Revista Mexicana de
Ingeniería Biomédica

RESEARCH ARTICLE

Predictive models of anthropometric parameters for primary screening of sarcopenia based on Machine Learning

Modelos predictivos de parámetros antropométricos para el cribado primario de sarcopenia basado en Aprendizaje Automático

Santiago Arceo-Díaz[1], Elena Elsa Bricio-Barrios[2], Sergio Sánchez-García[3], Jaime Alberto Bricio-Barrios[4], Mónica Rios-Silva[4], Miguel Huerta-Viera[1], Xóchitl Angélica Rosio Trujillo-Trujillo[1,*]

[1] *Universidad de Colima, Centro Universitario de Investigaciones Biomédicas, Av. 25 de Julio No. 965, Colonia Villas de San Sebastián, C.P. 28045, Colima, Colima, México.*

[2] *Tecnológico Nacional de México/Instituto Tecnológico de Colima, Departamento de Ciencias Básicas, Av. Tecnológico No. 1, Colonia Liberación, C.P. 28976, Villa de Álvarez, Colima, México.*

[3] *Instituto Mexicano del Seguro Social, Centro Médico Nacional Siglo XXI, Unidad de Investigación Epidemiológica y en Servicios de Salud, Área Envejecimiento, Av. Cuauhtémoc No. 330, Edificio CORSE, 3er piso, Colonia Doctores, Alcaldía Cuauhtémoc, C.P. 06720, Ciudad de México, México.*

[4] *Universidad de Colima, Facultad de Medicina, Av. Universidad No. 333, Colonia Las Víboras, C.P. 28040, Colima, Colima, México.*

[*] Corresponding author(s); contact in the panel below.

Abstract

This work reports a free-access primary screening system for detecting sarcopenia risk in older Mexican adults, using machine learning and anthropometric variables obtained through accessible instruments such as measuring tapes. An observational, retrospective, and analytical study was conducted based on records from beneficiaries of the Mexican Social Security Institute from the year 2019, with a sample of 1,678 participants. The models, developed using data from individuals without comorbidities, followed a structured machine learning workflow that included data preprocessing, variable transformation and clustering, and supervised classification using decision-tree-based models. The optimal variable combinations for men and women achieved F1-scores above 0.94, accurately classifying the risk levels of sarcopenia and severe sarcopenia. The current models need to be expanded to include individuals with comorbidities such as type 2 diabetes, hypertension, and arthritis, which have been associated with greater muscle mass loss. This proposal does not replace clinical diagnostic testing but serves as a complementary tool to rule out low-risk individuals and prioritize specialized evaluation for those who may be affected by sarcopenia.

Resumen. Este trabajo reporta un sistema de cribado primario de libre acceso para la detección del riesgo de sarcopenia en personas adultas mayores mexicanas, utilizando aprendizaje automático y variables antropométricas obtenidas mediante instrumentos asequibles como cintas métricas. Se llevó a cabo un estudio observacional, retrospectivo y analítico, basado en registros de personas derechohabientes del Instituto Mexicano del Seguro Social correspondientes al año 2019, con una muestra de 1,678 participantes. Los modelos, construidos a partir de personas sin comorbilidades, se desarrollaron mediante un flujo de trabajo estructurado de aprendizaje automático que incluyó el preprocesamiento de datos, transformación y agrupamiento de variables, y clasificación supervisada a través de modelos basados en árboles de decisión. Las combinaciones óptimas de variables para hombres y mujeres alcanzaron puntuaciones F1-score superiores a 0.94, clasificando con alta precisión los niveles de riesgo de sarcopenia y sarcopenia severa. Los modelos actuales requieren ser ampliados para incluir personas con comorbilidades como diabetes mellitus tipo 2, hipertensión arterial y artritis, las cuales se han asociado con mayor pérdida de masa muscular. Esta propuesta no sustituye una prueba diagnóstica clínica, sino que representa una herramienta complementaria para descartar a personas sin riesgo y priorizar la evaluación especializada en quienes podrían padecer sarcopenia.

Keywords: Decision trees, Machine Learning, Elderly people, Primary screening, Sarcopenia. *(Palabras clave: Árboles de decisión, Aprendizaje automático, Personas adultas mayores, Cribado primario, Sarcopenia.)*

Revista Mexicana de Ingeniería Biomédica
ISSN 2395-9126 • SOMIB

Vol./No./Year	47/2/2026
Article no.	e2026-1565
DOI	10.17488/RMIB.47.2.1565
Received	01 May 2025
Accepted	18 April 2026
Published	29 June 2026
Corresponding	rosio@ucol.mx

1. Introduction

The loss of muscle mass is a natural process associated with aging. Muscle mass has been reported to decline around the age of 40, with an average reduction of approximately 8% per decade [1]. While muscle loss can be caused by comorbidities such as cancer, sarcopenia also involves a progressive loss of muscle strength and functional independence in those who suffer from it. This condition has been associated with functional decline, falls, and, indirectly, increased mortality [2].

Depending on the level of functional decline, sarcopenia is classified into different degrees of severity. Decreased muscle strength is indicative of probable sarcopenia; If this also coincides with a decrease in muscle mass, the diagnosis of sarcopenia is confirmed; and if the functional impairment is such that the person needs constant assistance to perform activities of daily living, it is considered severe sarcopenia [3].

The clinical diagnosis of sarcopenia is mainly carried out using two devices: (i) the bioimpedance scale, which uses a painless electric current to quantify body mass, and (ii) dual-energy X-ray absorptiometry (DEXA), which emits low-intensity X-rays to distinguish and quantify body fat, muscle mass, and water content [4, 5]. These techniques have an accuracy of 95% and 98%, respectively. However, the cost of each test ranges from 25 to 150 USD, which limits their availability in most healthcare centers located in rural communities.

Screening or primary tests have proven to be useful tools for identifying, within a group of individuals, those who are not at risk of developing a disease, allowing healthcare providers to focus attention on those showing potential risk indicators. The SARC-F and SARC-CALF tests are questionnaires consisting of six questions (the first five relate to sarcopenia symptoms, and the sixth records calf circumference). While these tests have shown high performance in diagnosing severe sarcopenia compared to clinical diagnosis using DEXA, their main limitation lies in the fact that the responses depend on the participant's self-perception, introducing a risk of bias that can affect the accuracy of the primary diagnosis [6].

In contrast, the recording of anthropometric measurements offers a more objective and reproducible approach for developing primary diagnostic systems for sarcopenia. As a first approach in Mexico, Gómez de la Garza in [7] collected data from a sample of 202 participants aged 60 to 85 who attended recreational centers for older adults in the city of Monterrey. For each participant, variables such as muscle mass quantified by DEXA, body weight (kg), height (cm), calf circumference, arm circumference, waist circumference, age, and sex were recorded. Using this information, a multiple regression equation was developed, showing a diagnostic accuracy of 93% compared to the reference standard. Subsequently, Ramírez *et al.* in [8] replicated this approach in 251 participants, using muscle mass estimated by bioelectrical impedance as the output variable, achieving an accuracy of 91% in comparison with conventional clinical diagnosis.

Although anthropometric equations have shown high accuracy in predicting muscle mass and the primary screening of sarcopenia, their coefficients are specific to the population of origin, significantly limiting their applicability in other regions. Therefore, alternative approaches are needed to improve generalizability.

Machine learning offers a flexible and adaptive alternative to regression equations. This branch of artificial intelligence enables the construction of predictive models that automatically learn from data, identifying patterns that are not necessarily linear, as well as the relationships between the variables recorded for each participant [9, 10]. Among these, models based on decision trees are predictive tools that organize anthropometric variables hierarchically, using specific cutoff points to guide the classification process. Based on the magnitude recorded for each variable by the patient, the model selects the appropriate next branch, establishing a logical sequence based on the severity or relevance of the symptoms. In this way, decision trees enable a structured progression toward a preliminary diagnosis based on the values obtained. Furthermore, these models support adaptive strategies: if the tree indicates a positive diagnosis at an early level, the patient can be referred for a more comprehensive evaluation using validated instruments [11].

In the specific case of sarcopenia, decision trees offer a methodology aligned with the approach proposed by the European Working Group on Sarcopenia in Older People (EWGSOP2), which introduced a hierarchical diagnostic framework with cutoff points across four stages: initial screening using the SARC-F questionnaire, assessment of muscle strength (e.g., grip strength), confirmation through muscle mass measurement via DEXA, and classification of severity through physical performance tests such as gait speed [3].

However, the cutoff points were designed based on studies that collected anthropometric measurements from non-institutionalized older adults living in Western European countries such as Spain, France, Germany, Belgium, the United Kingdom, Italy, Sweden, and Switzerland. Subsequent validation was conducted among older adults living in the United States, Japan, and Australia. The results showed that the four-stage framework was consistent across these populations; however, the cutoff points may vary according to the anthropometric characteristics of the population under study [12-14].

In Mexico, Arceo *et al.* in [15] collected anthropometric variables from 122 older adults over the age of 60 in the metropolitan area of Colima-Villa de Álvarez and developed three systems represented through decision trees based on different combinations of muscle mass, grip strength, gait speed, and body fat percentage. The predictive accuracy of these models, compared to the bioimpedance technique, ranged from 80% to 93%. To enhance accessibility, two free-access web applications were created for those interested in the topic. The application registered under the name SARC-OPEN-IA, available at https://1introsarcopeniapy-prsmum5dlbfejjg3mmgspu.streamlit.app/, allows users to input anthropometric variables and returns a sarcopenia risk assessment based on cutoff points specific to the Mexican population.

Given that methods like DEXA or specialized dynamometers are not always available in primary care settings, the aim of this study was to develop predictive models based on accessible anthropometric measurements to estimate muscle mass, body fat percentage, and sarcopenia risk. A free-access web application is presented, which enables both healthcare professionals and general users to identify the level of sarcopenia risk in older Mexican adults. To validate the web application (available at https://mainpy-lj2lpzdm34zbxhp9swkzba.streamlit.app/), an observational, retrospective, and analytical study was conducted using databases of older adults affiliated with the Mexican Social Security Institute (IMSS), from the North and South clinics in Mexico City. The records, in .xlsx format, were processed and analyzed using Python 3.0, an open-source software. The application's visual interface was developed in Streamlit, a free web platform that allows the creation of interactive applications using dropdown menus and buttons, making data entry accessible to users with no programming experience. The application outputs a primary screening of sarcopenia risk level.

2. Materials and Methods

Records from 2019 were analyzed to minimize potential bias due to the COVID-19 pandemic. This year was selected as the baseline for pre-pandemic records, as the first confirmed case of COVID-19 in Mexico was reported on February 27, 2020. The predictive models developed in this study were constructed using data from participants without reported comorbidities, although the contribution of comorbidities in the study population is described later. It was decided to develop initial models under baseline health conditions to minimize the possible influence of non-muscular factors on sarcopenia-related functional measures. This is important because respiratory or cardiometabolic diseases can affect physical performance independent of muscle mass or strength. For example, reduced lung function has been reported in older adults with hypertension [16]. The finding by Lammie *et al.* [17] that older adults reported fewer chronic illnesses during the pandemic is a result that should be cautiously interpreted in this regard, because reduced access to healthcare services could have resulted in underreporting.

On the other hand, it remains too early to determine whether the public health impact of the pandemic will be enduring or merely transient. For this reason, this study proposes a methodology to develop predictive models that, in their initial stage, evaluate the health status of older adults prior to the pandemic, an approach that can then be extended to post-pandemic data.

The present work used data from the 2019 Mexican Older Adult Obesity, Sarcopenia, and Frailty Cohort (COSFAMM) [18], which are periodically collected by staff of the IMSS under a collaborative agreement that authorizes data use, subject to approval by the corresponding ethics committee. The objective of the study behind the COSFAMM is to evaluate the health patterns of older adults enrolled in the IMSS North and South clinics in Mexico City, focusing on monitoring specific conditions such as obesity, sarcopenia, and frailty, within a population in which long-term health inequalities have been reported among older Mexican adults [19]. Sociodemographic variables (education, occupation, and family structure); comorbidities (Elixhauser index); self-perceived health; frailty assessment (Fried criteria); sarcopenia; cognition (MMSE and verbal tests); locomotion (SPPB), physical activity (PASE); mental health (depression with CESD-R, anxiety with SAST), quality of life (WHOQOL-OLD); activities of daily living (Katz, IADL, and Barthel); and nutrition, among others, are considered. Anthropometric and metabolic measurements, such as bioimpedance phase angle, which is an indicator of nutritional and metabolic status, are also included. Of these, we consider the patient's medical record, age, anthropometric variables, locomotion, and comorbidity history.

The dataset was provided by the staff of the Epidemiological and Health Services Research Unit, Aging Division, at IMSS. The study protocol was reviewed and approved by the Research Ethics Committee CONBIOÉTICA-09-CEI-009-20160601. Data analysis was carried out in accordance with the ethical principles established in the Declaration of Helsinki for medical research involving human subjects.

To ensure confidentiality, each COSFAMM booklet includes a unique folio that allows traceability without revealing personal information. In line with this privacy protection, free and publicly accessible descriptive predictive models were developed. These models do not allow access to any volatile or identifiable data. However, researchers or institutions interested in consulting the original database may contact the authors directly.

2.1 Data curation

The study used records from 1,678 individuals enrolled in the IMSS in Mexico City. The sample corresponds to the 2019 edition of the COSFAMM.

This database contains anthropometric measurements of the registered individuals, including some of those used in the second consensus statement of the European Working Group on Sarcopenia in Older Adults (EWGSOP2). Some of these variables included body weight (kg) and height (cm); calf, arm, waist, and neck circumferences; triceps, biceps, and calf skinfold thicknesses; and handgrip strength (kg) and gait speed (m/s).

The inclusion criteria were participants born in Mexico, at least 60 years of age, with complete anthropometric measurements, and with sufficient mobility to perform the physical performance tests included in the COSFAMM booklet. To ensure data quality, data from participants with at least one outlier were excluded from the study. Although the proportion of participants with comorbidities in the sample is described later, the predictive models were constructed only from data from participants without reported comorbidities.

2.2 Development of a sarcopenia risk screening system

The data processing phase and the creation of the predictive models based on artificial intelligence techniques for building the primary screening system for sarcopenia risk in older Mexican adults are shown in Figure 1.

Figure 1 outlines the stages that make up the screening system for predicting the level of sarcopenia risk in older adults associated with comorbidities. This process consists of three phases (from left to right): data preprocessing, risk estimation, and patient screening.

(1) Data preprocessing: The 2019 database was uploaded and processed on the Google Colab platform using the Python 3 programming language. Data filtering, cleaning, and manipulation were performed using the pandas library. Each patient was represented as a single row in the dataset, while anthropometric and sociodemographic variables were orga-

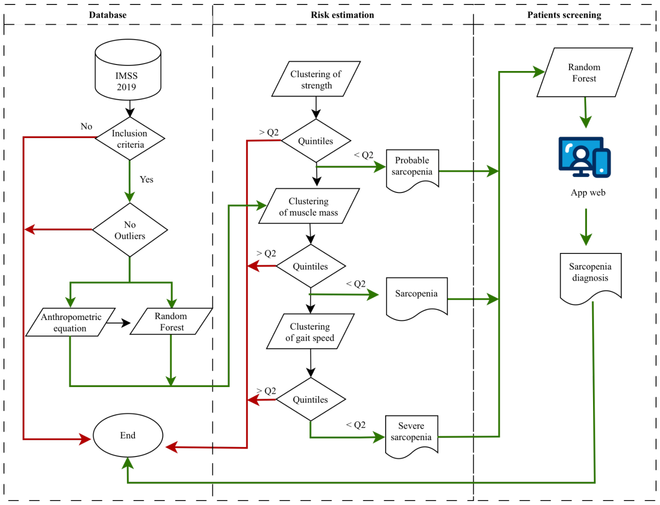

Figure 1. Flowchart of the predictive models for sarcopenia risk level.

nized as columns. To protect personal information, patient names were replaced with unique identifiers (IDs). The anthropometric variables were normalized to ensure scale comparability.

The cleaned records were subjected to the previously established inclusion criteria, and those who met them proceeded to the analysis phase. We explored the relative variability of all anthropometric variables using the coefficient of variation (CV), defined as the standard deviation divided by the mean (CV = SD / mean), which provides a standardized measure of dispersion and facilitates comparisons between men and women. We also generated correlation networks using Pearson's coefficients to visually illustrate the strength and structure of associations among anthropometric variables. These networks were not used to define inclusion or exclusion thresholds, as the aim was not variable selection but rather to describe sex-specific variability patterns and the interrelationships between measurements.

COSFAMM did not include direct records of muscle mass; this variable was indirectly estimated using the anthropometric equation (hereinafter referred to as the Ramírez equation) developed in [8], which estimates the Appendicular Skeletal Muscle Mass Index (ASMI) in older Mexican adults using a multiple regression model (Equation 1):

$$
\begin{aligned}
ASMI = 0.215 \times calfcircumference(cm) \\
+ 0.093 \times handgripstrength(kg) \\
+ 0.061 \times weight(kg) \\
+ 3.637 \times sex(male = 1; female = 0) \\
+ 0.112 \times heigh(cm) - 16.448 \quad (1)
\end{aligned}
$$

This model employs predictors such as calf circumference, handgrip strength, weight, height, and sex, achieving a group-level accuracy of ±0.1 kg and an individual precision of ±2.5 kg compared to DEXA. The equation could be used to account for reported differences in body composition between men and women [20].

We conducted an exploratory analysis to determine which anthropometric variables provide the most efficient estimation of ASMI, taking the Ramirez equation as the comparative benchmark. Two practical scenarios were evaluated: (1) the availability of a dynamometer to measure handgrip strength, and (2) the absence of such equipment. For each scenario, all possible combinations of 2 to 5 predictor variables were generated. These combinations were modeled using a decision-tree regressor due to its low computational requirements and its suitability for capturing nonlinear relationships without distributional assumptions. The datasets were randomly split into training (80%) and testing (20%) partitions using a fixed random seed to ensure methodological reproducibility. The performance of models was assessed using the root mean square error (RMSE), following recommendations in prior anthropometric modeling studies [21, 22].

The predictor subsets identified in this exploratory phase were subsequently used to develop Random Forest regression models for ASMI estimation. The modeling approach followed the performance metrics reported by Shasaif in [23] and employed a structured hyperparameter optimization strategy using RandomizedSearchCV with five-fold cross-validation. The number of decision trees (300-500) was selected to ensure convergence of the prediction error without incurring excessive computational cost. Tree depth was limited ('max depth' = 4-10) to prevent overfitting, while 'min_samples_split' (2-10) and 'min_samples_leaf' (1-4) provided additional control over tree growth and terminal node purity. Candidate configurations also varied the number of predictors considered at each split ('max_features' = 'sqrt' and 'log2') and included settings with and without bootstrap resampling ('max_samples' = 0.6-0.8). Both squared error and absolute error were evaluated as split criteria.

The primary optimization metric was mean absolute error (MAE), chosen for its interpretability in clinical settings, as it directly reflects the average deviation in ASMI units. The best-performing configuration featured 400 trees, a maximum depth of 10, and minimal constraints on sample splitting ('min_samples_split' = 2, 'min_samples_leaf' = 1). This setup favored high-resolution modeling of complex anthropometric relationships while controlling for overfitting.

In addition, the models used the sqrt method for feature selection at each split, enhancing tree diversity and predictive stability. Bootstrap resampling was enabled with a maximum sample fraction of 0.8 ('max_samples' = 0.8), further reducing variance and overfitting. These optimized Random Forest models were subsequently used in comparative analyses against the Ramírez equation for evaluating the relative importance of each predictor, with combinations of 2 to 5 predictive variables using 300 estimators to predict ASMI.

To contextualize and benchmark the Random Forest model, two additional machine-learning algorithms commonly used in clinical prediction were implemented: Gradient Boosting and Support Vector Regression (SVR) with an RBF kernel. All models were trained and evaluated under identical data partitions and preprocessing conditions. Performance was compared using standard regression metrics (RMSE, MAE, and R^2), allowing assessment of consistency, agreement, and methodological robustness across algorithms.

Agreement between machine-learning predictions and the Ramírez equation was evaluated using identity-line comparison and Bland-Altman methodology, following recommended procedures for assessing concordance between clinical estimation methods [24]. Variable importance analyses were conducted using the Random Forest framework to examine the relative contribution of anthropometric predictors and to evaluate the biological coherence of the final models.

(2) Risk estimation: The clustering stage involved categorizing the anthropometric variables following the sequential structure proposed by EWGSOP2 [3], combined with a patient-classification algorithm based on hierarchical clustering. Handgrip strength, ASMI, and gait speed were standardized using z-score normalization, after which hierarchical clustering was applied using the implementation provided by the scikit-learn library, with Ward's linkage and Euclidean distance. A range of cluster solutions (k = 2-14) was explored, and the Silhouette index, the Calinski-Harabasz coefficient, and the elbow method were used to evaluate internal validity. Across this range, the analysis revealed that both $k = 4$ and $k = 6$ provided competitive and physiologically coherent structures: $k = 6$ produced the highest silhouette values, whereas $k = 4$ yielded a clear elbow in the intra-cluster inertia plot, marking a point of diminishing returns in cluster separation. The differences in performance between these configurations were small, and adopting the four-cluster solution did not reduce the model's discriminatory ability or its capacity to separate clinically distinct subgroups. The choice of the elbow method aligns with previous stratification work in older adults from Colima, Mexico, who participated in functional dependence evaluations and indirect sarcopenia estimation through anthropometry [25], ensuring consistency with prior methodological criteria used in comparable populations.

Once the number of clusters was defined, the clustering algorithm was refined through a sequential three-stage process. First, handgrip strength was used as the sole variable for clustering, generating preliminary patient groups. In this stage, the second quintile of the full handgrip strength distribution was used as the threshold, and patients within each cluster whose values exceeded this cutoff were removed. The procedure was then repeated for ASMI and subsequently for gait speed, each time using the corresponding second quintile as the threshold for exclusion, thereby ensuring that only individuals below the Q2 cutoff for each variable remained in the final refined clusters. Participants falling below the second quintile were classified as cases of low strength, in accordance with the diagnostic criteria proposed by EWGSOP2 [3]. The use of quintiles as a classification criterion was due to the heterogeneity in the distribution of the data and the anthropometric variables analyzed, as well as the importance of identifying subgroups with the lowest recorded values.

To assess whether the classification pipeline depended excessively on the selection of the second-quintile thresholds, a sensitivity analysis was conducted by systematically shifting each physiological cutoff within a ±20% window around the reference value. The thresholds for handgrip strength, ASMI, and gait speed were modified in 5% increments, and at each shift, the full EWGSOP2-based screening process was recalculated. The number of participants assigned to the categories Probable, Sarcopenia, Severe, and Remaining was recorded for every increment.

(3) Screening of patients: With the availability of anthropometric variables corresponding to handgrip strength, muscle mass (indirectly estimated), and gait speed in participants with suspected probable sarcopenia, a hierarchical agglomerative clustering algorithm was employed, grouping individuals based on the similarity in the magnitude of the selected anthropometric variables. To identify the optimal number of clusters, the elbow method was used, which relates the number of clusters to the internal variability of the data within each cluster [26].

Subsequently, cluster groupings were generated based on three combinations of anthropometric variables. Unlike Phase 2, these combinations were selected according to parameters previously reported by Arceo *et al.* in [25] and based on the recommendations of EWGSOP2 [3]. However, instead of using the classic appendicular muscle mass index, two alternative models were designed: (i) one considering arm muscle mass, and (ii) another based on leg muscle mass.

Once the hierarchical decision tree with defined cutoff points was established, the next step was to automate the primary diagnostic process for the sarcopenia risk level. To this end, a free-access web application was developed using Streamlit. This tool enables healthcare professionals or general users without requiring programming knowledge to input anthropometric measurements through dropdown menus and obtain as output the sarcopenia risk level, regardless of the patient's comorbidity status.

3. Results and Discussion

3.1 Participant characteristics

The sample consisted of 1,678 patients, of whom 50.1% were women. Figure 2 shows the distribution of comorbidities. Some patients had more than one comorbidity at the same time, totaling 2,016 records. The largest category corresponded to participants without recorded comorbidities (38.9%). Among the categories with comorbidities, the largest contribution (49%) corresponds to records with hypertension and mild type 2 diabetes mellitus (DM2). The remaining 12% is detailed at the bottom of the figure, where rheumatoid arthritis (30.2%), depression (19.0%), and complicated DM2 (11.6%) are the most prevalent conditions.

This result is consistent with Arteaga *et al.* in [27], who reviewed comorbidities associated with diabetes mellitus in adult populations. The analysis indicates that, in Latin American countries, the most frequent comorbidities among this population group are hypertension, DM2, dyslipidemia, and overweight.

Figure 3 shows the distribution of comorbidities for participants separated by sex. Since some participants had more than one condition, a priority criterion was used to assign them to a single comorbidity group for the sex-stratified analysis. Among male participants (n = 838), 48.0% had no comorbidities, 23.2% reported mild DM2, and 37.5% had uncomplicated hypertension. Among women (n = 840), 45.5% had no comorbidities, while 17.5% reported mild DM2 and 39.8% had uncomplicated hypertension. The prevalence of other comorbidities, including complicated DM2 and complicated hypertension, was relatively low in both groups. Overall, the burden of comorbidities was slightly higher in women, particularly in the “Other” category (14.9% in women vs. 7.4% in men). For the predictive models, we kept only participants without comorbidities, resulting in a sample of 765, after also removing participants with outlier values in their anthropometric measurements.

3.2 Assessment of anthropometric variability and correlation

The COSFAMM booklet recorded multiple anthropometric measurements, as well as handgrip strength and gait speed. As an exploratory step, we evaluated the variability of all anthropometric parameters. This metric provides a direct comparison across variables with different units and scales. In addition, we computed the interquartile range (IQR) as a robust, outlier-insensitive descriptor of the absolute spread of values in their natural units. Figure 4 presents the CV values for men and women, ranked according to their mean CV values across sexes. Although several variables displayed low variability, the CV and IQR were used exclusively for descriptive purposes, and no anthropometric variables were removed based on these metrics. This ensured that all potentially informative predictors were preserved for subsequent analyses.

Subsequently, correlation networks were constructed among the selected anthropometric variables (Figure 5). Each node represents a variable, and the edges indicate statistically significant correlations above 0.3. The color of the edges denotes the sex-specificity of the correlation: blue edges correspond to men, pink to women, and magenta (purple) to correlations present in both sexes. Additionally, edge styles reflect correlation strength: dotted lines indicate correlations between 0.3 and 0.5, dashed lines represent values between 0.5 and 0.7, and solid lines indicate strong correlations above 0.7.

Figure 5 highlights that, although most associations are consistent across sexes, certain anthropometric relationships are sex-dependent (particularly those involving muscular strength and gait speed). In the male group (blue edges), grip strength shows a stronger correlation with weight, gait speed, and upper limb circumferences. Meanwhile, in women (pink edges), skinfold measurements such as triceps, biceps, and subscapular exhibit stronger internal associations and weak correlations with calf circumference. This result is consistent with López and Arroyo [20], who evaluated 7,865 older Mexican adults and found that men are taller, have a higher BMI, and have a greater waist circumference compared to women. In contrast, women exhibited higher levels of obesity and adiposity in the hips, waist, and thighs.

3.3 Adjustment of the appendicular skeletal muscle mass index

Table 1 summarizes the variable combinations evaluated to replicate the ASMI estimates produced by the Ramirez equation. The table reports results from models that include handgrip strength as a predictor as well as those that exclude it, allowing comparison between scenarios with and without access to a dynamometer, similar to what has been reported in previous studies [28, 29]. Each random forest configuration is labeled with the letter M followed by a model number, and the columns to the right display the standardized root mean square error (RMSE) obtained during both training and testing. The variable sets are ordered according to their relative predictive contribution, with sex, handgrip strength, BMI, and calf circumference emerging as the most frequently selected predictors. Although it is not listed among the top-performing five-variable models shown in the table, model M2 corresponds to the exact combination of predictors used in the Ramírez equation and yields an RMSE of approximately 0.07 during training and 0.14 during testing.

Figure 6 shows the performance indices for the Random Forest models predicting adjusted ASMI. The left panel displays the RMSE, the middle panel the MAE, and the right panel the coefficient of determination (R^2). The models are arranged in ascending order according to the magnitude of their errors. It can be observed that there is no pronounced imbalance between training and test errors (indicating the absence of severe overfitting). Additionally, model M4 (which uses sex, calf circumference, BMI, and handgrip strength as predictors) exhibits error magnitudes similar to the five-variable models with the smallest errors. All models, except M9, have a coefficient of determination above 0.75, and those with the smallest errors achieve an R^2 greater than 0.95.

On the left side of Figure 7, a strong linear correlation can be observed between the ASMI values predicted by the Random Forest model (Y-axis) and those estimated using the Ramírez equation (X-axis), as evidenced by the clustering of points along the identity line (y = x, red dashed line). This

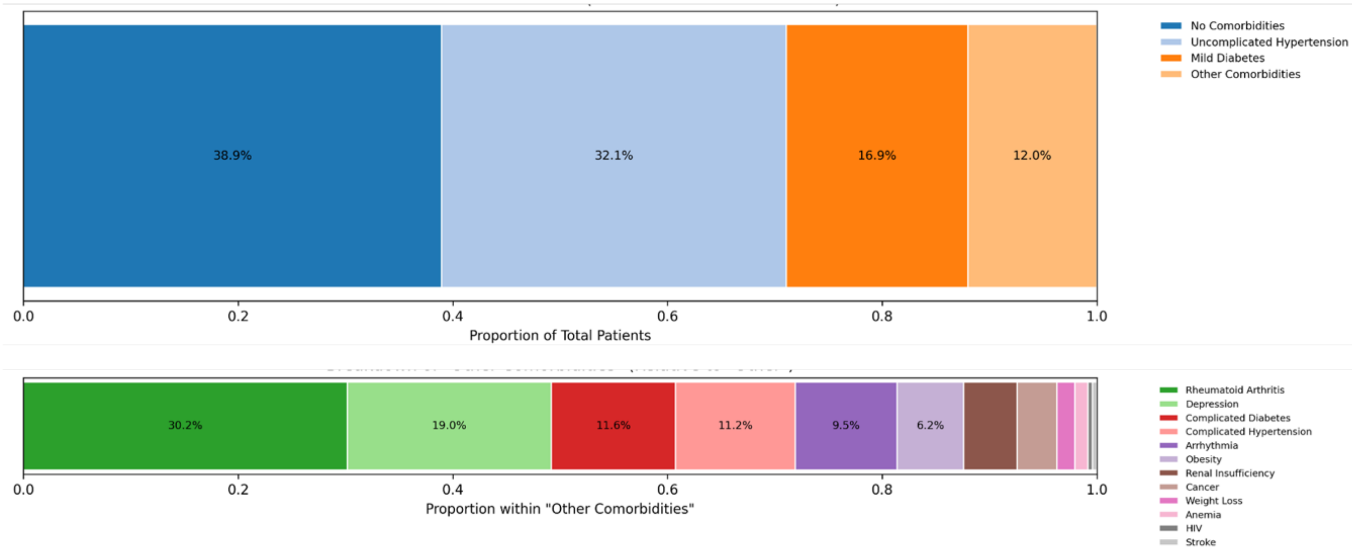

Figure 2. Distribution of the main comorbidities in the sample (up) and breakdown of the minoritary group labeled as “Other Comorbidities" (down)

Table 1. Performance of ASMI prediction models with and without grip strength as a predictor (left and right, respectively).

Best Model (with grip strength)	RMSE (Train)	RMSE (Test)	Best Model (without grip strength)	RMSE (Train)	RMSE (Test)
Sex, height, calf circumference, BMI, and, grip strength (M1)	0.059	0.110	Sex, weight, calf circumference, BMI, and gait speed (M3)	0.119	0.200
Sex, calf circumference, BMI, and grip strength (M4)	0.086	0.137	Sex, weight, calf circumference, and BMI (M5)	0.145	0.204
Sex, BMI, and grip strength (M7)	0.234	0.290	Sex, calf circumference, and BMI (M6)	0.164	0.208
Sex, and grip strength (M9)	0.479	0.518	Sex, and calf circumference (M8)	0.298	0.295

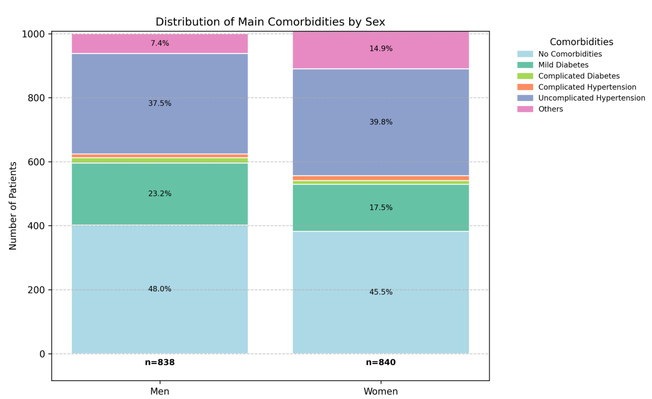

Figure 3. Distribution of the main comorbidities in the samples of each sex.

close alignment indicates that the model achieved a fit with R² 0.98, meaning the proposed model shows a high capacity to replicate the values obtained from the Ramirez equation. When analyzing the points by sex, it is noticeable that men (in blue) are grouped in a higher ASMI range, while women (in pink) are concentrated at lower values. This separation reflects expected sex differences in appendicular muscle mass and confirms that the model respects this physiological trend.

Figure 7 on the right shows the Bland-Altman analysis, which illustrates the difference between the two estimations as a function of their average value. The blue line represents the mean bias, which is practically zero (0.00), indicating no systematic overestimation or underestimation. The green and orange lines indicate the limits of agreement (LOA = ±1.96 standard deviations), which are quite narrow (±0.16), indicating that the discrepancies are small and clinically acceptable across the entire ASMI range. The residual plot confirms the absence of systematic patterns and a homogeneous distribution around zero, indicating that the model does not exhibit heteroscedasticity or dependency on the predicted value. The variable importance plot shows that sex and handgrip strength are the most influential predictors within the Random Forest, which aligns with the physiology of the muscle component and supports the biological plausibility of the model’s estimates.

In addition to developing and evaluating the proposed random forest model, this study also compared its performance against two advanced machine-learning techniques commonly applied in clinical predictive modeling. The first method, Gradient Boosting, sequentially builds weak learners, typically decision trees, with each learner aiming to correct the errors of its predecessor. This ensemble strategy is known for its robustness and has shown promising results in complex, multivariable clinical systems as predictive models for diabetes [30].

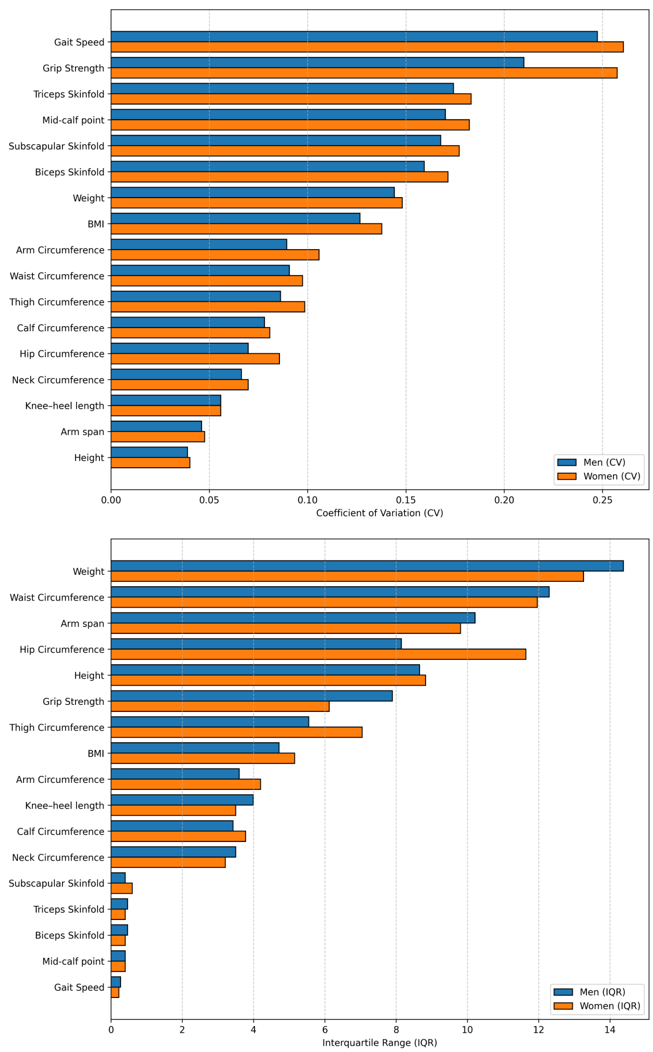

Figure 4. Coefficient of variation (up) and interquartile range (down) for the variables in the sample.

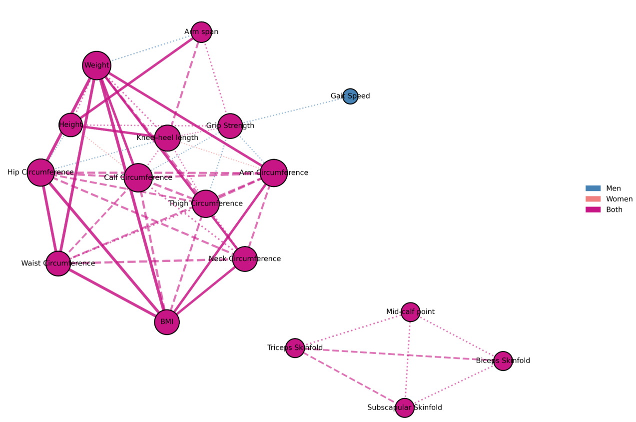

Figure 5. Correlation network of anthropometric variables in the sample.

The second method evaluated was Support Vector Regression (SVR) with a Radial Basis Function (RBF) kernel. SVR-RBF excels in datasets with non-linear trends by mapping input features into higher-dimensional spaces where linear regression becomes feasible [31].

Gradient Boosting and SVR were evaluated by Castillo *et al.* in [32], who conducted a study on sarcopenia levels in older adults attending the Tijuana General Hospital. Among the classifiers tested, the Random Forest ('max_depth' = 3, 'n_estimators' = 10) achieved a strong performance with an F1-score of 0.901 and an accuracy of 0.825. Similarly, the RBF SVM model yielded an F1-score of 0.902 and an accuracy of 0.825. Overall, models such as decision trees, k-nearest neighbors, and Gaussian processes also demonstrated comparable predictive power, highlighting the utility of machine learning-based approaches in assessing sarcopenia in clinical settings.

Figure 8 shows the Random Forest model adequately reproduces the predicted ASMI values ($R^2 = 0.981$) and maintains a consistent agreement pattern between men and women, although with higher errors (RMSE = 0.139, MAE = 0.104). Gradient Boosting improved these indicators (RMSE = 0.060, MAE = 0.047, $R^2 = 0.996$), but its training time was close to three times the time taken by the other two algorithms, making it less practical in clinical scenarios that require immediate computation. The SVR model with an RBF kernel showed the best overall performance (RMSE = 0.015, MAE = 0.009, $R^2 = 0.999$) and had training times comparable to Random Forest, along with a notable reduction in bias and residual dispersion.

Nevertheless, Random Forest was retained as the primary model. This decision is justified by its balance between accuracy, sex-specific stability, lack of overfitting, computational efficiency, and greater interpretability. Furthermore, unlike SVR, which is more sensitive to hyperparameters and less transparent, Random Forest allows for extraction of variable importance and maintains more reliable behavior for clinical extrapolation.

3.4 Clustering

Figure 9 illustrates the clustering results obtained with these indices. As shown in the upper panel in Figure 9, although higher silhouette scores were observed for $k > 10$, these solutions were considered oversegmented; within the plausible range, the highest score was obtained for $k = 6$, indicating a slightly more cohesive clustering structure. However, the lower panel in Figure 9 shows that $k = 4$ forms a clear elbow in the plot of intra-cluster inertia, suggesting diminishing returns in cluster separation beyond four groups. To determine which solution was more appropriate for the purposes of this study, both $k = 4$ and $k = 6$ were tested empirically in the subsequent EWGSOP2-based screening pipeline. After applying the sequential physiological filters (low handgrip strength, low ASMI, and slow gait speed), we observed that the final number of individuals classified into the "Probable sarcopenia", "Sarcopenia," and "Severe sarcopenia" categories was identical under both clustering solutions. Given this empirical equivalence

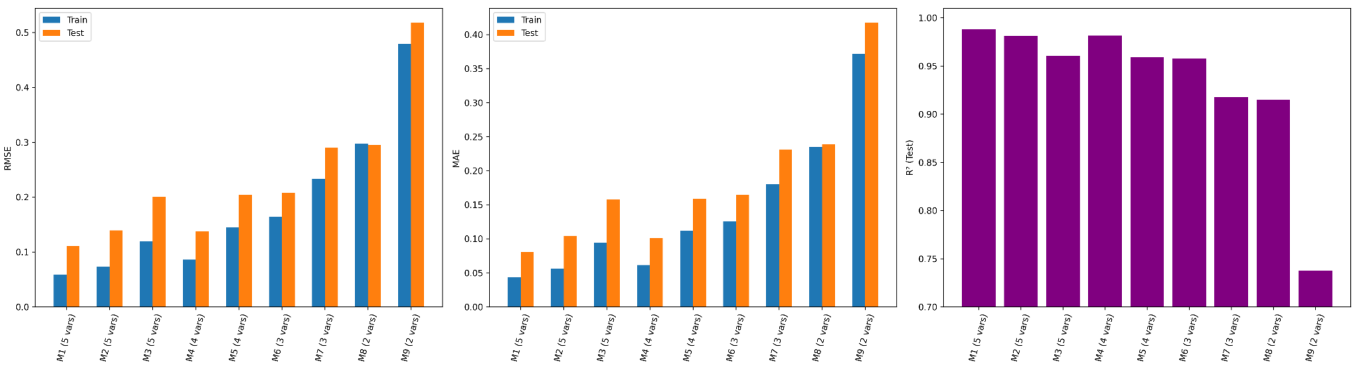

Figure 6. Performance of the models based on RMSE, MAE, and R².

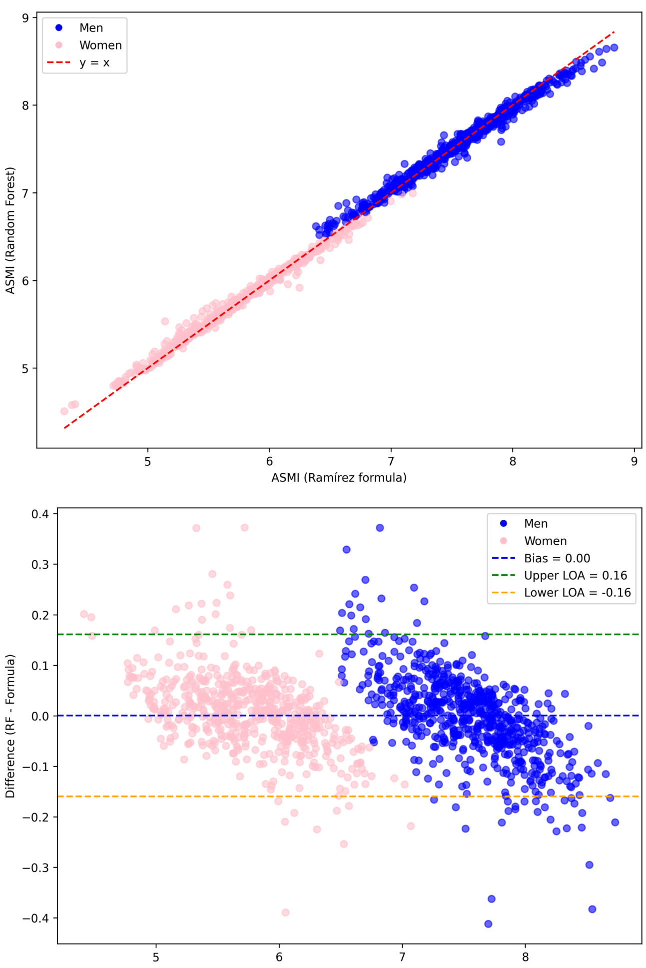

Figure 7. Comparison between predictions of the random forest model by sex with 4 predictive variables and the multiple regression model reported by Ramírez et al. [8].

and to maintain methodological simplicity and interpretability, we report the 4-cluster solution in the manuscript.

Tables 2 and 3 show the percentage of participants in each cluster who were excluded by applying the criterion of having values above the second quintile of the full range of values for handgrip strength (first filter), ASMI (second filter), and gait speed (third filter) for each sex. The header of each column shows the threshold values corresponding to these filters. In both cases, the second quintile shows values that are close to typical clinical cutoff points for these parameters (according to the EWGSOP2 [3]).

Once the number of clusters was identified, box-and-whisker plots were analyzed with the aim of exploring the relationship between anthropometric variables and the risk level of sarcopenia for each sex (Figure 10). These plots allow for a visual assessment of how each variable behaves across the defined clusters, highlighting central tendencies, dispersion, and potential outliers.

Figure 10 is organized into three sections, illustrating the relationship between sarcopenia levels and the corresponding anthropometric variables. In the top section, handgrip strength shows marked differences between sexes. In both men and women, there is a progressive decline in median grip strength from participants classified as at no risk to those with severe sarcopenia, although the category-based differences are more pronounced in men, where the values associated with severe sarcopenia diverge more sharply from the rest of the population. In contrast, the pattern for gait speed is less evident. Although there is a general tendency toward a decline from the no-risk group to the severe category, the distinctions are less clear, particularly among women, suggesting that gait speed may be less effective for differentiating sarcopenia severity in female patients. This aligns with Salinas-Rodríguez *et al.* [33], who used an operational definition of sarcopenia in older Mexican adults based on low muscle quantity accompanied by either weak handgrip strength or slow gait speed. For ASMI, a consistent downward trend is observed as sarcopenia severity increases in both sexes. Men exhibit higher ASMI values across all diagnostic categories with clearer separation between groups, whereas women show greater overlap despite presenting the same overall decreasing pattern.

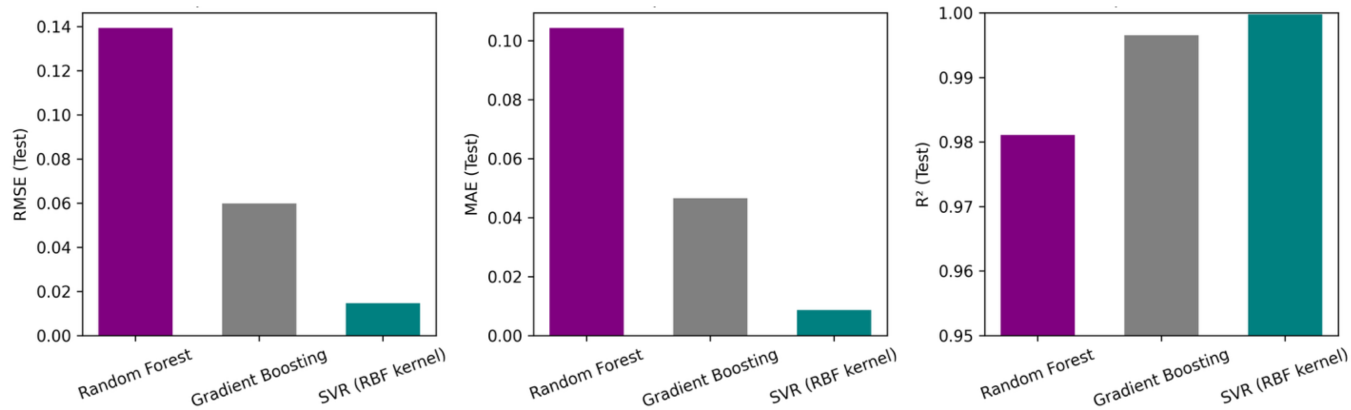

Figure 8. Performance of Random Forest, Gradient Boosting, and SVR based on RMSE, MAE, and R^2.

Table 2. Percentage of patients removed by cluster in the male sample after applying the cutoff points corresponding to the second quintile of each variable.

Cluster	Grip strength >30 kg (%)	ASMI > 7.1kg/m2 (%)	Gait speed > 0.8 m/s (%)
0	0	100	0
1	100	16	52
2	50	100	100
3	100	0	100

Additionally, the scatterplot points in Figure 10 are color-coded to reflect the number of parameters in which each participant exhibited low values: red for those with deficits in all three parameters (handgrip strength, ASMI, and gait speed), orange for deficits in two, yellow for deficits in one, and green for participants without deficits. This color scheme was designed to show how many individuals with deficits in variables not yet considered by the EWGSOP2 sequential cascade might be inadvertently classified as "Remaining" (i.e., without problems) due to the stepwise nature of the screening. The figure shows that, although some individuals in the "Remaining" category do not exhibit low handgrip strength, they may still present reduced ASMI and gait speed. However, this pattern does not persist in the downstream categories: participants classified as Probable, Sarcopenia, or Severe consistently display the expected accumulation of deficits, with those in the sarcopenia and severe sarcopenia groups predominantly presenting red or orange markers

To compare anthropometric characteristics across the four mutually exclusive groups derived from the sequential clustering process (Remaining, Probable, Sarcopenia, and Severe), the non-parametric Kruskal-Wallis H test was applied. This test evaluates whether at least one group differs significantly from the others without assuming normality, making it appropriate for skewed or non-homogeneous distributions. Given that the analysis involved four independent groups, Kruskal-Wallis allowed us to detect global differences for each variable. P-values below 0.05 were considered statistically significant. Under this criterion, the variables showing the strongest group separation were handgrip strength, ASMI, gait speed, BMI, calf circumference, and several circumferences and skinfolds, while variables such as triceps skinfold, biceps skinfold, and some peripheral measurements showed weaker or non-significant differences. This pattern reflects the expected physiological gradient across the classification, where muscle strength, muscle mass, and physical performance concentrate most of the discriminatory power.

Table 4 shows that the Kruskal–Wallis test revealed statistically significant differences ($p < 0.05$) across the four sarcopenia groups for all anthropometric and physical performance variables in both men and women. The most pronounced differences were observed in handgrip strength, ASMI, and gait speed, with extremely low p-values supporting their central role in the diagnostic criteria for sarcopenia. Significant differences were also found in other variables such as body weight, BMI, and the circumferences of the calf, arm, thigh, waist, and hip. These findings indicate that both muscle mass indicators and physical performance measures contribute substantially to the differentiation between sarcopenia stages.

To assess the robustness of this primary screening approach, we performed a sensitivity analysis by shifting the percentile-based thresholds used for handgrip strength, ASMI, and gait speed around the 40th percentile (Q40) and re-running the full classification pipeline (Figure 11). Across shifts from 0.20 to +0.20, the distribution of patients among the four screening categories (Remaining, Probable, Sarcopenia, and Severe) remained generally stable, indicating low susceptibility to minor threshold changes. In men, Remaining decreased smoothly with positive shifts, and Severe increased as expected in a sequential screening system, with Probable and Sarcopenia showing minimal fluctuation; women exhibited an even more attenuated version of the same pattern. However, beyond ±0.10,

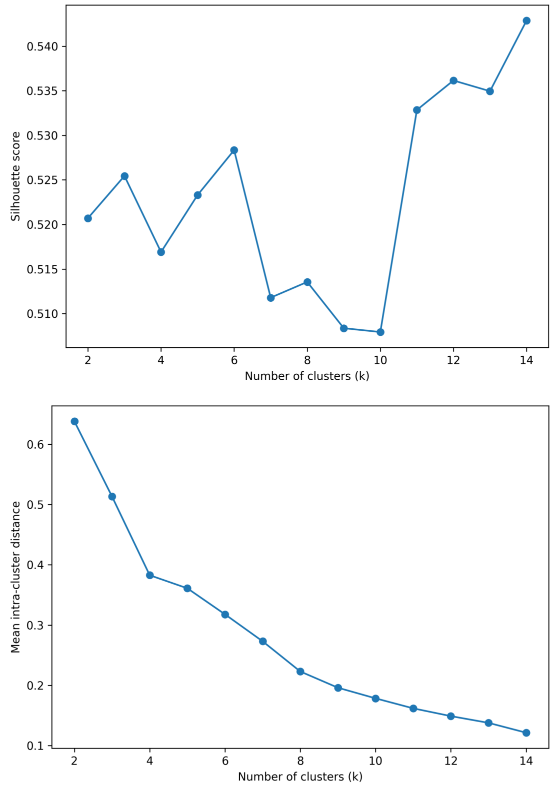

Figure 9. Silhouette (up) and elbow (down) plots of the study participants.

Table 3. Percentage of patients removed by cluster in the female sample after applying the cutoff points corresponding to the second quintile of each variable.

Cluster	Grip strength >17 kg (%)	ASMI > 5.6kg/m2 (%)	Gait speed > 0.67 m/s (%)
0	0	0	52
1	100	100	100
2	57	82	0
3	100	100	100

Table 4. P-values for differences in anthropometric variables for male and female samples by the Kruskal-Wallis test.

Variable	Men	Female
Grip strength	1.63e-61	1.14e-57
ASMI	3.17e-21	1.14e-57
Gait speed	5.85e-17	1.70e-11
Calf	8.26e-23	5.54e-20
Arm	1.39e-16	4.61e-15
BMI	7.49e-22	4.43e-19
Weight	6.02e-19	8.49e-16
Thigh	5.78e-17	1.15e-15
Waist	3.53e-12	1.64e-11
Hip	9.93e-12	7.67e-13

both sexes showed sharper increases in Severe and marked declines in Remaining, suggesting that larger deviations begin to alter the clinical balance of categories rather than reflecting acceptable flexibility around the cut points.

Figure 11 shows that the classification system demonstrates high stability against small variations in the cutoff points, particularly within the interval between -0.10 and +0.10. Among men, as the cutoff points were adjusted, there was a progressive decrease in the percentage of individuals classified as non-sarcopenic and a corresponding increase in those categorized as having severe sarcopenia, while the proportions of the intermediate groups remained relatively stable. This pattern reflects a behavior consistent with the cascade structure of the classification proposed by the EWGSOP2 [3]. In contrast, among women, the shifts in cutoff points produced more modest changes in the distribution of individuals across diagnostic categories, although a gradual transition toward more severe classifications was also observed.

Table 5 shows the distribution of participants according to their primary sarcopenia risk level, disaggregated by sex. One of the most notable findings is the higher proportion of men classified in the "severe sarcopenia" category, where they represent approximately 60%. In contrast, there is a higher

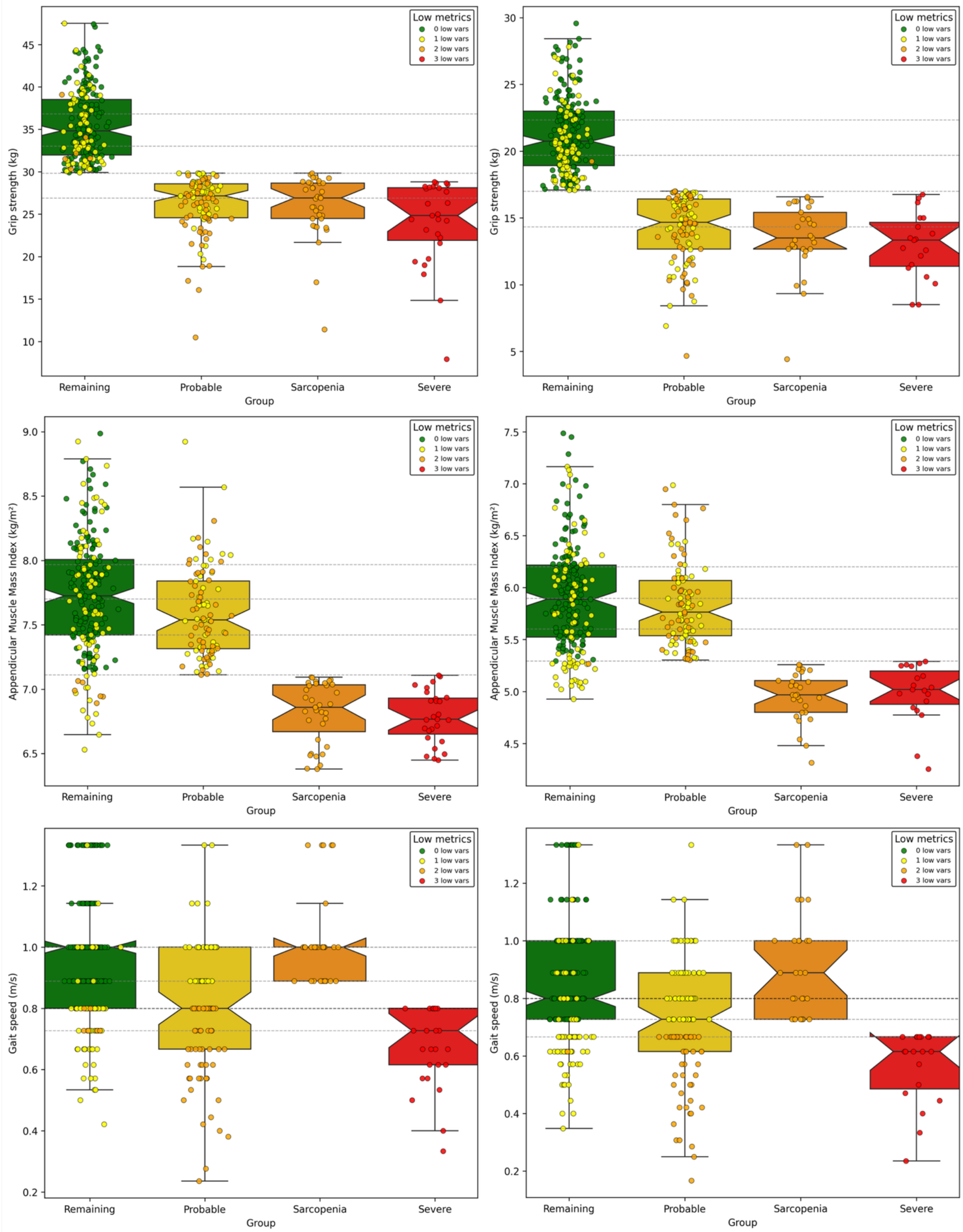

Figure 10. Boxplots of the relationship between anthropometric variables and the risk level of sarcopenia for each sex.

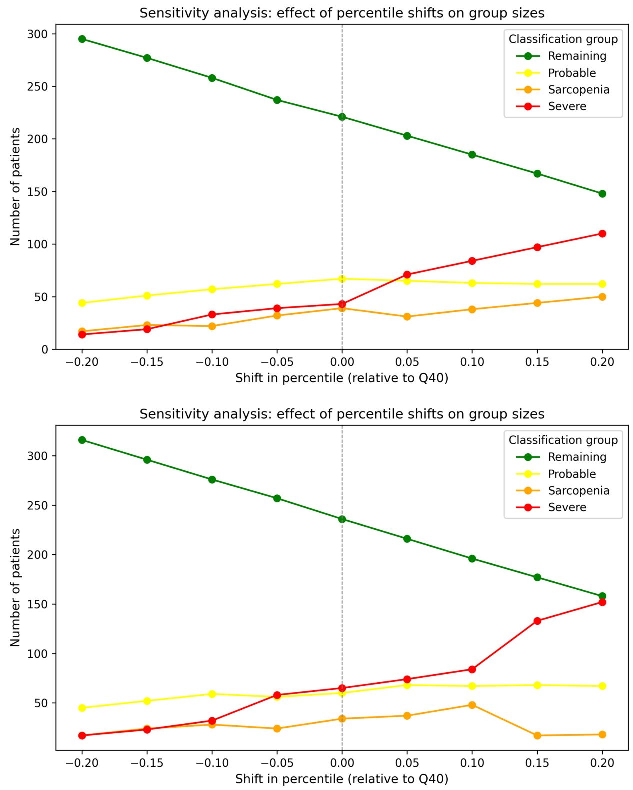

Figure 11. Sensitivity analysis of percentile-based thresholds for primary sarcopenia screening for men (up) and women (down).

Table 5. Estimates of patients with sarcopenia by sex, with each risk group shown as mutually exclusive.

Risk	Female	Male	n
Severe sarcopenia	43	65	108
Sarcopenia	39	34	73
Probable	67	60	127
Remaining	221	236	457

proportion of women in the categories of sarcopenia and probable sarcopenia. This trend may reflect earlier detection due to differences in healthcare access or health-seeking behavior. However, it is important to note that previous studies have found that the prevalence of sarcopenia measured in study groups depends on factors related to the studied population and the cutoff points [34], so the observed proportions should not be taken as an estimate of sarcopenia prevalence in Mexico.

3.5 Random forest for sarcopenia risk estimation

A Random Forest model was developed to classify patients into three categories: "No Sarcopenia", "Sarcopenia" and "Severe Sarcopenia," using handgrip strength, gait speed, and ASMI as predictive variables, following previous machine-learning approaches for sarcopenia classification [35]. The dataset was split into training and testing sets using a stratified 70:30% partition to preserve class proportions. The model was trained using 100 estimators, with a fixed random seed to ensure reproducibility. The feature importance analysis showed that handgrip strength had the highest predictive value, followed by gait speed and ASMI. The Random Forest model achieved an overall accuracy of 99% and a weighted F1-score of 0.99, indicating excellent overall performance. Class-specific analysis showed a recall of 1.00 and F1-score of 0.99 for patients without sarcopenia, a recall of 0.94 and F1-score of 0.97 for sarcopenia, and perfect recall and F1-score values of 1.00 for severe sarcopenia.

Figures 12 and 13 show the most representative decision trees generated from Random Forest models trained to classify the risk of sarcopenia in female (up) and male (down) patients from the IMSS 2019 sample. In both cases, the models rely on three key predictors: muscle strength, gait speed, and ASMI.

The classification follows a hierarchical structure that begins with a threshold in handgrip strength, followed by cutoffs in gait speed and ASMI, consistent with the diagnostic algorithm proposed by the EWGSOP2 working group. Notably, the initial strength threshold differs by sex: 16.83 kg for women and 29.29 kg for men, accurately reflecting physiological differences between the two groups.

Each node in the tree represents a decision point and includes the number of participants meeting that condition, the distribution of classes at that point (no sarcopenia, sarcopenia, severe sarcopenia), and the final predicted class. Terminal leaves are color-coded as follows: orange for no risk, green for sarcopenia, and purple for severe sarcopenia.

Although the overall structure of the trees is similar for

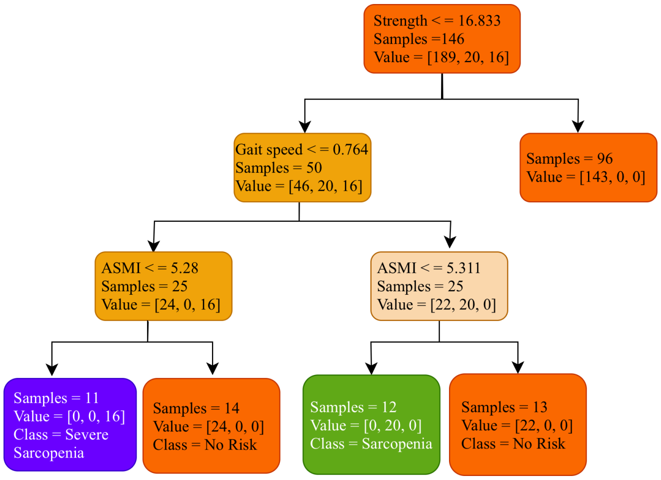

Figure 12. Best trees within the Random Forest models for predicting sarcopenia risk level in the female group.

both sexes, there are key differences in the thresholds and decision paths. In men, the model places greater weight on handgrip strength as the initial discriminative factor, followed by gait speed. In contrast, among women, gait speed tends to appear earlier in some paths, which may reflect a different functional profile or a stronger association between mobility and sarcopenia risk in this group.

Additionally, the ASMI threshold values used by the models differ between men and women, highlighting the importance of sex-specific criteria in clinical assessment. For women, even with acceptable handgrip strength, a low gait speed combined with reduced ASMI can lead to a classification of severe sarcopenia. In men, however, a sufficiently high ASMI may partially offset lower functional performance. While diagnostic criteria are shared, their relative weight and effective combinations may vary by sex. This finding reinforces the need to adapt screening and prevention strategies with a gender-sensitive approach.

The relevant thresholds obtained from Partial Dependence Plots (PDPs) for sarcopenia risk classification were identified through partial dependence plots generated from Random Forest models to estimate the average marginal effect of each predictor on the probability of belonging to a class while conditioning on the other variables. PDPs are used to interpret machine learning models [36]. They illustrate how the model's prediction changes when a particular variable is modified, while the effects of other variables are averaged out or held constant. This technique is particularly useful in clinical contexts, as it helps identify critical thresholds in functional and morphological variables associated with the risk of complex conditions such as sarcopenia [37, 38]. Figures 14 and 15 show how three central variables: handgrip strength, gait speed, and ASMI, influence the probability of presenting different degrees of sarcopenia in both men and women.

In men (Figure 14), handgrip strength shows a critical threshold around 30 kg: below this value, the probability of sarcopenia, especially in its severe form, increases considerably, while

Table 6. Variable combinations including anthropometry and at least one EWGSOP2 parameter used in Random Forest models for male and female samples.

Sex	Variable combination	Weighted F1-score
Male	Gait speed, ASMI, weight and triceps skinfold	0.945
Male	Gait speed and ASMI, arm and subscapular skinfold	0.944
Male	Gait speed, ASMI, weight, thigh, and calf	0.940
Female	Grip strength, gait speed, waist, and BMI	0.950
Female	Grip strength, gait speed, waist, and weight	0.940
Female	Grip strength, gait speed, waist, and arm	0.923

exceeding it favors classification as not at risk. Gait speed presents a clear inflection point at 0.75 m/s, with a marked reduction in risk beyond this threshold. Finally, ASMI shows a relevant change near 6.7 kg/m², serving as a lower limit below which the probability of sarcopenia rises.

In women (Figure 15), although the overall trends are similar, the thresholds are notably different. Handgrip strength reaches its critical point around 17.2 kg, beyond which the probability of sarcopenia decreases significantly. As in men, gait speed exhibits a very consistent pattern: below 0.75 m/s, the risk increases, whereas higher values are associated with a lower probability of sarcopenia. Regarding ASMI, the threshold with the most significant change is approximately 5.3 kg/m², reflecting physiological differences in body composition compared to men.

Although the Random Forest model based on variables defined by the EWGSOP2 consensus showed adequate performance, in clinical or population-based contexts, the full set of these measurements is not always available. To address this limitation, an exhaustive search approach was implemented in two scenarios: a) partial availability of one or two functional variables from the European consensus in combination with anthropometric variables, and b) complete absence of these variables, relying solely on anthropometric measurements.

In each scenario, random combinations of predictors were generated, and the performance of Random Forest models was evaluated using the weighted F1-score. This metric provides a more robust evaluation of model performance compared to accuracy, as it penalizes both false positives and false negatives [39]. To mitigate class imbalance, the SMOTE (Synthetic Minority Over-sampling Technique) method was applied during the training phase. This method generates synthetic instances of the minority class by interpolating between each sample and its nearest neighbors within the same class, rather than duplicating existing data, and may improve classification performance under imbalanced conditions [40].

Table 6 shows the F1-score of the Random Forest models using different combinations of anthropometric variables and those included by the EWGSOP2 consensus. These combinations yielded high classification performance, with weighted F1-scores equal to or greater than 0.94 in men and up to 0.95 in women.

Table 7 compares the efficiency of the predictive model developed in this study with that of the main working groups dedicated to primary sarcopenia screening, as reported by

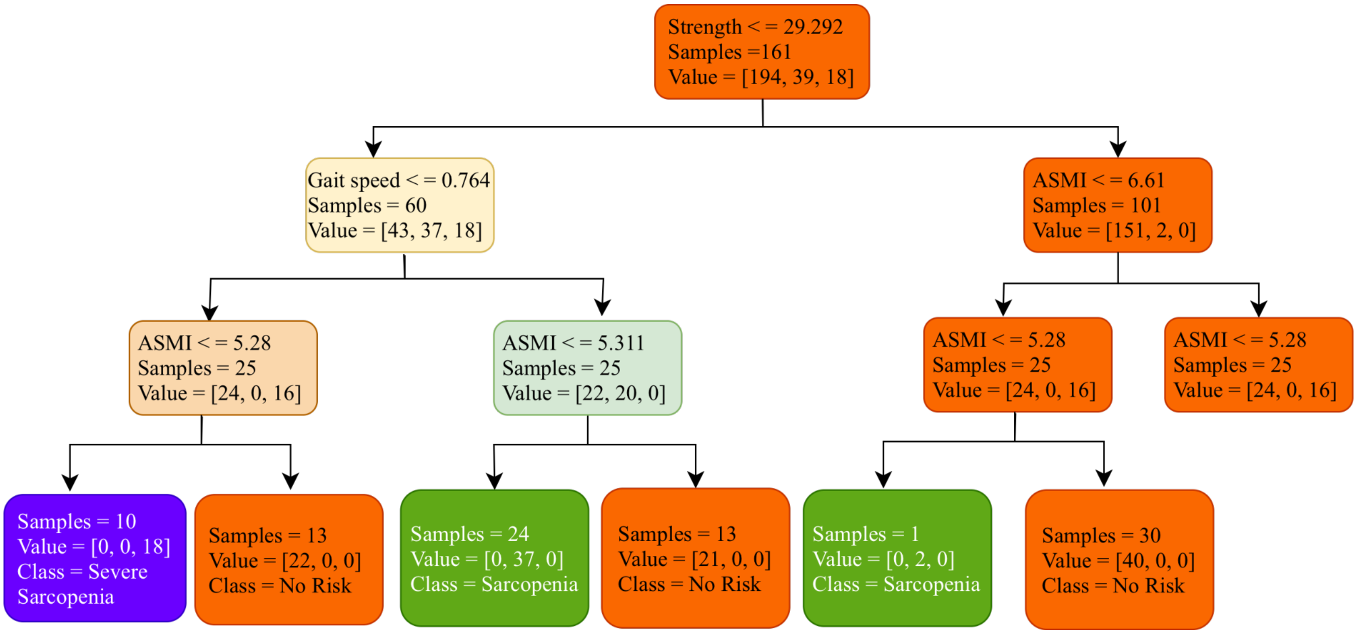

Figure 13. Best trees within the Random Forest models for predicting sarcopenia risk level in the male group.

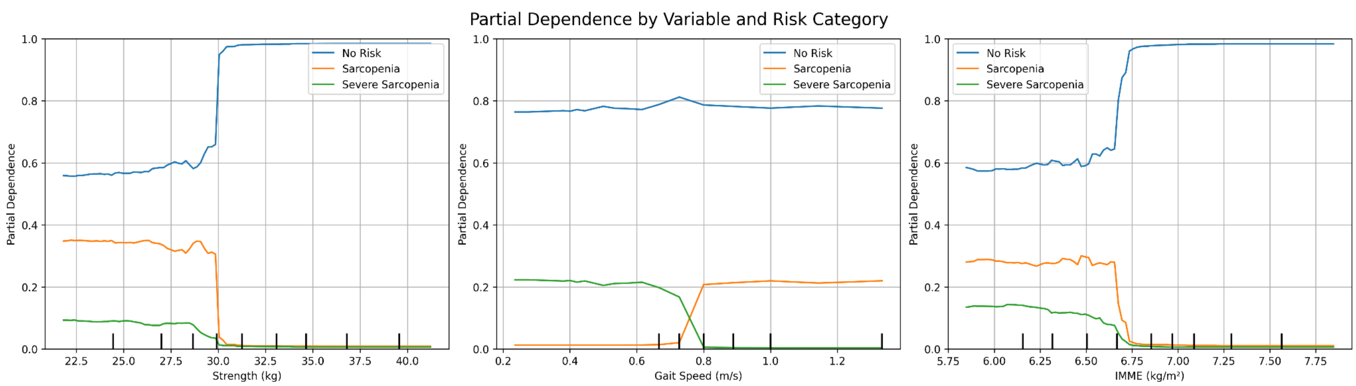

Figure 14. Partial dependence plots for risk level in the male sample.

Sepúlveda *et al.* in [41]. It is worth noting that alternative proposals exist, such as those from the Society on Sarcopenia and Cachexia and the International Working Group on Sarcopenia, which define their ASMI cutoff points based on sample standard deviation criteria that fall outside the scope of the present study.

Table 7 shows that performance metrics improve when the diagnostic model includes a larger number of variables. Specifically, the combination of ASMI, handgrip strength, and gait speed yields markedly superior predictive performance (F1-score, accuracy, and recall close to or equal to 1.00 in both sexes). In contrast, models relying solely on ASMI and gait speed demonstrate lower performance. This highlights that the effectiveness of a predictive model depends not only on the number of variables included but also on their discriminative value for identifying sarcopenia.

It is important to note that excluding handgrip strength significantly reduces its predictive capacity. This finding is consistent with the results reported by Blanquet *et al.* [42], who assessed anthropometric variables in 118 older adults in France between 2014 and 2018 and determined sarcopenia risk using the cutoff points proposed by EWGSOP. The authors found that low handgrip strength is strongly associated with an increased risk of sarcopenia, highlighting the clinical relevance of including this variable in screening systems.

4. Conclusions

This study presents the design, evaluation, and validation of a free-access digital tool developed to estimate the risk of sarcopenia in older adults using anthropometric variables. The system is based on a supervised classification models trained with the Random Forest algorithm, which showed adequate performance, especially when at least one or two functional parameters recommended by the EWGSOP2 consensus were included.

The tool offers two main functionalities: (i) consultation of

doi: 10.17488/RMIB.47.2.1565

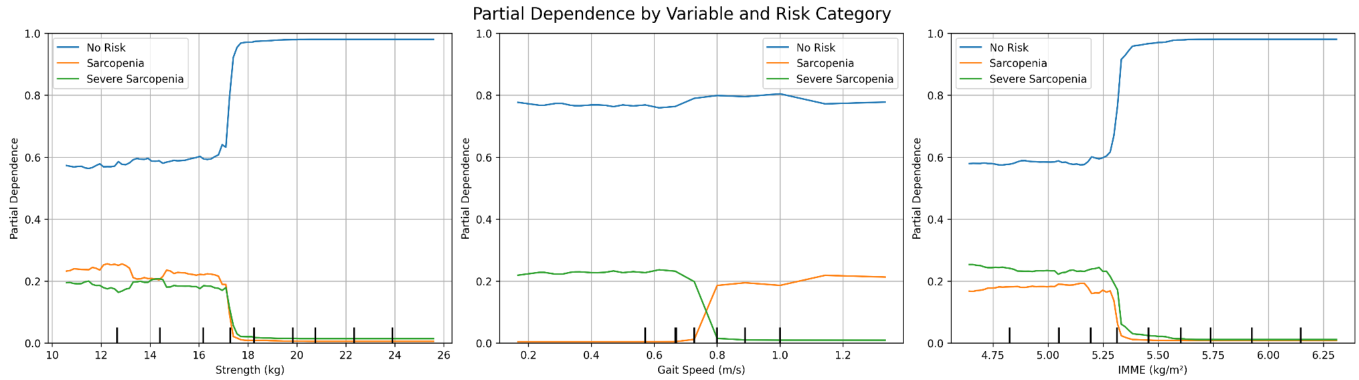

Figure 15. Partial dependence plots for risk levels in the female sample.

Table 7. Comparison of diagnostic criteria and predictive performance across sarcopenia models

Working Group	Variable combination	Weighted F1-score (male/female)	Accuracy (male/female)	Recall (male/female)
European/Asian Working Group on Sarcopenia in Older People	ASMI, handgrip strength and gait speed	0.99/1.00	0.99/1.00	0.97/1.00
International Sarcopenia Group/Special Interest Group	ASMI and gait speed	0.65/0.75	0.62/0.72	0.79/0.82

predefined or customized anthropometric variables and cutoff points to estimate primary sarcopenia risk in older Mexican adults without comorbidities; and (ii) input of individual data or complete datasets to obtain a preliminary classification by risk level. Additionally, the application allows exporting results, facilitating its use in both clinical and research settings. This approach may support diagnostic prioritization, particularly in primary care contexts where advanced clinical infrastructure is limited.

The practical feasibility of implementing this tool lies in its reliance on accessible, low-cost measurements that are already common in primary care and community health programs across Mexico. By leveraging routine anthropometric and simple functional assessments, the system can be integrated into existing workflows without additional equipment, potentially facilitating early identification of individuals at elevated risk and guiding appropriate referrals for confirmatory evaluations or preventive interventions. Nevertheless, the tool is intended to complement standardized diagnostic procedures, especially in contexts where more precise reference methods such as DEXA are available.

Several limitations must be acknowledged. The models were developed exclusively with data from individuals without comorbidities, which may restrict their applicability to broader clinical populations. Muscle mass was estimated indirectly using anthropometric equations rather than measured with gold-standard techniques, potentially affecting accuracy. The quintile-based cutoff values derived from the study sample may limit external validity, and although oversampling methods such as SMOTE were used to mitigate class imbalance, such strategies do not fully guarantee comparable performance under real-world conditions. In addition, gait speed was recorded with low numerical resolution (one significant digit), potentially affecting the the accuracy of risk categorization.

Future research will incorporate data from individuals with common comorbidities to assess model performance in more representative clinical scenarios. External validation in different populations and geographic contexts is also necessary to evaluate generalizability. Comparative analyses with alternative supervised learning algorithms may provide insights into potential performance improvements. In addition, integrating longitudinal data could allow assessment of the stability of the models over time in predicting sarcopenia progression. Finally, implementation studies in primary healthcare settings are recommended to evaluate the practical utility and impact of the proposed digital tool in clinical workflows.

5. Funding

This work was supported by the Secretaría de Ciencia, Humanidades, Tecnología e Innovación through a postdoctoral fellowship.

6. Acknowledgements

The authors thank the Secretaria de Ciencia, Humanidades, Tecnología e Innovación for the postdoctoral fellowship granted to CVU 231269, which made this work possible.

7. Author Contributions (CRediT)

Conceptualization, Arceo-Díaz, S. and Trujillo-Trujillo, X.A.R.; Methodology, Sánchez-García, S.; Investigation, Bricio-Barrios, J.A. and Rios-Silva, M.; Data curation, Arceo-Díaz, S.; Software, Arceo-Díaz, S. and Bricio-Barrios, E.E.; Formal analysis, Trujillo-Trujillo, X.A.R.; Writing – review & editing, Arceo-Díaz, S., Bricio-Barrios, E.E. and Huerta-Viera,

M.; Funding acquisition, Trujillo-Trujillo, X.A.R. and Huerta-Viera, M.; Project administration, Trujillo-Trujillo, X.A.R.; Resources, Sánchez-García, S.; Supervision, Trujillo-Trujillo, X.A.R. and Huerta-Viera, M.

8. Conflict of Interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and / or publication of this article.

9. Data Availability Statement

The data are available from the corresponding author upon reasonable request.

10. Bioethics

The project was approved by the Bioethics Committee of the State Cancer Institute and verified by the National Bioethics Commission in Mexico under registration number: CONBIOÉTICA-09-CEI-009-20160601.

11. Artificial Intelligence Use Disclosure

During the preparation of this manuscript, generative AI and language editing tools were used for editorial assistance and coding assistance to implement analyses defined by the authors. These tools were not used to perform autonomous data analysis, make methodological decisions, interpret findings, formulate conclusions, or generate or modify results, tables, or figures. All AI-assisted text, code, and outputs were reviewed, validated, and approved by the authors, who take full responsibility for the manuscript.

12. References

[1] Chen C, Liao DM. *Mechanisms for the development of sarcopenia and its medical and nutritional prevention: A narrative review.* Tungs' Medical Journal, vol. 19, supl. 1, pp. S11–S15, 2025. https://doi.org/10.4103/etmj.etmj-d-25-00003

[2] Aguilera Reguera D, Fuentes Díaz Z, Rodríguez Salazar O. *Estimación de factores predictivos de la sarcopenia secundaria al cáncer en el paciente quirúrgico*. Gaceta Médica Espirituana, vol. 26, s/p, 2024. Consulted November 2025. https://dialnet.unirioja.es/servlet/articulo?codigo=9972492

[3] Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, Cooper C, Landi F, Rolland Y, Sayer AA, Schneider S M, Sieber CC, Topinkova E, Vandewoude M, Visser M, Zamboni M. *Sarcopenia: revised European consensus on definition and diagnosis*. Age and Ageing, vol. 48, no. 1, pp. 16–31, 2019. https://doi.org/10.1093/ageing/afy169

[4] Cheng K, Chow S, Hung V, Wong C, Wong R, Tsang C, Cheung W. *Diagnosis of Sarcopenia by evaluating Skeletal muscle mass by adjusted Bioimpedance analysis validated with Dual-Energy X-Ray Absorptiometry*. Journal of Cachexia, Sarcopenia and Muscle, vol. 12, no. 6, pp. 2163–2173, 2021. https://doi.org/10.1002/jcsm.12825

[5] Tseng Y, Mo S, Zeng Y, Zheng W, Song H, Zhong B, Luo F, Rong L, Liu J, Luo Z. *Machine Learning Model in Predicting Sarcopenia in Crohn's Disease Based on Simple Clinical and Anthropometric Measures*. International Journal of Environmental Research and Public Health, vol. 20, no. 1, p. 656, 2023. https://doi.org/10.3390/ijerph20010656

[6] Rocha A, Teixeira N, Santos A, Fonseca A, Rocha JS. *Risk of Sarcopenia in Older People: Association with Illiteracy and Negative Self-Perception of Health*. Revista Psicologia e Saúde, vol. 17, s/p, 2025. https://doi.org/10.20435/pssa.v1i1.2842

[7] Gómez de la Garza M. *Desarrollo y validación de una ecuación antropométrica para estimar la masa apendicular muscular utilizando como referencia la absorciometría dual de rayos x en adultos mayores*. Universidad Autónoma de Nuevo León, tesis, 2013. Consulted November 2025. http://eprints.uanl.mx/11871/

[8] Ramírez E, Enríquez-Reyna M, Garza-Sepúlveda G, Tijerina-Sáenz A, Ramos-Peña E, Gómez de la Garza M. *Puntos de corte y validación de una ecuación antropométrica para estimar la masa muscular en población mexicana*. Salud Pública de México, vol. 57, no. 6, pp. 485–486, 2015. Consulted November 2025. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0036-36342015000600003

[9] Javadpour N. *Machine Learning Approaches Towards Cybersickness Prediction: An Updated Systematic Review*. Iowa State University, Master thesis, 2024. https://doi.org/10.31274/td-20250502-278

[10] Olshvang D, Harris C, Chellappa R, Santhanam P. *Predictive modeling of lean body mass, appendicular lean mass, and appendicular skeletal muscle mass using machine learning techniques: A comprehensive analysis utilizing NHANES data and the Look AHEAD study*. PLoS One, 19(9), e0309830. 2024. https://doi.org/10.1371/journal.pone.0309830

[11] Charbuty B, Abdulazeez A. *Classification based on decision tree algorithm for machine learning*. Journal of Applied Science and Technology Trends, vol. 2, no. 1, pp. 20–28, 2021. https://doi.org/10.38094/jastt20165

[12] Zhang S, Yang X, An N, Lv M, Yang L, Liu R, Hu S, Chen W, Feng W, Mao Y. *Risk factors and predictive models for sarcopenia in older adults*. Aging Medicine. 2025;8(3): 192-199. https://doi.org/10.1002/agm2.70012

[13] He X, Song Y, Ma L, Ainsworth BE, Liu Y, Chen N. *Prevalence and factors influencing sarcopenia among community-dwelling older adults using the Asian Working Group for Sarcopenia Definition*. Clinical Interventions in Aging, pp. 1707–1727, 2022. https://doi.org/10.2147/cia.s388319

[14] Gómez Salazar L, Arango Hoyos G, Dueñas E, Reyes Ortíz C. *Probabilidad de sarcopenia asociada a funcionalidad y riesgo de caídas en adultos mayores de Colombia*. Retos, no. 63, pp. 63–73, 2025. https://doi.org/10.47197/retos.v63.100622

[15] Arceo-Díaz S, Bricio-Barrios EE, Trujillo-Trujillo XAR, González-Farías JR, Bricio-Barrios JA, Ríos-Silva M, Huerta-Viera M. *Sistema de Cribado Primario para la*

Sarcopenia en Personas Adultas Mayores Basado en Inteligencia Artificial. Revista Mexicana de Ingeniería Biomédica, vol. 44, no. SPE1, pp. 53–69, 2023. https://doi.org/10.17488/rmib.44.4.4

[16] Brandao-Rangel M, Brill B, de Souza Carvalho E, Melamed D, Moraes-Ferreira R, Silva-Reis A, Leonardo P, Frison CR, De Angelis K, Vieira R. *Physically Active Lifestyle Attenuates Impairments on Lung Function and Mechanics in Hypertensive Older Adults*. Advances in Respiratory Medicine, vol. 92, no. 4, pp. 278–290, 2024. https://doi.org/10.3390/arm92040027

[17] Lammie S, Rodríguez-Araña S, Posada Rodríguez C, Flores-Cuadra J, Pérez-Lao A, Britton GB, Oviedo DC, Tratner AE. *A Cross-Sectional and Longitudinal Analysis of Cognitive Function and Well-Being of Older Adults in Panama During the COVID-19 Pandemic*. COVID, vol. 5, p. 128, 2025. https://doi.org/10.3390/covid5080128

[18] Instituto Mexicano del Seguro Social. *Cuadernillo de Obesidad, Sarcopenia y Fragilidad en Adultos Mayores (OSFAMM)*. IMSS, México, 1ª ed., 2018.

[19] Salinas-Rodríguez A, Rojas-Botero M, Rivera-Almaraz A, Fernández-Niño J, Montañez-Hernández J, Manrique-Espinoza B. *Long-term inequalities in health among older Mexican adults: An outcome-wide analysis*. SSM–Population Health, vol. 26, p. 101684, 2024. https://doi.org/10.1016/j.ssmph.2024.101684

[20] López-Ortega M, Arroyo P. *Anthropometric characteristics and body composition in Mexican older adults: age and sex differences*. British Journal of Nutrition, vol. 115, no. 3, pp. 490–499, 2016. https://doi.org/10.1017/S0007114515004626

[21] Ko JB, Kim KB, Shin,Y. S, Han H, Han S. K, Jung D. Y, Hong, JS. *Predicting sarcopenia of female elderly from physical activity performance measurement using machine learning classifiers*. Clinical Interventions in Aging, 1723-1733.2021. 2021. https://doi.org/10.2147/CIA.S323761

[22] MacEwan JP, Gill TM, Johnson K, Sullivan J, Shim J, Goldman DP. *Measuring sarcopenia severity in older adults and the value of effective interventions*. The journal of nutrition, health & aging, 22(10), 1253-1258. 2018. https://doi.org/10.1007/s12603-018-1104-7

[23] Shasaif M, Shaik A. *Enhancements in Random Forest Algorithms for Improving Healthcare Applications*. In Proceedings of the International Conference on Industrial Engineering and Operations Management; IEOM Society International: Southfield, MI, USA. 2024. https://doi.org/10.46254/IN04.20240192

[24] Silveira P, Vieira, J, Siqueira J. *Is the Bland-Altman plot method useful without inferences for accuracy, precision, and agreement?* Revista de Saúde Pública, 58, 01. 2024. https://doi.org/10.11606/s1518-8787.2024058005430

[25] Arceo-Díaz S, Bricio-Barrios EE, González-Farías JR, De la Madrid-Trevizo L, Trujillo-Trujillo XAR, Bricio-Barrios JA, Huerta-Viera M. *Sistema Biomédico Basado en Inteligencia Artificial para Estimar Indirectamente Sarcopenia en Personas Adultas Mayores Mexicanas*. Revista ELECTRO, vol. 46, no. 1, pp. 18–24, 2024. Available at: https://itchihuahua.mx/revista_electro/2024/A53_18-24.html

[26] Fritz M, Beutel A, et al. *LOG-Means: Efficiently Estimating the Number of Clusters in Large Datasets*. Proceedings of the VLDB Endowment, 2020.

[27] Arteaga-Zambrano A, Solorzano-Veliz S, Castro-Jalca J. *Comorbilidades asociadas a diabetes mellitus en población adulta.* Conocimiento Global, vol. 9, supl. 1, pp. 234–244, 2024. Available at: https://conocimientoglobal.org/revista/index.php/cglobal/article/view/544

[28] Güleryüz D, Efe F, Efe B. *Estimation of Anthropometric Measurements Using Optimized Machine Learning Models with Bayesia Algorithm*. Scientia Iranica, s/p, 2023. https://doi.org/10.24200/sci.2023.61224.7209

[29] Park M, Kim Y, Lee J. *Machine Learning-Based Model for Grip Strength Prediction in Healthy Adults: A Nationwide Dataset-Based Study*. Journal of Clinical Medicine, vol. 14, no. 5, p. 1542, 2025. https://doi.org/10.3390/jcm14051542

[30] Seto H, Oyama A, Kitora S, Toki H, Yamamoto R, Kotoku J, Haga A, Shinzawa M, Yamakawa M, Fukui S, Moriyama T. *Gradient boosting decision tree becomes more reliable than logistic regression in predicting probability for diabetes with big data*. Scientific Reports. 2022;12:15889. https://doi.org/10.1038/s41598-022-20149-z

[31] Turimov-Mustapoevich D, Kim W. *Machine learning applications in sarcopenia detection and management: a comprehensive survey*. In Healthcare (Vol. 11, No. 18, p. 2483). MDPI. 2023. https://doi.org/10.3390/healthcare11182483

[32] Castillo-Olea C, García-Zapirain A, Soto B, Carballo Lozano C, Zuñiga C. Automatic classification of sarcopenia level in older adults: a case study at Tijuana General Hospital. International Journal of Environmental Research and Public Health, 16(18), 3275. 2019. https://doi.org/10.3390/ijerph16183275

[33] Salinas-Rodríguez A, Palazuelos-González R, Rivera-Almaraz A, Manrique-Espinoza B. *Longitudinal association of sarcopenia and mild cognitive impairment among older Mexican adults*. J Cachexia Sarcopenia Muscle. 2021 Dec;12(6):1848-1859. https://doi.org/10.1002/jcsm.12787

[34] Yuan S, Larsson S. *Epidemiology of sarcopenia: Prevalence, risk factors, and consequences*. Metabolism, 144, 155533. 2023. https://doi.org/10.1016/j.metabol.2023.155533

[35] Kim S, Kim H, Yoo J. *Sarcopenia classification model for musculoskeletal patients using smart insole and artificial intelligence gait analysis*. Journal of Cachexia, Sarcopenia and Muscle, vol. 14, no. 6, pp. 2793–2803, 2023. https://doi.org/10.1002/jcsm.13356

[36] Friedman JH. *Greedy function approximation: a gradient boosting machine*. Ann Stat. 2001;29(5):1189–1232. https://doi.org/10.1214/aos/1013203451

[37] Wen Y, Wan Z, Ren H, Wang X, Wang W. *Interpretable Machine Learning Model for Predicting and Assessing the Risk of Diabetic Nephropathy*. JMIR Medical Informatics, vol. 13, e64979, 2025. https://doi.org/10.2196/64979

[38] Kim J, Jeong K, Lee S, *et al. Machine-learning model predicting quality of life using multifaceted lifestyles in middle-aged South Korean adults: a cross-sectional study*. BMC Public Health, vol. 24, p. 159, 2024. https://doi.org/10.1186/s12889-023-17457-y

[39] Yáñez-Sepúlveda, R., Vásquez-Bonilla, A., Olivares, R. *et al*. *Supervised machine learning algorithms for the classification of obesity levels using anthropometric indices derived from bioelectrical impedance analysis*. Scientific Reports, vol. 15, p. 30681, 2025. https://doi.org/10.1038/s41598-025-15264-6

[40] Karimov S, Turimov D, Kim W, Kim J. *Comparative study of imputation strategies to improve the sarcopenia prediction task*. Digital Health, vol. 11, 20552076241301960. 2025. https://doi.org/10.1177/20552076241301960

[41] Sepúlveda Loyola W, Luna Corrales G, Ganz F, González-Caro H, Suziane Probst V. *Sarcopenia, definición y diagnóstico: ¿Necesitamos valores de referencia para adultos mayores de Latinoamérica?* Revista Chilena de Terapia Ocupacional, vol. 20, no. 2, pp. 259–267, 2020. https://doi.org/10.5354/0719-5346.2020.53583

[42] Blanquet M, Massoulié G, Boirie Y, Guiguet-Auclair C, Mulliez A, Anker S, Eschalier R, *et al*. *Handgrip strength to screen early-onset sarcopenia in heart failure*. Clinical Nutrition ESPEN, vol. 50, pp. 183–190, 2022. https://doi.org/10.1016/j.clnesp.2022.05.019