, RBC, triglycerides,

and HDL cholesterol. Functional dimensions such as HGS and SPPB scoring were also worse in the clinical subset. These data fit well within a geriatric outpatient/clinical healthcare setting, as the latter assumes that subjects come to the hospital to fix some medical issue, whereas in a population-based setting, subjects are usually recruited to participate in data collection for research purposes because they fit the purposes of the study well, and thus come to the hospital for a visit without a specific health concern to fix. In our specific case, subjects from the north were recruited in a clinical setting that included both outpatients and inpatients, and were therefore expected to be more frail and physically impaired than the general population recruited in the southern sample. This view most likely explains the poorer general health status of the clinical subset. Moreover, the physiological path of reduction in testosterone, the hormone that drives protein synthesis and muscle development, may then elucidate the higher male incidence of sarcopenia already acknowledged by the scientific community [21]. Basically, testosterone works as the fuel that powers muscle building.

The RF plot of important variables for sarcopenia algorithm showed SMI, HGS, FFM (arms and legs), sex, and BMI as the most relevant. Due to the predominance of skeletal muscle in the arms and legs, is not surprising that both lean soft tissues are key and are actually embedded in the SMI as muscle proxies in the consensus panel. For BMI, the role in relation to sarcopenia can be explained in both excess and deficit values. On the one hand, weight gain and obesity can speed up the onset and the progression of sarcopenia directly or indirectly. For example, in fat tissue of obesity phenotypes [22], the accumulation of pro-inflammatory macrophages and other immune cells, as well as the dysregulated production of various adipokines together with cytokines released by immune cells, create a prolonged local pro-inflammatory state [23,24]. In addition, overproduction and impaired lipid storage capacity is a feature of adipose tissue in obesity, which accumulates ectopically in skeletal muscle. These intramuscular lipids and their products can result in mitochondrial dysfunction and increased secretion of certain pro-inflammatory myokines with the potential to induce muscle dysfunction. On the other hand, reduced BMI may be equally troublesome as it would mean a reduced fat mass, that is believed to be an energy reserve in older people and helps them survive disease and chronic conditions [25]. It has also been hypothesized that individuals with higher fat mass may have higher protein intake, which is a protective factor against sarcopenia [26]. Therefore, maintaining a healthy weight is important for older adults to preserve muscle mass and strength.

In a further RF plot of significance, we found overlapping top-ranked association variables across the two population settings (i.e., SMI, HGS, FFM of legs and arms, and sex) showing how the clinical and population-based subsets were not much divergent. Therefore, an additional RF plot was built without splitting the two samples to get a selection of variables most associated with sarcopenia, and found that albumin, CRP, and folate ranked high. A further comparative logistic analysis found sex, folate, and vitamin D among the most closely associated, although the latter lacked significance. These findings are discussed below.

It is well-acknowledged that serum albumin concentration may be an indicator of individual nutritional status [27,28], with lower values indicating a decrease in protein reserve, stimulating catabolic processes that lead to muscle breakdown [29]. There is also a body of research indicating the antioxidant properties of albumin, showing that albumin is a specific modulator of cellular glutathione, one of the body's major antioxidants [30]. Oxidative damage may play a crucial role in skeletal muscle decline with aging. Furthermore, increased concentrations of free cortisol have been observed in hypoalbuminemic individuals, and this other biological pathway potentially stimulates muscle breakdown, especially in inactive people. Albumin has also been shown to activate the phosphatidyl-inositol 3-kinase pathway, thus mediating muscle breakdown.

For serum folate, our findings are in line with other studies revealing serum folate levels significantly correlated with reduced lower limb strength and grip in the elderly over

65, especially in women [31]. Findings from other research suggest that folate deficiency is associated with a decline in muscle strength, and that a reduced dietary micronutrient intake [32,33], including folate, has an important impact on muscle health, as indicated by decreased ability to generate strength and endurance as well as reduced physical activity. We therefore suggest that the effect of folate deficiency on measures of strength and physical performance involves biological mechanisms that could include specific folate activities in addition to the homocysteine pathway, such as neurotransmitter synthesis, myelination, DNA and protein synthesis, DNA methylation, and epigenetic regulation [34].

For vitamin D, epidemiological studies showed clinical correlations between vitamin D deficiency and sarcopenia, such that this lipoprotein is attracting more and more attention among the scientific community. In the general population, low serum vitamin D concentration has been significantly found to be associated with a higher prevalence of sarcopenia and loss of physical performance, such as walking speed [35,36]. Several meta-analyses and RCT studies have also demonstrated the positive effects of vitamin D supplementation, such as improved overall muscle strength (particularly of lower limb muscles) and decreased standing time [37–39].

Last, regarding the inflammatory pathway, it has long been hypothesized that higher levels of inflammatory markers play a role in functional decline in aging populations [40,41]. Cross-sectional and longitudinal studies consistently demonstrated associations between high levels of interleukins, especially interleukin-6 (IL-6) [42], or CRP and poor physical performance and disability. The causal pathway running from inflammation to disability is suggested to involve catabolic effects of inflammatory markers on muscles. To further substantiate our findings, there is evidence that in older men and women, high levels of CRP are associated with a 2- to 3-fold increased risk of losing more than 40% of muscle strength over 3 years [43].

When comparing sarcopenia to other studies, we can look at what has been found in other reports. For example, among elderly people aged  $\geq 65$  years admitted to daycare centres in Taiwan [44], calf circumference, Mini Nutritional Assessment, dementia, and BMI were factors associated with sarcopenia, diagnosed, however, using Asian Working Group for Sarcopenia (AWGS) criteria. In another report using data from elderly people in Brazil and the EWGSOP criteria for sarcopenia diagnosis [45], older age, cognitive impairment, lower income, smoking, and undernutrition and undernutrition risk were factors associated with sarcopenia. In a report that used machine learning like ours as a data analysis approach [46], the top risk factors in men were BMI, red blood cell count (RBC), blood urea nitrogen (BUN), vitamin D, ferritin, fibre intake (g/d), primary diastolic blood pressure, white blood cell count (WBC), fat intake (g/d), age, glutamic-pyruvic transaminase, niacin intake (mg/d), protein intake (g/d), fasting blood glucose, and water intake (g/d). Inevitably, the most important risk factors in women were BMI, water intake (g/d), WBC, RBC count, iron intake (mg/d), BUN, high-density lipoprotein, protein intake (g/d), fibre consumption (g/d), vitamin C (mg/d), parathyroid hormone, niacin intake (mg/d), carotene intake ( $\mu\text{g}/\text{d}$ ), potassium intake (mg/d), calcium intake (mg/d), sodium intake (mg/d), retinol intake ( $\mu\text{g}/\text{d}$ ), and age. The population setting, study setting, exposome, and diagnostic criteria used for diagnosis certainly create heterogeneity in the findings, and making it unfeasible to draw consistent conclusions.

#### *Strengths and Limitations*

The strengths of this study include the large sample size, the multiple anthropometric and clinico-metabolic variables collected and correlated with sarcopenia, the use of DXA as the gold standard for body composition and thus for reliability of findings, and the employment of robust statistical algorithms and evidence-based references. Further, double recruitment and data collection on two different populations lacks any previous research to compare to in Italy. However, some limitations should be acknowledged. The cross-sectional nature precludes causal inference on outcomes, and although comprehensive, the database lacked a broader assessment of fluid markers. Moreover, by skimming the dataset

for variables shared by the two population samples, a more comprehensive assessment of the biological markers panel was not allowed.

## 5. Conclusions

Albumin, CRP, vitamin D, and serum folate should not be overlooked when screening for sarcopenia in geriatric settings, particularly in the male population. Improving the health burden and quality of life of the aging population is urgently needed. Employing multidimensional methodology to model risk management pathways may provide a way to stratify the risk of sarcopenia in preventive medicine settings, and thus ease the identification of a deteriorating health status in the aged population.

**Author Contributions:** Conceptualization, C.G., F.C. (Fabio Castellana), L.L., and S.P.; methodology, R.S.; software, F.C. (Fabio Castellana) validation, M.L.C., R.S., and M.R.; formal analysis, F.C. (Fabio Castellana); investigation, F.C. (Feliciano Catino), C.G., F.C. (Fabio Castellana), L.L., and S.P.; resources, R.S.; data curation, F.C. (Fabio Castellana); writing—original draft preparation, R.Z. and A.M.; writing—review & editing, R.Z. and A.M.; visualization, M.L.C., R.S., and M.R.; supervision, M.L.C. and M.R.; project administration, R.S.; funding acquisition, M.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of IRCCS S. De Bellis (protocol code n. 68, 9 April 2019).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study. Written informed consent was obtained from the patients to publish this paper.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author Fabio Castellana. The data are not publicly available due to privacy reasons.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Chang, A.Y.; Skirbekk, V.F.; Tyrovolas, S.; Kassebaum, N.J.; Dieleman, J.L. Measuring Population Ageing: An Analysis of the Global Burden of Disease Study 2017. *Lancet Public Health* **2019**, *4*, e159–e167. [[CrossRef](#)]
2. The Lancet Healthy Longevity. Care for Ageing Populations Globally. *Lancet Healthy Longev.* **2021**, *2*, e180. [[CrossRef](#)]
3. Zhang, X.-M.; Chen, D.; Xie, X.-H.; Zhang, J.-E.; Zeng, Y.; Cheng, A.S. Sarcopenia as a Predictor of Mortality among the Critically Ill in an Intensive Care Unit: A Systematic Review and Meta-Analysis. *BMC Geriatr.* **2021**, *21*, 339. [[CrossRef](#)] [[PubMed](#)]
4. Cruz-Jentoft, A.J.; Bahat, G.; Bauer, J.; Boirie, Y.; Bruyère, O.; Cederholm, T.; Cooper, C.; Landi, F.; Rolland, Y.; Sayer, A.A.; et al. Sarcopenia: Revised European Consensus on Definition and Diagnosis. *Age Ageing* **2019**, *48*, 16–31. [[CrossRef](#)]
5. Beaudart, C.; McCloskey, E.; Bruyère, O.; Cesari, M.; Rolland, Y.; Rizzoli, R.; Araujo de Carvalho, I.; Amuthavalli Thiyagarajan, J.; Bautmans, I.; Bertiè, M.-C.; et al. Sarcopenia in Daily Practice: Assessment and Management. *BMC Geriatr.* **2016**, *16*, 170. [[CrossRef](#)] [[PubMed](#)]
6. Guglielmi, G.; Ponti, F.; Agostini, M.; Amadori, M.; Battista, G.; Bazzocchi, A. The Role of DXA in Sarcopenia. *Aging Clin. Exp. Res.* **2016**, *28*, 1047–1060. [[CrossRef](#)]
7. Curcio, F.; Ferro, G.; Basile, C.; Liguori, I.; Parrella, P.; Pirozzi, F.; Della-Morte, D.; Gargiulo, G.; Testa, G.; Tocchetti, C.G.; et al. Biomarkers in Sarcopenia: A Multifactorial Approach. *Exp. Gerontol.* **2016**, *85*, 1–8. [[CrossRef](#)]
8. Papadopoulou, S.K.; Voulgaridou, G.; Kondyli, F.S.; Drakaki, M.; Sianidou, K.; Andrianopoulou, R.; Rodopaios, N.; Pritsa, A. Nutritional and Nutrition-Related Biomarkers as Prognostic Factors of Sarcopenia, and Their Role in Disease Progression. *Diseases* **2022**, *10*, 42. [[CrossRef](#)] [[PubMed](#)]
9. Chen, L.-K.; Arai, H.; Assantachai, P.; Akishita, M.; Chew, S.T.H.; Dumlao, L.C.; Duque, G.; Woo, J. Roles of Nutrition in Muscle Health of Community-Dwelling Older Adults: Evidence-Based Expert Consensus from Asian Working Group for Sarcopenia. *J. Cachexia Sarcopenia Muscle* **2022**, *13*, 1653–1672. [[CrossRef](#)]
10. Wang, Y.; Gu, Y.; Huang, J.; Wu, H.; Meng, G.; Zhang, Q.; Liu, L.; Zhang, S.; Wang, X.; Zhang, J.; et al. Serum Vitamin D Status and Circulating Irisin Levels in Older Adults with Sarcopenia. *Front. Nutr.* **2022**, *9*, 1051870. [[CrossRef](#)] [[PubMed](#)]
11. Cesari, M.; Kuchel, G.A. Role of Sarcopenia Definition and Diagnosis in Clinical Care: Moving from Risk Assessment to Mechanism-Guided Interventions. *J. Am. Geriatr. Soc.* **2020**, *68*, 1406–1409. [[CrossRef](#)]

12. Lampignano, L.; Bortone, I.; Castellana, F.; Donghia, R.; Guerra, V.; Zupo, R.; de Pergola, G.; di Masi, M.; Giannelli, G.; Lozupone, M.; et al. Impact of Different Operational Definitions of Sarcopenia on Prevalence in a Population-Based Sample: The Salus in Apulia Study. *Int. J. Environ. Res. Public Health* **2021**, *18*, 12979. [[CrossRef](#)]
13. NCD Risk Factor Collaboration (NCD-RisC). Trends in Adult Body-Mass Index in 200 Countries from 1975 to 2014: A Pooled Analysis of 1698 Population-Based Measurement Studies with 19.2 Million Participants. *Lancet* **2016**, *387*, 1377–1396. [[CrossRef](#)] [[PubMed](#)]
14. Zupo, R.; Castellana, F.; Donghia, R.; Lampignano, L.; Guerra, V.; de Pergola, G.; Lozupone, M.; Bortone, I.; de Nucci, S.; Tatoli, R.; et al. Liver Frailty and All-Cause Mortality in the Older Participants of the Salus in Apulia Study. *Geroscience* **2022**, *44*, 835–845. [[CrossRef](#)] [[PubMed](#)]
15. Zupo, R.; Sardone, R.; Donghia, R.; Castellana, F.; Lampignano, L.; Bortone, I.; Misciagna, G.; de Pergola, G.; Panza, F.; Lozupone, M.; et al. Traditional Dietary Patterns and Risk of Mortality in a Longitudinal Cohort of the Salus in Apulia Study. *Nutrients* **2020**, *12*, 1070. [[CrossRef](#)]
16. Sardone, R.; Castellana, F.; Bortone, I.; Lampignano, L.; Zupo, R.; Lozupone, M.; Griseta, C.; Dibello, V.; Seripa, D.; Guerra, V.; et al. Association between Central and Peripheral Age-Related Hearing Loss and Different Frailty Phenotypes in an Older Population in Southern Italy. *JAMA Otolaryngol. Head Neck Surg.* **2021**, *147*, 561–571. [[CrossRef](#)]
17. Guralnik, J.M.; Simonsick, E.M.; Ferrucci, L.; Glynn, R.J.; Berkman, L.F.; Blazer, D.G.; Scherr, P.A.; Wallace, R.B. A Short Physical Performance Battery Assessing Lower Extremity Function: Association with Self-Reported Disability and Prediction of Mortality and Nursing Home Admission. *J. Gerontol.* **1994**, *49*, M85–M94. [[CrossRef](#)] [[PubMed](#)]
18. 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.; 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. [[CrossRef](#)] [[PubMed](#)]
19. Ellis, P.D. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*; Cambridge University Press: Cambridge, UK, 2010; ISBN 9781139488150.
20. Grissom, R.J.; Kim, J.J. *Effect Sizes for Research: A Broad Practical Approach*; Lawrence Erlbaum Associates Publishers: Mahwah, NJ, USA, 2005; Volume 253.
21. Shigehara, K.; Kato, Y.; Izumi, K.; Mizokami, A. Relationship between Testosterone and Sarcopenia in Older-Adult Men: A Narrative Review. *J. Clin. Med. Res.* **2022**, *11*, 6202. [[CrossRef](#)]
22. Donini, L.M.; Busetto, L.; Bischoff, S.C.; Cederholm, T.; Ballesteros-Pomar, M.D.; Batsis, J.A.; Bauer, J.M.; Boirie, Y.; Cruz-Jentoft, A.J.; Dicker, D.; et al. Definition and Diagnostic Criteria for Sarcopenic Obesity: ESPEN and EASO Consensus Statement. *Obes. Facts* **2022**, *15*, 321–335. [[CrossRef](#)]
23. Zamboni, M.; Rubele, S.; Rossi, A.P. Sarcopenia and Obesity. *Curr. Opin. Clin. Nutr. Metab. Care* **2019**, *22*, 13–19. [[CrossRef](#)] [[PubMed](#)]
24. Perna, S.; Peroni, G.; Faliva, M.A.; Bartolo, A.; Naso, M.; Miccono, A.; Rondanelli, M. Sarcopenia and Sarcopenic Obesity in Comparison: Prevalence, Metabolic Profile, and Key Differences. A Cross-Sectional Study in Italian Hospitalized Elderly. *Aging Clin. Exp. Res.* **2017**, *29*, 1249–1258. [[CrossRef](#)] [[PubMed](#)]
25. Yu, R.; Wong, M.; Leung, J.; Lee, J.; Auyeung, T.W.; Woo, J. Incidence, Reversibility, Risk Factors and the Protective Effect of High Body Mass Index against Sarcopenia in Community-Dwelling Older Chinese Adults. *Geriatr. Gerontol. Int.* **2014**, *14* (Suppl. S1), 15–28. [[CrossRef](#)]
26. Yoo, M.C.; Won, C.W.; Soh, Y. Association of High Body Mass Index, Waist Circumference, and Body Fat Percentage with Sarcopenia in Older Women. *BMC Geriatr.* **2022**, *22*, 937. [[CrossRef](#)]
27. Keller, U. Nutritional Laboratory Markers in Malnutrition. *J. Clin. Med. Res.* **2019**, *8*, 775. [[CrossRef](#)]
28. Pergola, G.D.; de Pergola, G.; Zupo, R.; Lampignano, L.; Paradiso, S.; Murro, I.; Bartolomeo, N.; Cecere, A.; Ciccone, M.M.; Giannelli, G.; et al. Effects of a Low Carb Diet and Whey Proteins on Anthropometric, Hematochemical and Cardiovascular Parameters in Subjects with Obesity. *Endocr. Metab. Immune Disord. Drug Targets* **2020**, *20*, 1719–1725. [[CrossRef](#)]
29. Visser, M.; Kritchevsky, S.B.; Newman, A.B.; Goodpaster, B.H.; Tylavsky, F.A.; Nevitt, M.C.; Harris, T.B. Lower Serum Albumin Concentration and Change in Muscle Mass: The Health, Aging and Body Composition Study. *Am. J. Clin. Nutr.* **2005**, *82*, 531–537. [[PubMed](#)]
30. Cantin, A.M.; Paquette, B.; Richter, M.; Larivée, P. Albumin-Mediated Regulation of Cellular Glutathione and Nuclear Factor Kappa B Activation. *Am. J. Respir. Crit. Care Med.* **2000**, *162*, 1539–1546. [[CrossRef](#)] [[PubMed](#)]
31. Wee, A.K.H. Serum Folate Predicts Muscle Strength: A Pilot Cross-Sectional Study of the Association between Serum Vitamin Levels and Muscle Strength and Gait Measures in Patients >65 Years Old with Diabetes Mellitus in a Primary Care Setting. *Nutr. J.* **2016**, *15*, 89. [[CrossRef](#)]
32. Zupo, R.; Castellana, F.; de Nucci, S.; Sila, A.; Aresta, S.; Buscemi, C.; Randazzo, C.; Buscemi, S.; Triggiani, V.; de Pergola, G.; et al. Role of Dietary Carotenoids in Frailty Syndrome: A Systematic Review. *Biomedicine* **2022**, *10*, 632. [[CrossRef](#)]
33. Kochlik, B.; Stuetz, W.; Pérès, K.; Pilleron, S.; Féart, C.; García García, F.J.; Bandinelli, S.; Gomez-Cabrero, D.; Rodriguez-Mañas, L.; Grune, T.; et al. Associations of Fat-Soluble Micronutrients and Redox Biomarkers with Frailty Status in the FRAILOMIC Initiative. *J. Cachexia Sarcopenia Muscle* **2019**, *10*, 1339–1346. [[CrossRef](#)]
34. Miller, A.L. The Methylation, Neurotransmitter, and Antioxidant Connections between Folate and Depression. *Altern. Med. Rev.* **2008**, *13*, 216–226.

35. Wang, J.; Wang, X.; Gu, Y.; Liu, M.; Chi, V.T.Q.; Zhang, Q.; Liu, L.; Meng, G.; Yao, Z.; Wu, H.; et al. Vitamin D Is Related to Handgrip Strength in Adult Men Aged 50 Years and over: A Population Study from the TCLSIH Cohort Study. *Clin. Endocrinol.* **2019**, *90*, 753–765. [[CrossRef](#)] [[PubMed](#)]
36. Wicherts, I.S.; van Schoor, N.M.; Boeke, A.J.P.; Visser, M.; Deeg, D.J.H.; Smit, J.; Knol, D.L.; Lips, P. Vitamin D Status Predicts Physical Performance and Its Decline in Older Persons. *J. Clin. Endocrinol. Metab.* **2007**, *92*, 2058–2065. [[CrossRef](#)]
37. Beaudart, C.; Buckinx, F.; Rabenda, V.; Gillain, S.; Cavalier, E.; Slomian, J.; Petermans, J.; Reginster, J.-Y.; Bruyère, O. The Effects of Vitamin D on Skeletal Muscle Strength, Muscle Mass, and Muscle Power: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *J. Clin. Endocrinol. Metab.* **2014**, *99*, 4336–4345. [[CrossRef](#)]
38. Gkekas, N.K.; Anagnostis, P.; Paraschou, V.; Stamiris, D.; Dellis, S.; Kenanidis, E.; Potoupnis, M.; Tsiridis, E.; Goulis, D.G. The Effect of Vitamin D Plus Protein Supplementation on Sarcopenia: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Maturitas* **2021**, *145*, 56–63. [[CrossRef](#)] [[PubMed](#)]
39. Rondanelli, M.; Faliva, M.; Monteferrario, F.; Peroni, G.; Repaci, E.; Allieri, F.; Perna, S. Novel Insights on Nutrient Management of Sarcopenia in Elderly. *BioMed Res. Int.* **2015**, *2015*, 524948. [[CrossRef](#)]
40. Cesari, M.; Penninx, B.W.J.; Pahor, M.; Lauretani, F.; Corsi, A.M.; Williams, G.R.; Guralnik, J.M.; Ferrucci, L. Inflammatory Markers and Physical Performance in Older Persons: The InCHIANTI Study. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **2004**, *59*, M242–M248. [[CrossRef](#)] [[PubMed](#)]
41. Bano, G.; Trevisan, C.; Carraro, S.; Solmi, M.; Luchini, C.; Stubbs, B.; Manzano, E.; Sergi, G.; Veronese, N. Inflammation and Sarcopenia: A Systematic Review and Meta-Analysis. *Maturitas* **2017**, *96*, 10–15. [[CrossRef](#)] [[PubMed](#)]
42. Ferrucci, L.; Harris, T.B.; Guralnik, J.M.; Tracy, R.P.; Corti, M.C.; Cohen, H.J.; Penninx, B.; Pahor, M.; Wallace, R.; Havlik, R.J. Serum IL-6 Level and the Development of Disability in Older Persons. *J. Am. Geriatr. Soc.* **1999**, *47*, 639–646. [[CrossRef](#)]
43. Schaap, L.A.; Pluijm, S.M.F.; Deeg, D.J.H.; Visser, M. Inflammatory Markers and Loss of Muscle Mass (sarcopenia) and Strength. *Am. J. Med.* **2006**, *119*, 526.e9–526.e17. [[CrossRef](#)] [[PubMed](#)]
44. Chang, C.-F.; Yeh, Y.-L.; Chang, H.-Y.; Tsai, S.-H.; Wang, J.-Y. Prevalence and Risk Factors of Sarcopenia among Older Adults Aged  $\geq 65$  Years Admitted to Daycare Centers of Taiwan: Using AWGS 2019 Guidelines. *Int. J. Environ. Res. Public Health* **2021**, *18*, 8299. [[CrossRef](#)] [[PubMed](#)]
45. Del Mar Bibilonic, M.; Tur, J.A. Anthropometry, Body Composition and Resting Energy Expenditure in Human. *Nutrients* **2019**, *11*, 1891.
46. Kang, Y.-J.; Yoo, J.-I.; Ha, Y.-C. Sarcopenia Feature Selection and Risk Prediction Using Machine Learning: A Cross-Sectional Study. *Medicine* **2019**, *98*, e17699. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.