

Original Paper

Early Identification of Mobility Limitations in Community-Dwelling Middle-Aged and Older Adults: Development of a Prediction Model Based on a Prospective Cohort

Alberto Conde Freniche^{1*}, MSc, MVetMed; Wei Hu^{2*}, MD; Mo Chen¹, MSc; Dantong Wang¹, PhD; Denis Breuillé¹, PhD; Yu-ming Chen², PhD

¹Nestlé Institute of Health Sciences, Nestlé Research, Société des Produits Nestlé, Lausanne, Switzerland

²Department of Epidemiology, School of Public Health, Sun Yat-sen University, Guangzhou, China

*these authors contributed equally

Corresponding Author:

Yu-ming Chen, PhD

Department of Epidemiology

School of Public Health, Sun Yat-sen University

No. 135, Xingang Xi Road

Guangzhou

China

Phone: 86 20-87330605

Email: chenyum@mail.sysu.edu.cn

Abstract

Background: With the aging of the global population, preventing the onset of mobility limitations is considered a worldwide public health priority.

Objective: This study aimed to develop a predictive model for incident early mobility limitations (EMLs) in late middle-aged and older adults, based on a simple functional test and modifiable lifestyle factors to facilitate a home-based self-assessment of early mobility decline and promote lifestyle intervention strategies.

Methods: Our study population was community dwellers aged 45 years and above who participated in the second and fourth waves of the Guangzhou Nutrition and Health Study. The included participants were healthy, nonfrail adults reporting no limitations in activities of daily living at baseline. At the 6-year follow-up, participants with poor physical performance (walking speed <1 m/s or handgrip strength <28 kg for males and <18 kg for females) or reporting some difficulty in walking and/or climbing stairs were classified as experiencing EMLs. Least absolute shrinkage and selection operator (LASSO) was used to identify predictors from various factors, and 6 machine learning models were trained and evaluated for EML prediction, using bootstrap-based techniques to address class imbalance. Predictive ability was quantified using the area under the receiver operating characteristic curve (AUC). Variable importance analysis was used to identify key predictors.

Results: A total of 1344 participants were included in the analysis, of which 206 (15.33%) developed EMLs after a median follow-up of 6.67 (IQR 5.91-7.26) years. Those who developed EMLs were older, had a higher BMI, a lower alternate Mediterranean diet score, and poorer performance in the sit-to-stand (STS) test, as well as lower estimated muscle power from the STS at baseline. The final models included 6 out of 9 predictors: age, sex, BMI, alternate Mediterranean diet score, STS power, and dietary calcium intake. Four machine learning models (logistic regression, LASSO, support vector machine, and neural network) achieved an acceptable AUC value (≥ 0.70) or slightly below in the testing dataset, with the neural network performing the best (AUC 0.70, 95% CI 0.63-0.77). Curve analysis showed positive net benefits for LASSO and logistic regression. Bootstrapping did not improve classification performance. Advanced age, lower adherence to Mediterranean diet, lower muscle power estimated from STS test, and higher BMI at baseline emerged as the most important predictors of EMLs.

Conclusions: This study shows that EMLs in Chinese adults aged 45 years and older can be predicted using easy-to-obtain physical performance measures, age, sex, BMI, and specific nutritional factors. The combination of prediction models and variable importance analysis provides valuable insights for the early identification and intervention of EMLs. More efforts are needed to validate our findings in external cohorts.

Trial Registration: ClinicalTrials.gov NCT03179657; <https://clinicaltrials.gov/study/NCT03179657>

International Registered Report Identifier (IRRID): RR2-10.2188/jea.JE20230108

JMIR Aging 2026;9:e77187; doi: [10.2196/77187](https://doi.org/10.2196/77187)

Keywords: prediction model; prognostic; China; self-assessment; mobility; mobility limitations; sarcopenia; frailty; diet; machine learning; healthy aging

Introduction

Maintaining optimal locomotive capacity—in this study referred to as “mobility”—is essential to successful aging. Optimal mobility determines the capacity to perform activities of daily living (ADLs) and engage in social and leisure activities. However, aging has a negative impact on mobility, underpinned by multiple mechanisms affecting the musculoskeletal, cardiorespiratory, and neurological systems. Along with the onset of chronic conditions, the aging process increases the likelihood of developing mobility limitations, such as difficulties in running, walking, or climbing stairs. These limitations often cause older adults to lose independence, reducing their quality of life and dramatically increasing health care costs. Therefore, with the aging of the global population, preventing the onset of mobility limitations is considered a worldwide public health priority [1,2].

Mobility limitations are often the first signs of further functional decline. The interplay between the emergence of mobility limitations, sarcopenia, and physical frailty involves intricate relationships. Sarcopenia, defined as “a progressive and generalized loss of skeletal muscle mass and strength” [3,4], requires the presence of at least 1 mobility limitation for its diagnosis, assessed through objective physical performance tests. In turn, sarcopenia is an important component of the physical frailty syndrome [5], where mobility limitations make up 2 out of 5 total diagnostic criteria, together with weight loss, low energy expenditure, and exhaustion [6]. The development of mobility limitations is therefore a critical component at the intersection of sarcopenia and physical frailty, serving as a primary target for interventions. In accordance, the World Health Organization has recently called health care professionals and gerontologists to focus on strategies for maximizing and preserving mobility across the entire life course independent of any specific disease.

The early stage of mobility decline is a window of opportunity for interventions, such as adequate nutrition intake and active exercise, to positively modify the trajectory in later life. Once physical frailty becomes manifest in older adults, it is increasingly challenging to reverse the development of mobility limitations [7]. Notably, the progressive age-related mobility decline may not clinically manifest at early stages because during the aging process, people unconsciously modify the way tasks are performed routinely to compensate for the accumulating physiological deficits [8]. For example, older adults typically increase the step width to preserve their balance while walking [9]. This strategy is maintained until the capacity reserve is exhausted and limitations become manifest. Therefore, to prevent mobility limitations, it is important to identify people who do not yet manifest them but who are at risk for their onset

in the near future [10]. A recent scoping review and expert consensus supports the use of the term preclinical or early mobility limitation (EML), referring to such subtle and often unconscious changes in performance [11].

Models to predict EMLs in late middle-aged adults are scarce. Most of the models available were developed to predict mobility disability in older adults based on common physical performance measures, such as gait speed, balance tests, and handgrip strength [12-15]. However, these measures have poor sensitivity to detect mobility deficits in younger adults (under 65 y) [16,17]. Notably, measuring muscle power may increase the sensitivity of prognostic models compared to more commonly used muscle strength measures, such as handgrip strength. Muscle power declines at an earlier and faster rate than muscle strength and is likely to be a more discriminant predictor of mobility limitations [18,19]. Specifically, sit-to-stand (STS) transitions have been shown to decline about 2 decades earlier than handgrip strength and gait speed [16]. Therefore, it is possible that new predictive models built with power-based measures can improve their prognostic value among middle-aged adults. Recently, a formula to calculate muscle power from the STS test has been developed and validated, overcoming the difficulties of evaluating muscle power without any instrument [20,21].

Similarly, the importance of nutrition in preserving mobility has been studied [22,23]. Including nutrition elements in the prediction model would enable a holistic intervention approach to reduce the risk of mobility limitation. To date, few studies have investigated the association between muscle power-based physical performance measures or dietary factors and the development of EML among late middle-aged adults.

In the present study, we aimed to develop and internally validate a model to predict future preclinical mobility limitations in late middle-aged and older adults for patient self-assessment, based on simple-to-obtain measurements, including self-reported data and the STS test, using a cohort of healthy adults aged 45 years and above recruited in a community-based study in the south of China. This model could be used to identify which persons may benefit from preventive strategies at middle and older age to decelerate or delay the onset of mobility limitations in the Chinese population.

Methods

Guidelines for Model Development and Validation

We adhered to the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or

Diagnosis) guidelines, which provide transparent reporting standards for developing and validating prediction models for individual prognosis or diagnosis [24,25].

Study Design and Participants

This retrospective cohort study used data from the Guangzhou Nutrition and Health Study (GNHS), which is a longitudinal community-based cohort conducted by the School of Public Health, Sun Yat-sen University, Guangzhou, China. Between 2008 and 2013, a total of 4048 participants aged between 40 and 80 years, who lived in Guangzhou city for at least 5 years, were enrolled in the GNHS and have been followed up every 3 years approximately. Detailed information about the cohort protocol and baseline characteristics can be found elsewhere [26]. Briefly, data were collected from study visits from healthy participants, including physical performance tests, body composition measurements, dietary, and lifestyle factors. A validated semiquantitative Food Frequency Questionnaire was used to collect dietary intake information [27]. Exclusion criteria such as myocardial infarction, stroke, diabetes, cancer, low-energy fractures, mental and physical disability, and chronic liver and kidney diseases were confirmed at the first enrollment.

Participants in the current analysis participated in 2 follow-ups of the GNHS (waves 2 and 4). At wave 2, from October 2013 to July 2017, 2970 participants were asked to take the 5-time STS test. Therefore, this study used wave 2 as the baseline. In addition, participants completed the ADL Scale questionnaire, grip strength test, and 4-minute walking test in wave 2 and subsequent follow-ups. The latest follow-up with available ADL and performance data was selected to allow EMLs to develop. At wave 4, from October 2020 to December 2023, 1952 participants had been followed-up.

Ethical Considerations

Written informed consent was obtained from all participants. The GNHS was approved by the Ethics Committee of the School of Public Health, Sun Yat-sen University (approval code: SYSU-GWYL2018048; approval date: July 17, 2018) and registered at ClinicalTrials.gov (NCT03179657). All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the World Medical Association Declaration of Helsinki. Data were anonymized and were accessible only to the research team. No financial compensation was provided.

Predicted Outcome

Mobility limitations were assessed using objective mobility tests and perceived difficulties by ADL questionnaire. Participants rated their ease at walking and climbing stairs as independent, have some difficulties, need help, and unable to mobilize. In the predictive model, EML was determined for those individuals meeting at least 1 of the following 3 criteria: (1) some difficulty in walking and/or climbing stairs; (2) poor handgrip strength (handgrip strength <28 kg for males and <18 kg for females); and (3) slow walking speed (<1 m/s).

Validated preclinical thresholds for middle-aged adults were lacking at the time of the study. Gait speed <1 m/s is a common threshold of slow walking speed among well-functioning older adults, and is often treated as a functional benchmark in middle-aged and older adults [28-31]. This threshold is recommended by the Asian Working Group on Sarcopenia (AWGS 2019) to screen for possible sarcopenia and has been applied to cohorts that include midlife participants across Asian countries, including China [32,33]. Similarly, for low handgrip strength, the threshold proposed by the AWGS 2019 was used in this study.

Baseline Candidate Predictors

The STS test is used to assess functional lower limb strength, transitional movements, balance, and fall risk in older adults. STS test score is based on the amount of time (to the nearest decimal in seconds) a participant is able to transfer from a seated to a standing position and back to sitting 5 times. The equipment needed to perform the STS test includes a stopwatch and standard height chair with straight back (42 cm high). Estimated STS power is calculated based on the results of the STS test, using the subject's height and BMI and the height of the chair.

In addition to STS power, other prefrailty phenotype-related predictors were considered, such as age, sex, BMI, alternate Mediterranean diet score (aMDS), energy-adjusted daily intake of vitamin D, protein, and calcium (Ca), and the use of vitamin supplements. Age, gender, and vitamin supplement use were self-reported by participants. BMI was measured by physical examination, the details of which are described in our previous study [26]. Nutrient intake levels were estimated based on a semiquantitative Food Frequency Questionnaire. The adherence to the alternate Mediterranean diet was measured by aMDS, calculated by summing the dichotomous points for the items of higher intakes of whole grain, vegetables, fruits, legumes, nuts, fish, and the ratio of monounsaturated fatty acids:saturated fatty acids, lower red meat, and moderate ethanol consumption, as described previously [34].

Finally, a least absolute shrinkage and selection operator (LASSO) regression model was built to screen the predictors for machine learning (ML) model. Feature selection was performed only in the training dataset. After feature selection, 6 variables (age, sex, BMI, aMDS, STS power, and dietary Ca intake) were selected from 9 predictors to construct a prediction model (Figure S1 in [Multimedia Appendix 1](#)).

Statistical Analysis

The continuous measurement data were represented as mean (SD). Student *t* test was used to compare differences between the EML and the normal (no EML) groups. The chi-squared test was used to compare categorical variables. Data were analyzed using R software (version 4.1.0; R Foundation for Statistical Computing).

ML Stages

The data were preprocessed before the models were trained. The predictors exhibit nonzero variance, and there is no

significant correlation between them (Pearson correlation coefficient <0.75). Subsequently, the original dataset was randomly divided into a training dataset with 70% (940/1344) of the sample (796 normal and 144 EMLs) and a testing dataset with 30% (404/1344) of the sample (342 normal and 62 EMLs). Some predictor variables had a percentage of missing data of $<10\%$ (Figure S2 in [Multimedia Appendix 1](#)). For missing data, the training dataset was imputed with the random forest (RF) method by “mice” package [35]. Missing values of the testing dataset were imputed using median (for numerical) and mode (for categorical), with imputation parameters estimated from the training dataset to avoid leakage. This study performed variable (categorical and numerical) standardization on both the training and testing datasets.

Class imbalance is a common challenge encountered in medical datasets, and it can significantly impact the performance of classification models. The presence of imbalanced classes often leads to overfitting of the majority class data by learning algorithms, resulting in decreased classification performance [36]. Sampling is a widely used technique for adjusting the size of training datasets, and undersampling and oversampling are common approaches [37]. To deal with binary classification problems in the presence of imbalanced classes, synthetic balanced samples were generated according to “ROSE (Random Over-Sampling Examples)” package [38]. ROSE is a bootstrap-based technique that aids the task of binary classification in the presence of rare classes. We combined oversampling and undersampling to rectify the imbalance in data classification. The distribution of the observations by category-wise in the balanced and imbalanced training datasets is shown in Figure S3 in [Multimedia Appendix 1](#).

After the work of preprocessing, the imbalanced and balanced datasets were divided into 5 random parts in the cross-validation test, which was repeated 10 times to randomize the training dataset and improve model performance. In this stage, the models were manually adjusted during training and selected based on the best value of the receiver operating characteristic (ROC), specificity, and sensitivity. The values of these metrics and the parameters

used are shown in Table S1 in [Multimedia Appendix 1](#). Calibration curve and decision curve analyses (DCA) were used to evaluate the calibration and clinical utility of the models. The following 6 ML models were trained: logistics regression (LR), RF, support vector machine (SVM), neural network (NNET), lasso regression (LASSO), and Light Gradient-Boosting Machine (LightGBM).

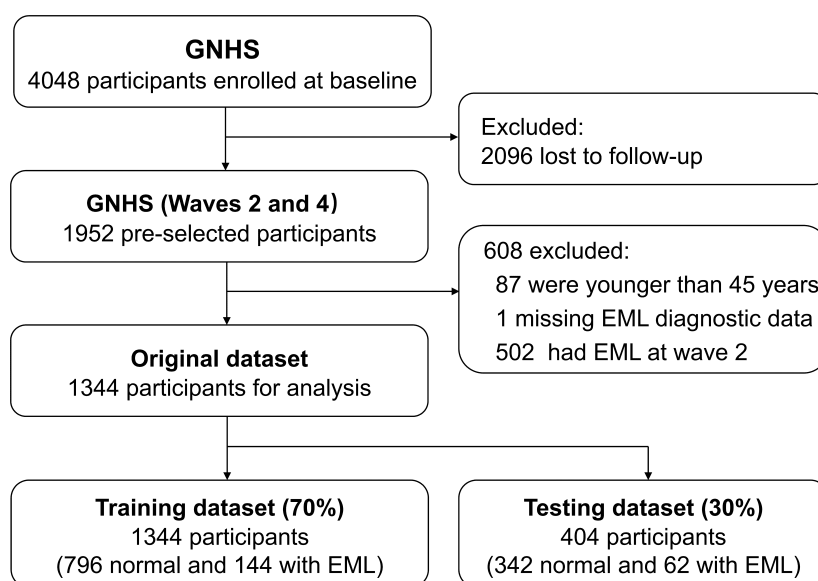
Finally, the testing dataset was used to evaluate the performance of the models trained on the imbalanced and balanced datasets, to assess how well these models fitted the data that was not used during the training stage into the testing dataset. The ROC and precision-recall curves and the metrics area under the receiver operating characteristic curve (AUC), F_1 -score, accuracy, balanced accuracy, specificity, sensitivity, positive predictive value (PPV), and negative predictive value were used to assess the performance of the 6 models. The Youden index was used to determine the optimal threshold for the prediction of EML by optimizing specificity and sensitivity. The “tidymodels” package was used in the ML stages.

The Shapley Additive Explanations (SHAP) analysis was used to interpret the model and assess the marginal contribution of each feature to the model output. The waterfall of individualized SHAP was used to interpret single-sample data, illustrating the direction and strength of influence that an individual’s variable values have on the predictive outcome.

Results

Population Characteristics

In the final analysis, 1344 participants attended the follow-up and were included. At baseline, the median age was 61.99 (IQR 59.03-65.68) years, and most participants were women (945/1344, 70.3%). After a median follow-up of 6.67 (IQR 5.91-7.26) years, 15.33% (206/1344) participants developed EMLs at 6-year follow-up ([Figure 1](#)). Those who developed EMLs were older, had a higher BMI, a lower aMDS, and poorer performance in the STS test, as well as lower estimated muscle power from the test (STS power) at baseline ([Table 1](#)).

Figure 1. Flowchart of the machine learning stages in this study. EML: early mobility limitation; GNHS: Guangzhou Nutrition and Health Study.**Table 1.** Baseline candidate variables selected for machine learning by mobility limitations at follow-up in the dataset.^{a,b,c,d}

Baseline variables	Preclinical mobility limitations (EML) ^e			P value
	Total (N=1344)	No EML (n=1138)	EML (n=206)	
Age, y, mean (SD)	62.56 (5.14)	62.21 (5.03)	64.46 (5.33)	<.001
Sex, n (%)				.29
Female	945 (70.31)	807 (70.91)	138 (66.99)	
Male	399 (29.69)	331 (29.09)	68 (33.01)	
BMI (kg/m ²), mean (SD)	23.38 (3.17)	23.30 (2.99)	23.84 (4.02)	.02
Protein (g/d), mean (SD)	69.07 (10.50)	69.15 (10.54)	68.66 (10.31)	.54
Vitamin D (IU/d), mean (SD)	83.11 (57.96)	84.03 (59.89)	78.07 (45.62)	.17
Ca (mg/d), mean (SD)	547.46 (164.82)	548.81 (158.83)	540.05 (194.89)	.48
aMDS ^f , mean (SD)	4.11 (1.71)	4.15 (1.71)	3.86 (1.69)	.02
Vitamin supplementation, n (%)				>.99
No	1117 (83.11)	946 (83.13)	171 (83.01)	
Yes	227 (16.89)	192 (16.87)	35 (16.99)	
STS ^g power (W), mean (SD)	268.41 (92.38)	271.66 (94.40)	250.49 (78.12)	.002
STS (s), mean (SD)	7.68 (1.88)	7.62 (1.92)	8.01 (1.67)	.003

^aContinuous variables are presented as mean (SD), and categorical variables are presented as n (%).

^bStudent *t* tests were used to compare the mean differences between no EML and EML groups.

^cChi-square tests were used to count data.

^dStatistics after imputation of missing data.

^eEML: early mobility limitation.

^faMDS: alternate Mediterranean diet score.

^gSTS: sit-to-stand test.

Models' Performance

Figure 2 shows the AUC values for the 6 ML models trained with imbalanced (AUC range 0.62-0.70; Figure 2A) and balanced (AUC range 0.56-0.69; Figure 2B) datasets. All 6 ML models selected all 6 variables (age, sex, BMI, aMDS, STS power, and dietary calcium intake) with different

importance (Figure 3), except the LightGBM model for which age, BMI, and dietary calcium intake were not selected. The performance of most models on unseen data remained stable or decreased when the balanced dataset was used, and the LR model maintained the best performance across both datasets.

Figure 2. Receiver operating characteristic curves of the models trained with imbalanced (A) and balanced (B) datasets. AUC: area under the receiver operating characteristic curve; LASSO: Lasso regression; LightGBM: Light Gradient-Boosting Machine; LR: logistic regression; NNET: neural network; RF: random forest; SVM: support vector machine.

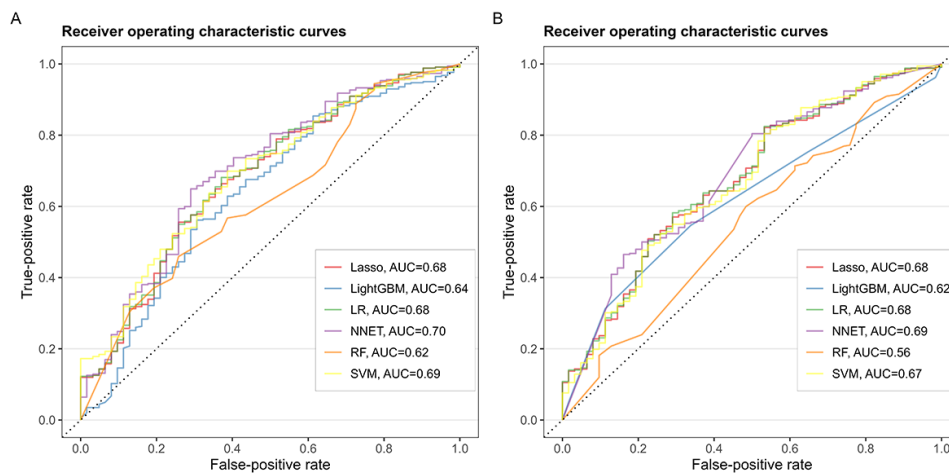


Figure 3. The SHAP for feature importance of the models trained with balanced datasets: (A) logistic regression; (B) random forest; (C) support vector machine; (D) neural network; (E) least absolute shrinkage and selection operator regression; and (F) Light Gradient-Boosting Machine. aMDS: alternate Mediterranean diet score; SHAP: Shapley Additive Explanations; STS: sit-to-stand test.

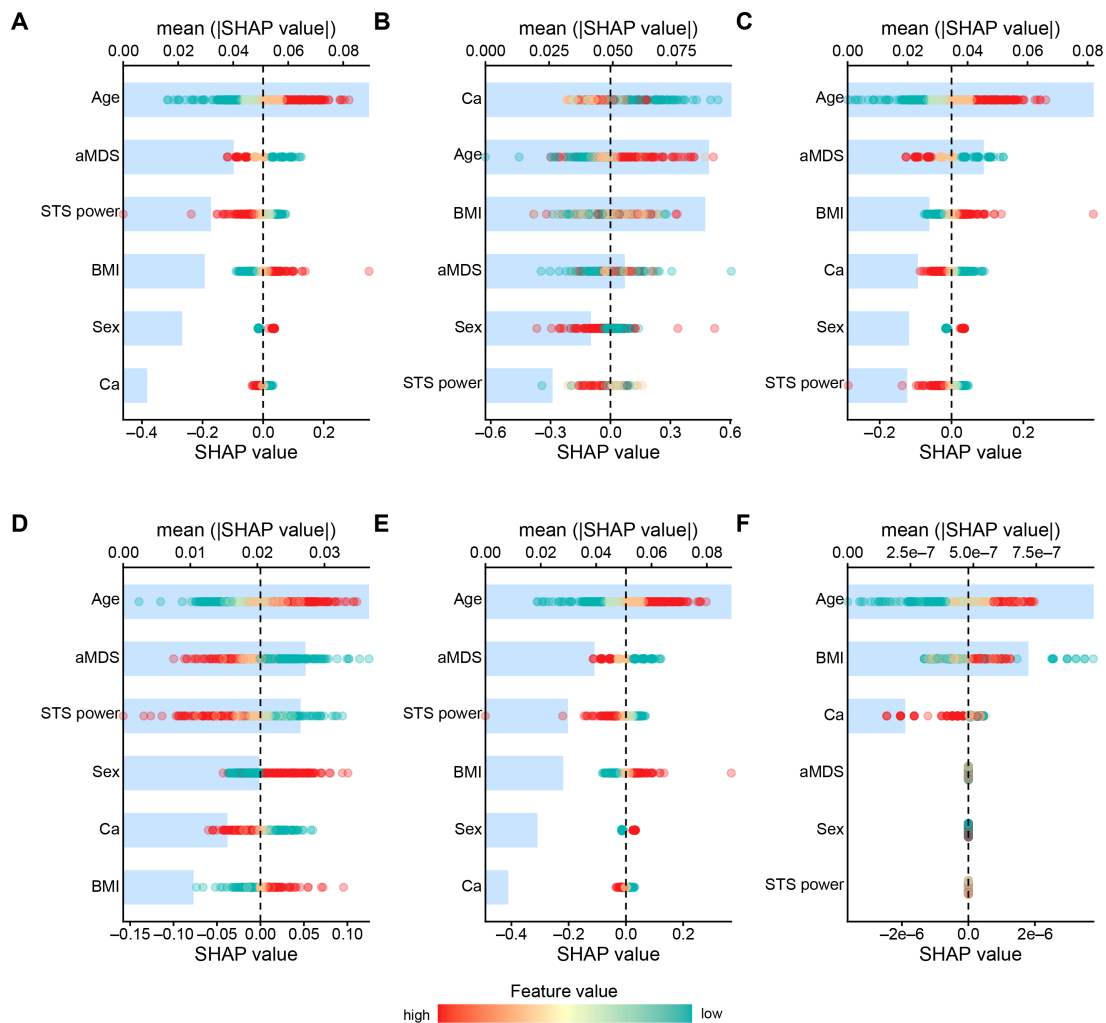


Table 2 shows the performance of the ML models on unseen data, trained with imbalanced and balanced datasets, based on different metrics. Some models (RF, SVM, and NNET) yielded better values for metric accuracy when they were

trained with balanced data, whereas the unbalanced dataset showed better performance in all the other metrics and the rest of the models. LR, NNET, SVM, and LASSO had acceptable values for AUC (≥ 0.70) or slightly below when

trained with imbalanced data, among which NNET had the best values for AUC (0.70; 95% CI 0.63-0.77) and balanced accuracy (0.68) compared with all the other models trained with imbalanced data. However, NNET models are typically more complex and prone to overfitting. Hence, in conditions

of similar performance, simpler models like LR or LASSO may be preferred due to their interpretability and lower risk of overfitting. Specifically, the straightforward coefficients in LR can be easily understood in terms of odds ratios.

Table 2. Performance of the models trained with imbalanced and balanced datasets based on different metrics.^a

Models	Accuracy	Balanced accuracy	F_1 -score	PR ^a -AUC ^b	AUC (95% CI)	Threshold	Specificity	Sensitivity	PPV ^c	NPV ^d
Imbalanced dataset										
LR ^e	0.67	0.66	0.92	0.29	0.68 (0.61-0.76)	0.16	0.68	0.63	0.26	0.91
RF ^f	0.50	0.60	0.91	0.23	0.62 (0.55-0.70)	0.08	0.46	0.74	0.20	0.91
SVM ^g	0.69	0.66	0.92	0.27	0.69 (0.62-0.76)	0.15	0.70	0.61	0.27	0.91
NNET ^h	0.66	0.68	0.92	0.29	0.70 (0.63-0.77)	0.32	0.65	0.71	0.27	0.92
LASSO ⁱ	0.58	0.65	0.92	0.29	0.68 (0.61-0.76)	0.14	0.56	0.74	0.23	0.92
LightGBM ^j	0.58	0.62	0.92	0.23	0.64 (0.56-0.72)	0.18	0.56	0.69	0.22	0.91
Balanced dataset										
LR	0.77	0.64	0.69	0.28	0.68 (0.60-0.75)	0.60	0.82	0.47	0.33	0.90
RF	0.59	0.56	0.74	0.16	0.56 (0.48-0.63)	0.43	0.60	0.52	0.19	0.87
SVM	0.75	0.64	0.71	0.29	0.67 (0.60-0.74)	0.59	0.80	0.47	0.30	0.89
NNET	0.76	0.65	0.66	0.24	0.69 (0.62-0.76)	0.52	0.80	0.50	0.32	0.90
LASSO	0.77	0.64	0.69	0.28	0.68 (0.60-0.75)	0.60	0.82	0.47	0.32	0.90
LightGBM	0.56	0.60	0.92	0.17	0.62(0.55-0.69)	0.50	0.55	0.66	0.21	0.90

^aPR: precision-recall.

^bAUC: area under the receiver operating characteristic curves.

^cPPV: positive predictive value.

^dNPV: negative predictive value.

^eLR: logistic regression.

^fRF: random forest.

^gSVM: support vector machine.

^hNNET: neural network.

ⁱLASSO: least absolute shrinkage and selection operator.

^jLightGBM: Light Gradient-Boosting Machine.

DCA (Figure S8 in [Multimedia Appendix 1](#)) results showed that prediction model's effective range for test datasets are concentrated at low thresholds. All prediction model curves demonstrate positive net benefits exceeding the "no treatment" baseline only within the threshold range of 0% to 50%. When the threshold exceeds 50%, the model's net benefit drops below zero, losing clinical value.

In calibration curves (Figures S5 and S7 in [Multimedia Appendix 1](#)), LASSO regression and logistic regression exhibit relatively good calibration on imbalanced test datasets. However, in balanced datasets, most predictive models show that the predicted probabilities are higher than the actual probabilities, indicating that the models are overconfident in their predictions. For implementation in the real world, prediction models need to be further calibrated according to the specific application scenario.

In addition to discrimination metrics, we assessed model performance using the net reclassification improvement. Logistic regression was selected as reference because it represents a widely used baseline approach for binary risk prediction in preventive settings. Net reclassification improvement was decomposed into event and nonevent components to better understand potential gains and tradeoffs in risk stratification across modeling approaches, and the results are reported separately for models trained on

imbalanced and balanced datasets (Table S1 in [Multimedia Appendix 1](#)). Only RF achieved a modest improvement in risk reclassification over logistic regression, whereas other ML models did not demonstrate meaningful net reclassification gains.

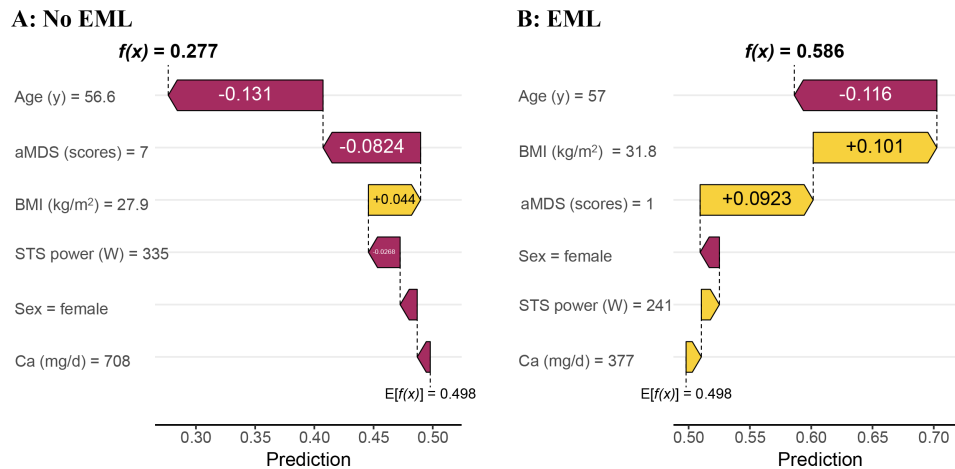
Model Interpretability

The SHAP model was used to provide both global and local interpretations of the models. Globally, the more significant contributors to the models trained with the balanced dataset were age, aMDS, STS power, and BMI. Age was among the top 3 most important variables in all 6 models, followed by aMDS in 4 models (LR, SVM, NNET, and LASSO) and STS power (LR, NNET, and LASSO) and BMI (RF, SVM, and light GBM) in 3 models ([Figure 3](#)).

In the single-sample waterfall plot, baseline SHAP value was 0.498, model output values for a randomly selected individual with no EML was 0.277 ([Figure 4A](#)), and for an EML individual was 0.586 ([Figure 4B](#)). In no EML individuals, lower age as well as higher aMDS scores and STS power were identified as features that had a negative impact on EML, while higher BMI was considered to have a positive impact ([Figure 4A](#)). Among EML individuals, lower age was identified as a feature that negatively affected EML, while

higher BMI and lower aMDS scores were considered to have a positive impact (Figure 4B).

Figure 4. Individualized SHAP waterfall plot for no EML (A) and EML participants (B), demonstrating the personalized contribution of each predictor to the risk score. aMDS: alternate Mediterranean diet score; EML: early mobility limitation; SHAP: Shapley Additive Explanations; STS: sit-to-stand test.



The SHAP bar plot ranks features by their overall importance, calculated as the mean absolute SHAP value across all participants.

The SHAP beeswarm plot illustrates the impact of each feature on the model's output for every participant. Each dot represents a participant; its position on the x -axis indicates the SHAP value (positive values increase the prediction of EML), and its color represents the feature's value (red for high, green for low).

Discussion

Principal Results

In this study, we developed 6 prediction models using ML for EML in healthy, mobility-intact middle-aged and older adults from a longitudinal Chinese population-based cohort who were followed-up over 6.67 years. A LASSO regression model was built to select the predictors. Finally, 6 variables (age, sex, BMI, aMDS, STS power, and dietary calcium intake) were identified and included in the model training process. The performance of the different ML models trained with balanced and imbalanced datasets was compared. Advanced age, lower adherence to Mediterranean diet, lower STS power, and higher BMI were identified as the most important predictors of EMLs.

Comparison to Prior Work

In this study, we have attempted to use commonly used ML models for binary variable prediction, including LR, RF, SVM, NNET, LASSO, and LightGBM. In a recent study with a large sample size (5466), ML models such as SVM, K-nearest neighbor, NNET, RF, and extreme gradient boosting demonstrated good results in predicting frailty among senior people in Canada [39]. Although our database may not have a comparable size, we were interested in

exploring the feasibility of applying these ML models to the cohort in Guangzhou, China.

The discrimination power of 4 models (LR, NNET, SVM, and LASSO) trained with the imbalanced data was found acceptable ($AUC \geq 0.70$) or slightly below in the testing dataset. The performance moderately decreased when trained with the balanced dataset, with NNET (0.69), LR (0.68), SVM (0.67), and LASSO (0.68) achieving the highest AUC values and, therefore, showing more robust discrimination across datasets (Table 2). Royston et al [40] noted that the AUC of prognostic models is typically between 0.6 and 0.85. The commonly used prognostic tests include the ratio of total plasma cholesterol to high-density lipoprotein cholesterol level, with an AUC of 0.72 for predicting the risk of coronary heart disease [41], and the 6-minute walking test, with an AUC of 0.60 for predicting increased risk of hospitalization among patients with chronic obstructive pulmonary disease [42]. Therefore, the predictive power of our models compares well with that of risk assessment tests widely used in clinical practice that are more invasive (eg, TC/HDL-C from a blood test) or burdensome (eg, 6-minute walking test requiring time and space). Additional DCA of the models trained with imbalanced datasets showed meaningful positive net benefits for LASSO and LR only, supporting the clinical utility of these 2 models within the threshold range of 0% to 35%, but no positive net benefits were found for the balanced trained models (Figure S8 in Multimedia Appendix 1). The incremental predictive benefit of other ML models compared to LR did not demonstrate meaningful net reclassification gains.

Similar ROC-AUC results were obtained by Chaves et al [43], who developed a screening tool from simple-to-obtain self-reported performance measures for predicting the probability of 18-month incident mobility difficulty in community-dwelling women aged 70 to 80 years, with an AUC of 0.73. However, a test is likely to be more sensitive when the prevalence of concomitant risk factors is higher,

as it would be expected from older individuals such as those included in that study [44]. Likewise, as the time to event decreases, the predictive accuracy of a model tends to improve [45], leading to less uncertainty in short-term predictions like in the study by Chaves et al [43]. Taken together, considering the challenges of predicting the onset of the early stages of an unspecific and multifactorial condition such as EML in mobility-intact adults who are generally healthy, drawing mostly from modifiable predictors, this study yielded acceptable results for the LASSO and LR imbalanced trained models, with positive net benefits. Given the models' intended use for early identification of individuals who may benefit from low-risk lifestyle interventions, their limited discrimination remains potentially valuable.

The models developed in this study were designed to maximize clinical applicability. All 6 predictors (age, sex, BMI, STS power, aMDS, and dietary calcium intake) can be easily collected by individuals using a standard chair, a stopwatch, and a short diet questionnaire. These inputs could be incorporated into a digital self-assessment platform or mobile health tool to support interpretation and reporting. Because the predictors include modifiable factors such as diet quality, micronutrient intake, physical function, and body weight, such a tool could help identify individuals at higher risk of mobility decline and provide personalized lifestyle recommendations. The decision threshold can be implemented within preventive care workflows, with thresholds in the range of 10% to 20%—as supported by DCA—offering a balance between early intervention benefits and the likelihood of false positives. This approach may empower individuals to monitor and preserve their mobility health without requiring specialized training or extensive resources.

As expected, age was the most important predictor of incident EML across the different ML models. Advanced age is a risk factor in developing mobility limitations and has been consistently identified as a predictor of mobility outcomes in many predictive equations [46,47]. Therefore, the finding that advanced age was the most important variable across the different ML models as determined by SHAP values (Figure 3) served as a positive control variable in developing our models to predict EML. Other important variables were adherence to the Mediterranean diet, STS power, and BMI. Among the rest of the baseline candidate predictors selected by LASSO regression, other less important variables across the different ML models included being female and lower energy-adjusted calcium intake. Many of these variables confirm some of the risk factors most frequently found to be independently linked to mobility limitations in older people, which include advanced age, higher BMI, and reduced muscle function measured through performance tests [48]. However, our results cast a new light on the predictive ability of estimated muscle power from STS and diet quality.

This study found support for STS power as an important predictor of incident EML in older, as well as in late middle-aged adults. The STS test has been consistently identified as a predictor of subsequent falls, mobility limitations [49], disability [14], osteoporotic hip fractures,

and mortality among older adults. Remarkably, we found a higher statistically significant difference and relative difference for STS power (7.79%, $P=.002$) than for STS time values (4.87%, $P=.003$) at baseline between the group developing EML and their counterparts (Table 1). Therefore, only STS power was included among the baseline candidate predictors. Whereas the STS test measures the minimum duration required to rise from a chair a specific number of times (eg, 5-time STS) or the maximum number of STS repetitions within a specified timeframe (eg, 30-s STS), the STS power formula also considers the chair height, and the subject's body mass and height. Previous evidence in older people suggests that STS power values are more clinically relevant in regard to physical function, cognitive function, and sarcopenia compared to traditional STS measures [20]. Consistent with these findings in older adults, our results extend the predictive value of STS power to identify prospective mobility limitations in late middle-aged adults.

Notably, although normative values for STS power in Asian populations are lacking, the 5-time STS values observed in our study suggest a rather optimal level of physical function in both the EML group (mean 8.01, SD 1.67 s) and the normal group (mean 7.62, SD 1.92 s) at baseline. Specifically, current guidelines from the AWGS 2019 recommend a cut-off of 12 seconds or more to screen for low physical performance and sarcopenia, based on a review of evidence from different Asian cohorts. The striking distance between that cut-off and the baseline functional performance observed in our cohort underlines the ability of our models to identify incident mobility limitations earlier in time, even in the absence of noticeable performance decline. Remarkably, a recent study redefined the cut-off value recommended by the AWGS to 9.2 seconds in men and 10.8 seconds in women aged 60 to 69 years, significantly improving the detection rate of sarcopenia among Chinese adults [50]. These lower thresholds tie well with the values observed in our study and further support the predictive value of subtle deficits in performance tests.

A further novel finding was the importance of diet quality, measured as adherence to the Mediterranean diet, to predict incident EML. Studying the influence of dietary factors on mobility limitation has practical importance, and the findings could provide opportunities for dietary intervention to reduce the potential risk. In this study, we identified adherence to the Mediterranean diet as a predictor of mobility limitations, with a higher adherence score associated with a lower probability of having mobility limitations. This finding is consistent with the previous studies. The aMDS was found positively associated with skeletal muscle mass index, which is an important aspect of mobility, in the Chinese population aged 40 to 75 years [34]. Silva et al [51] conducted a meta-analysis in community-dwelling older people and found that higher adherence to the Mediterranean diet was inversely associated with frailty. The anti-inflammatory and antioxidant properties of some nutrients typically found in higher amounts in Mediterranean diets, such as vitamin C and vitamin E or carotenoids, could contribute to that benefit [52]. Future

intervention studies are needed to provide direct evidence on the risk reduction of mobility limitation by adherence to aMDS and improvement of calcium intake among individuals aged 45 years and older.

Strengths and Limitations

Although many prediction models for mobility limitations or disability have been developed in older people with poorer baseline health and functional performance, to our knowledge, comparable models are lacking for mobility-intact adults under 65 years of age. Therefore, a major strength of the study is the sample selection of healthier and younger participants, allowing for the observation of the early transition from intact mobility to impairment. This design facilitates the identification of early predictors and risk factors of mobility limitations, which might be overlooked in studies restricted to older or already impaired individuals.

Another positive aspect of the study is the selected set of predictor variables, which can be easily self-assessed at home with a smartphone and comprises several modifiable risk factors such as diet, physical function, and weight. As a result, the information collected for the prognostic test could be used for guiding potential intervention strategies. Finally, despite the common practice among ML researchers of using extensive datasets and training models with numerous predictors, this research has shown the feasibility of applying ML in datasets with a reasonable number of predictors and in cases where the classes are initially unbalanced.

There are some limitations of the study. First, the composite outcome EML measured in this research integrated both patient-reported and performance-based measures, as recommended by Richardson et al [11]. The expert group conceptualized the term *early* or *preclinical* mobility limitations to capture subtle decrements in mobility; however, a standardized measure has not yet been established. Further research examining the reliability, validity, and responsiveness of outcome measures is necessary to strengthen the evidence base on preclinical mobility limitation.

Second, the PPV, which indicates the likelihood that a positive test result is a true positive, was not optimal across models. These results should be interpreted considering the relatively low prevalence of EML observed in our cohort (15.3%) compared to that reported in other studies with similar operationalizations of EML among Chinese (27.9%) [53]. Because predictive values are prevalence-dependent, this may have increased the proportion of false positives, lowering the PPVs of our models [54]. This prevalence discrepancy may stem from the high proportion of lost-to-follow-up participants in our study, over 50%. The cohort in this study was followed in a designated location, introducing attrition bias for participants who were more frail, developed disability, or moved to care facilities. To address this discrepancy, we recalculated the PPV using the prevalence in the target population (27.9% in China) for the LR, SVM, NNET, and LASSO models. The PPV increases to

approximately 43% to 44% indicating a substantially higher probability that individuals classified as high risk will truly experience the outcome when the model is applied in the intended population. In addition, proper calibration of the models may further optimize the PPV [55]. Notably, in the context of rule-in tests, it is crucial to balance the benefits of correctly identifying at-risk individuals against the potential harm of false positives [56]. The prognostic models developed in this study are aimed at promoting healthy lifestyle advice, like increased physical activity and improved nutrition, which are particularly relevant for the intended at-risk population. Therefore, the limited PPVs observed in our sample shall not limit the applicability of some of our models to identify individuals who may benefit the most from low-risk preventive interventions in the target Chinese population. In particular, a consistent net benefit was shown in DCA for SVM and LR across threshold probabilities roughly in the 10% to 35% band.

Another major limitation of the study was the validation of the models with an internal testing dataset. Further studies on independent cohorts need to be undertaken to establish the validity of our findings. Lastly, we suffered from an imbalanced dataset with only 206 subjects developing EML out of the 1344 sample. Similar imbalance problems have been observed in related studies [37], and this imbalance can lead to biased estimates of training performance [36]. This is because algorithms always tend to predict negative outcomes to achieve higher accuracy, which can also bring implementation challenges, such as when the primary interest lies in discovering the predictive results for EML, and biased predictions render the algorithm meaningless. However, in this study, we used the ROSE bootstrap-based technique [38] to remedy the imbalanced data and generate synthetic balanced samples, leading to more reliable and consistent evaluation of our models across 2 datasets.

Conclusions

This study is the first, to the authors' knowledge, to develop a prediction model for identifying the onset of EMLs in middle-aged and older adults based exclusively on at-home, simple-to-obtain measurements. The evaluation of model performance on a testing dataset, considering both imbalanced and balanced data training, indicates the potential utility of the ML approach in identifying individuals who developed EMLs over time. LASSO and logistic regression yielded the most meaningful positive net benefits, supporting their potential clinical utility for early identification of individuals who may benefit from low-risk lifestyle interventions, though external validation is still required. Across models, age, adherence to the Mediterranean diet, STS power, and BMI consistently emerged as the most important predictors. These models provide an initial foundation for developing a digital self-assessment tool to identify individuals at higher risk of mobility decline and support personalized preventive strategies in clinical and at-home settings.

Acknowledgments

The authors thank Fred (JZ) and Yu Kai for their work as scientific advisors.

Funding

The original cohort study was jointly supported by the Nutrition Research Foundation of the Chinese Nutrition Society (CNS-ZD2019049), National Natural Science Foundation of China (82574079, 82073533, and the 5010 Program for Clinical Researches (2007032) of Sun Yat-sen University (Guangzhou, China). No generative artificial intelligence tools were used at any stage in the preparation of this manuscript.

Data Availability

The data of Guangzhou Nutrition and Health Study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: ACF, DB, YC

Methodology: ACF, MC, WH

Data management and analysis: MC, WH

Writing – original draft: ACF

Writing – review & editing: ACF, DB, MC, DW, WH, YC

Conflicts of Interest

ACF, DW, and DB are employees of Société des Produits Nestlé SA. The authors declare no competing interests.

Multimedia Appendix 1

Supplementary figures and tables.

[\[DOCX File \(Microsoft Word File\), 1091 KB-Multimedia Appendix 1\]](#)

References

1. Inouye SK, Studenski S, Tinetti ME, Kuchel GA. Geriatric syndromes: clinical, research, and policy implications of a core geriatric concept. *J Am Geriatr Soc.* May 2007;55(5):780-791. [doi: [10.1111/j.1532-5415.2007.01156.x](https://doi.org/10.1111/j.1532-5415.2007.01156.x)] [Medline: [17493201](https://pubmed.ncbi.nlm.nih.gov/17493201/)]
2. Webber SC, Porter MM, Menec VH. Mobility in older adults: a comprehensive framework. *Gerontologist.* Aug 2010;50(4):443-450. [doi: [10.1093/geront/gnq013](https://doi.org/10.1093/geront/gnq013)] [Medline: [20145017](https://pubmed.ncbi.nlm.nih.gov/20145017/)]
3. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing.* Jan 1, 2019;48(1):16-31. [doi: [10.1093/ageing/afy169](https://doi.org/10.1093/ageing/afy169)] [Medline: [30312372](https://pubmed.ncbi.nlm.nih.gov/30312372/)]
4. Cruz-Jentoft AJ, Sayer AA. Sarcopenia. *Lancet.* Jun 29, 2019;393(10191):2636-2646. [doi: [10.1016/S0140-6736\(19\)31138-9](https://doi.org/10.1016/S0140-6736(19)31138-9)] [Medline: [31171417](https://pubmed.ncbi.nlm.nih.gov/31171417/)]
5. Cesari M, Landi F, Vellas B, Bernabei R, Marzetti E. Sarcopenia and physical frailty: two sides of the same coin. *Front Aging Neurosci.* 2014;6:192. [doi: [10.3389/fnagi.2014.00192](https://doi.org/10.3389/fnagi.2014.00192)] [Medline: [25120482](https://pubmed.ncbi.nlm.nih.gov/25120482/)]
6. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci.* Mar 2001;56(3):M146-M156. [doi: [10.1093/gerona/56.3.m146](https://doi.org/10.1093/gerona/56.3.m146)] [Medline: [11253156](https://pubmed.ncbi.nlm.nih.gov/11253156/)]
7. Kojima G, Taniguchi Y, Iliffe S, Jivraj S, Walters K. Transitions between frailty states among community-dwelling older people: a systematic review and meta-analysis. *Ageing Res Rev.* Mar 2019;50:81-88. [doi: [10.1016/j.arr.2019.01.010](https://doi.org/10.1016/j.arr.2019.01.010)] [Medline: [30659942](https://pubmed.ncbi.nlm.nih.gov/30659942/)]
8. Ferrucci L, Cooper R, Shardell M, Simonsick EM, Schrack JA, Kuh D. Age-related change in mobility: perspectives from life course epidemiology and geroscience. *J Gerontol A Biol Sci Med Sci.* Sep 2016;71(9):1184-1194. [doi: [10.1093/gerona/glw043](https://doi.org/10.1093/gerona/glw043)] [Medline: [26975983](https://pubmed.ncbi.nlm.nih.gov/26975983/)]
9. Aboutorabi A, Arazpour M, Bahramzadeh M, Hutchins SW, Fadayevatan R. The effect of aging on gait parameters in able-bodied older subjects: a literature review. *Aging Clin Exp Res.* Jun 2016;28(3):393-405. [doi: [10.1007/s40520-015-0420-6](https://doi.org/10.1007/s40520-015-0420-6)] [Medline: [26210370](https://pubmed.ncbi.nlm.nih.gov/26210370/)]
10. Ferrucci L, Guralnik JM, Studenski S, et al. Designing randomized, controlled trials aimed at preventing or delaying functional decline and disability in frail, older persons: a consensus report. *J Am Geriatr Soc.* Apr 2004;52(4):625-634. [doi: [10.1111/j.1532-5415.2004.52174.x](https://doi.org/10.1111/j.1532-5415.2004.52174.x)] [Medline: [15066083](https://pubmed.ncbi.nlm.nih.gov/15066083/)]
11. Richardson J, Beauchamp M, Bean J, et al. Defining and measuring preclinical mobility limitation: an expert consensus exercise informed by a scoping review. *J Gerontol A Biol Sci Med Sci.* Aug 27, 2023;78(9):1641-1650. [doi: [10.1093/gerona/glad143](https://doi.org/10.1093/gerona/glad143)] [Medline: [37300461](https://pubmed.ncbi.nlm.nih.gov/37300461/)]
12. Rantanen T, Guralnik JM, Foley D, et al. Midlife hand grip strength as a predictor of old age disability. *JAMA.* Feb 10, 1999;281(6):558-560. [doi: [10.1001/jama.281.6.558](https://doi.org/10.1001/jama.281.6.558)] [Medline: [10022113](https://pubmed.ncbi.nlm.nih.gov/10022113/)]

13. Cavanaugh EJ, Richardson J, McCallum CA, Wilhelm M. The predictive validity of physical performance measures in determining markers of preclinical disability in community-dwelling middle-aged and older adults: a systematic review. *Phys Ther.* Dec 1, 2018;98(12):1010-1021. [doi: [10.1093/ptj/pzy109](https://doi.org/10.1093/ptj/pzy109)] [Medline: [30256988](https://pubmed.ncbi.nlm.nih.gov/30256988/)]
14. Deshpande N, Metter EJ, Guralnik J, Bandinelli S, Ferrucci L. Predicting 3-year incident mobility disability in middle-aged and older adults using physical performance tests. *Arch Phys Med Rehabil.* May 2013;94(5):994-997. [doi: [10.1016/j.apmr.2012.10.032](https://doi.org/10.1016/j.apmr.2012.10.032)] [Medline: [23164980](https://pubmed.ncbi.nlm.nih.gov/23164980/)]
15. Perera S, Patel KV, Rosano C, et al. Gait speed predicts incident disability: a pooled analysis. *J Gerontol A Biol Sci Med Sci.* Jan 2016;71(1):63-71. [doi: [10.1093/gerona/glv126](https://doi.org/10.1093/gerona/glv126)] [Medline: [26297942](https://pubmed.ncbi.nlm.nih.gov/26297942/)]
16. Suetta C, Haddock B, Alcazar J, et al. The Copenhagen Sarcopenia Study: lean mass, strength, power, and physical function in a Danish cohort aged 20-93 years. *J Cachexia Sarcopenia Muscle.* Dec 2019;10(6):1316-1329. [doi: [10.1002/jcsm.12477](https://doi.org/10.1002/jcsm.12477)] [Medline: [31419087](https://pubmed.ncbi.nlm.nih.gov/31419087/)]
17. Rasmussen LJH, Caspi A, Ambler A, et al. Association of neurocognitive and physical function with gait speed in midlife. *JAMA Netw Open.* Oct 2, 2019;2(10):e1913123. [doi: [10.1001/jamanetworkopen.2019.13123](https://doi.org/10.1001/jamanetworkopen.2019.13123)] [Medline: [31603488](https://pubmed.ncbi.nlm.nih.gov/31603488/)]
18. Reid KF, Fielding RA. Skeletal muscle power: a critical determinant of physical functioning in older adults. *Exerc Sport Sci Rev.* Jan 2012;40(1):4-12. [doi: [10.1097/JES.0b013e31823b5f13](https://doi.org/10.1097/JES.0b013e31823b5f13)] [Medline: [22016147](https://pubmed.ncbi.nlm.nih.gov/22016147/)]
19. Lauretani F, Russo CR, Bandinelli S, et al. Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. *J Appl Physiol (1985).* Nov 2003;95(5):1851-1860. [doi: [10.1152/jappphysiol.00246.2003](https://doi.org/10.1152/jappphysiol.00246.2003)] [Medline: [14555665](https://pubmed.ncbi.nlm.nih.gov/14555665/)]
20. Alcazar J, Kamper RS, Aagaard P, et al. Relation between leg extension power and 30-s sit-to-stand muscle power in older adults: validation and translation to functional performance. *Sci Rep.* Oct 1, 2020;10(1):16337. [doi: [10.1038/s41598-020-73395-4](https://doi.org/10.1038/s41598-020-73395-4)] [Medline: [33004970](https://pubmed.ncbi.nlm.nih.gov/33004970/)]
21. Alcazar J, Losa-Reyna J, Rodriguez-Lopez C, et al. The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. *Exp Gerontol.* Oct 2, 2018;112:38-43. [doi: [10.1016/j.exger.2018.08.006](https://doi.org/10.1016/j.exger.2018.08.006)] [Medline: [30179662](https://pubmed.ncbi.nlm.nih.gov/30179662/)]
22. Mazza E, Ferro Y, Pujia R, et al. Mediterranean diet in healthy aging. *J Nutr Health Aging.* 2021;25(9):1076-1083. [doi: [10.1007/s12603-021-1675-6](https://doi.org/10.1007/s12603-021-1675-6)] [Medline: [34725664](https://pubmed.ncbi.nlm.nih.gov/34725664/)]
23. Shahar DR, Houston DK, Hue TF, et al. Adherence to mediterranean diet and decline in walking speed over 8 years in community-dwelling older adults. *J Am Geriatr Soc.* Oct 2012;60(10):1881-1888. [doi: [10.1111/j.1532-5415.2012.04167.x](https://doi.org/10.1111/j.1532-5415.2012.04167.x)] [Medline: [23035758](https://pubmed.ncbi.nlm.nih.gov/23035758/)]
24. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* Jan 6, 2015;162(1):W1-W73. [doi: [10.7326/M14-0698](https://doi.org/10.7326/M14-0698)] [Medline: [25560730](https://pubmed.ncbi.nlm.nih.gov/25560730/)]
25. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ.* Apr 16, 2024;385:e078378. [doi: [10.1136/bmj-2023-078378](https://doi.org/10.1136/bmj-2023-078378)] [Medline: [38626948](https://pubmed.ncbi.nlm.nih.gov/38626948/)]
26. Ling CW, Zhong H, Zeng FF, et al. Cohort profile: Guangzhou Nutrition and Health Study (GNHS): a population-based multi-omics study. *J Epidemiol.* Jun 5, 2024;34(6):301-306. [doi: [10.2188/jea.JE20230108](https://doi.org/10.2188/jea.JE20230108)] [Medline: [37813622](https://pubmed.ncbi.nlm.nih.gov/37813622/)]
27. Zhang CX, Ho SC. Validity and reproducibility of a food frequency questionnaire among Chinese women in Guangdong province. *Asia Pac J Clin Nutr.* 2009;18(2):240-250. [Medline: [19713184](https://pubmed.ncbi.nlm.nih.gov/19713184/)]
28. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. *JAMA.* Jan 5, 2011;305(1):50-58. [doi: [10.1001/jama.2010.1923](https://doi.org/10.1001/jama.2010.1923)] [Medline: [21205966](https://pubmed.ncbi.nlm.nih.gov/21205966/)]
29. Cesari M, Kritchevsky SB, Penninx B, et al. Prognostic value of usual gait speed in well-functioning older people--results from the Health, Aging and Body Composition Study. *J Am Geriatr Soc.* Oct 2005;53(10):1675-1680. [doi: [10.1111/j.1532-5415.2005.53501.x](https://doi.org/10.1111/j.1532-5415.2005.53501.x)] [Medline: [16181165](https://pubmed.ncbi.nlm.nih.gov/16181165/)]
30. Schrack JA, Zipunnikov V, Simonsick EM, Studenski S, Ferrucci L. Rising energetic cost of walking predicts gait speed decline with aging. *J Gerontol A Biol Sci Med Sci.* Jul 2016;71(7):947-953. [doi: [10.1093/gerona/glw002](https://doi.org/10.1093/gerona/glw002)] [Medline: [26850913](https://pubmed.ncbi.nlm.nih.gov/26850913/)]
31. Makino K, Raina P, Griffith LE, et al. Lifetime physical activity and late-life mild cognitive impairment in community-dwelling older adults. *J Am Med Dir Assoc.* Mar 2024;25(3):488-493.e3. [doi: [10.1016/j.jamda.2023.12.006](https://doi.org/10.1016/j.jamda.2023.12.006)] [Medline: [38246592](https://pubmed.ncbi.nlm.nih.gov/38246592/)]
32. Wu X, Li X, Xu M, Zhang Z, He L, Li Y. Sarcopenia prevalence and associated factors among older Chinese population: findings from the China Health and Retirement Longitudinal Study. *PLoS One.* 2021;16(3):e0247617. [doi: [10.1371/journal.pone.0247617](https://doi.org/10.1371/journal.pone.0247617)] [Medline: [33661964](https://pubmed.ncbi.nlm.nih.gov/33661964/)]

33. Cheong CY, Yap P, Gwee X, et al. Physical and functional measures predicting long-term mortality in community-dwelling older adults: a comparative evaluation in the Singapore Longitudinal Ageing Study. *Aging (Albany NY)*. Dec 2021;13(23):25038-25054. [doi: [10.18632/aging.203756](https://doi.org/10.18632/aging.203756)] [Medline: [34894397](https://pubmed.ncbi.nlm.nih.gov/34894397/)]
34. Tian HY, Qiu R, Jing LP, Chen ZY, Chen GD, Chen YM. Alternate Mediterranean diet score is positively associated with skeletal muscle mass index in middle-aged adults. *Br J Nutr*. Apr 2017;117(8):1181-1188. [doi: [10.1017/S0007114517001118](https://doi.org/10.1017/S0007114517001118)] [Medline: [28514984](https://pubmed.ncbi.nlm.nih.gov/28514984/)]
35. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45(3):1-67. [doi: [10.18637/jss.v045.i03](https://doi.org/10.18637/jss.v045.i03)]
36. Kim MS. An effective under-sampling method for class imbalance data problem. Presented at: Eighth International Symposium on Intelligent Systems (ISIS 2007); Sep 5-8, 2007; Sokcho, South Korea.
37. Li DC, Liu CW, Hu SC. A learning method for the class imbalance problem with medical data sets. *Comput Biol Med*. May 2010;40(5):509-518. [doi: [10.1016/j.combiomed.2010.03.005](https://doi.org/10.1016/j.combiomed.2010.03.005)] [Medline: [20347072](https://pubmed.ncbi.nlm.nih.gov/20347072/)]
38. Lunardon N, Menardi G, Torelli N. ROSE: a package for binary imbalanced learning. *R J*. 2014;6(1):79. [doi: [10.32614/RJ-2014-008](https://doi.org/10.32614/RJ-2014-008)]
39. Aponte-Hao S, Wong ST, Thandi M, et al. Machine learning for identification of frailty in Canadian primary care practices. *Int J Popul Data Sci*. 2021;6(1):1650. [doi: [10.23889/ijpds.v6i1.1650](https://doi.org/10.23889/ijpds.v6i1.1650)] [Medline: [34541337](https://pubmed.ncbi.nlm.nih.gov/34541337/)]
40. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. *BMJ*. Mar 31, 2009;338:b604. [doi: [10.1136/bmj.b604](https://doi.org/10.1136/bmj.b604)] [Medline: [19336487](https://pubmed.ncbi.nlm.nih.gov/19336487/)]
41. Grover SA, Coupal L, Hu XP. Identifying adults at increased risk of coronary disease. How well do the current cholesterol guidelines work? *JAMA*. Sep 13, 1995;274(10):801-806. [Medline: [7650803](https://pubmed.ncbi.nlm.nih.gov/7650803/)]
42. Spruit MA, Polkey MI, Celli B, et al. Predicting outcomes from 6-minute walk distance in chronic obstructive pulmonary disease. *J Am Med Dir Assoc*. Mar 2012;13(3):291-297. [doi: [10.1016/j.jamda.2011.06.009](https://doi.org/10.1016/j.jamda.2011.06.009)] [Medline: [21778120](https://pubmed.ncbi.nlm.nih.gov/21778120/)]
43. Chaves PH, Garrett ES, Fried LP. Predicting the risk of mobility difficulty in older women with screening nomograms: the Women's Health and Aging Study II. *Arch Intern Med*. Sep 11, 2000;160(16):2525-2533. [doi: [10.1001/archinte.160.16.2525](https://doi.org/10.1001/archinte.160.16.2525)] [Medline: [10979066](https://pubmed.ncbi.nlm.nih.gov/10979066/)]
44. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. Feb 20, 2007;115(7):928-935. [doi: [10.1161/CIRCULATIONAHA.106.672402](https://doi.org/10.1161/CIRCULATIONAHA.106.672402)] [Medline: [17309939](https://pubmed.ncbi.nlm.nih.gov/17309939/)]
45. Wang Z, Wang X. Evaluating the time-dependent predictive accuracy for event-to-time outcome with a cure fraction. *Pharm Stat*. Nov 2020;19(6):955-974. [doi: [10.1002/pst.2048](https://doi.org/10.1002/pst.2048)] [Medline: [32776646](https://pubmed.ncbi.nlm.nih.gov/32776646/)]
46. Neumann JT, Thao LTP, Murray AM, et al. Prediction of disability-free survival in healthy older people. *Geroscience*. Jun 2022;44(3):1641-1655. [doi: [10.1007/s11357-022-00547-x](https://doi.org/10.1007/s11357-022-00547-x)] [Medline: [35420334](https://pubmed.ncbi.nlm.nih.gov/35420334/)]
47. Speiser JL, Callahan KE, Ip EH, et al. Predicting future mobility limitation in older adults: a machine learning analysis of Health ABC study data. *J Gerontol A Biol Sci Med Sci*. May 5, 2022;77(5):1072-1078. [doi: [10.1093/gerona/glab269](https://doi.org/10.1093/gerona/glab269)] [Medline: [34529794](https://pubmed.ncbi.nlm.nih.gov/34529794/)]
48. Brown CJ, Flood KL. Mobility limitation in the older patient: a clinical review. *JAMA*. Sep 18, 2013;310(11):1168-1177. [doi: [10.1001/jama.2013.276566](https://doi.org/10.1001/jama.2013.276566)] [Medline: [24045741](https://pubmed.ncbi.nlm.nih.gov/24045741/)]
49. Pua YH, Tay L, Terluin B, et al. Estimating cutpoints of gait speed and sit-to-stand test values for self-reported mobility limitations in a cohort of community-dwelling older adults from Singapore: comparing receiver operating characteristic (ROC) analysis with adjusted predictive modelling. *Arch Gerontol Geriatr*. Sep 2023;112:105036. [doi: [10.1016/j.archger.2023.105036](https://doi.org/10.1016/j.archger.2023.105036)] [Medline: [37075584](https://pubmed.ncbi.nlm.nih.gov/37075584/)]
50. Li YH, Wang XH, Ya S, Jiaoling H, Hua N. The optimal cut-off value of five-time chair stand test for assessing sarcopenia among Chinese community-dwelling older adults. *J Cachexia Sarcopenia Muscle*. Apr 2024;15(2):756-764. [doi: [10.1002/jcsm.13441](https://doi.org/10.1002/jcsm.13441)] [Medline: [38342700](https://pubmed.ncbi.nlm.nih.gov/38342700/)]
51. Silva R, Pizato N, da Mata F, Figueiredo A, Ito M, Pereira MG. Mediterranean diet and musculoskeletal-functional outcomes in community-dwelling older people: a systematic review and meta-analysis. *J Nutr Health Aging*. 2018;22(6):655-663. [doi: [10.1007/s12603-017-0993-1](https://doi.org/10.1007/s12603-017-0993-1)] [Medline: [29806854](https://pubmed.ncbi.nlm.nih.gov/29806854/)]
52. Granic A, Sayer AA, Robinson SM. Dietary patterns, skeletal muscle health, and sarcopenia in older adults. *Nutrients*. Mar 30, 2019;11(4):745. [doi: [10.3390/nu11040745](https://doi.org/10.3390/nu11040745)] [Medline: [30935012](https://pubmed.ncbi.nlm.nih.gov/30935012/)]
53. Li L, Loo BP. Mobility impairment, social engagement, and life satisfaction among the older population in China: a structural equation modeling analysis. *Qual Life Res*. May 2017;26(5):1273-1282. [doi: [10.1007/s11136-016-1444-x](https://doi.org/10.1007/s11136-016-1444-x)] [Medline: [27796773](https://pubmed.ncbi.nlm.nih.gov/27796773/)]
54. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem*. Apr 1993;39(4):561-577. [Medline: [8472349](https://pubmed.ncbi.nlm.nih.gov/8472349/)]

55. Harrell FE, Lee KL, Mark DB. Prognostic/clinical prediction models: multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. In: *Tutorials in Biostatistics, Statistical Methods in Clinical Studies*. John Wiley & Sons, Ltd; 2004:223-249. [doi: [10.1002/0470023678.ch2b\(i\)](https://doi.org/10.1002/0470023678.ch2b(i))]
56. Romero-Brufau S, Huddleston JM, Escobar GJ, Liebow M. Why the C-statistic is not informative to evaluate early warning scores and what metrics to use. *Crit Care*. Aug 13, 2015;19(1):285. [doi: [10.1186/s13054-015-0999-1](https://doi.org/10.1186/s13054-015-0999-1)] [Medline: [26268570](https://pubmed.ncbi.nlm.nih.gov/26268570/)]

Abbreviations

ADL: activities of daily living
aMDS: alternate Mediterranean diet score
AUC: area under the receiver operating characteristic curve
DCA: decision curve analysis
EML: early mobility limitation
GNHS: Guangzhou Nutrition and Health Study
LASSO: least absolute shrinkage and selection operator
LightGBM: Light Gradient-Boosting Machine
LR: logistic regression
ML: machine learning
NNET: neural network
PPV: positive predictive value
RF: random forest
ROC: receiver operating characteristic
ROSE: Random Over-Sampling Examples
SHAP: Shapley Additive Explanations
STS: sit-to-stand
SVM: support vector machine
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis

Edited by Abdul-Fatawu Abdulai; peer-reviewed by Sonia Chen, Wenyan Li; submitted 12.May.2025; final revised version received 02.Feb.2026; accepted 08.Mar.2026; published 11.May.2026

Please cite as:

Freniche AC, Hu W, Chen M, Wang D, Breuillé D, Chen YM

Early Identification of Mobility Limitations in Community-Dwelling Middle-Aged and Older Adults: Development of a Prediction Model Based on a Prospective Cohort

JMIR Aging 2026;9:e77187

URL: <https://aging.jmir.org/2026/1/e77187>

doi: [10.2196/77187](https://doi.org/10.2196/77187)

© Alberto Conde Freniche, Wei Hu, Mo Chen, Dantong Wang, Denis Breuillé, Yu-ming Chen. Originally published in *JMIR Aging* (<https://aging.jmir.org>), 11.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Aging*, is properly cited. The complete bibliographic information, a link to the original publication on <https://aging.jmir.org>, as well as this copyright and license information must be included.