Brain, Behavior, and Immunity 124 (2025) 295–304

Contents lists available at ScienceDirect

Brain Behavior and Immunity

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

Full-length Article

Blood Biomarker Signatures for Slow Gait Speed in Older Adults: An Explainable Machine Learning Approach

Evrim Gökçe [1,*], Thomas Freret, Antoine Langeard

Université de Caen Normandie, INSERM, CYCERON, COMETE UMR 1075, Caen, France

ARTICLE INFO

Keywords:
Interleukin-6
C-reactive protein
Glycosylated hemoglobin
Interleukin-8
XGBoost

ABSTRACT

Maintaining physical function is crucial for independent living in older adults, with gait speed being a key predictor of health outcomes. Blood biomarkers may potentially monitor older adults' mobility, yet their association with slow gait speed still needs to be explored. This study aimed to investigate the relationship between blood biomarkers and gait speed using the Midlife in the United States (MIDUS) study biomarker dataset. A cross-sectional design was employed for analysis, involving 405 individuals aged 60 years and over. We used a machine learning framework, specifically the XGBoost algorithm, feature selection methods, and the Shapley Additive Explanations, to develop an explainable prediction model for slow gait speed. Our model demonstrated the highest cross-validation score with the six most important features among 35 variables, as elevated interleukin-6, C-reactive protein, glycosylated hemoglobin, interleukin-8, older age, and female sex were significantly associated with reduced gait speed (area under the curve = 0.75). Our findings suggest that blood biomarkers can play a critical role in integrated models to assess and monitor slow gait speed in older adults. Identifying key blood biomarkers provides valuable insights into the underlying physiological mechanisms of mobility decline and offers promising avenues for early intervention to preserve mobility in the aging population.

1. Introduction

Physical function is closely related to an individual's ability to perform daily activities. As individuals age, it is crucial to maintain adequate physical function levels to ensure independent daily living. One well-established measure of physical function is gait speed, which can predict frailty and health-related quality of life among older adults (Jung et al., 2018; Kim et al., 2016). Over the past two decades, gait speed has become a simple, safe, and inexpensive tool used in research and clinical settings to assess older adults at risk of adverse outcomes (Rydwik et al., 2012). As a responsive measure, gait speed is appropriate to predict and diagnose a variety of conditions, including orthopedic, neurodegenerative, and psychiatric disorders (Middleton et al., 2015), as well as adverse outcomes such as hospitalization and all-cause mortality (Abellan van Kan et al., 2009; Studenski et al., 2011). Moreover, the World Falls Guidelines (WFG) Task Force strongly recommends using gait speed assessment to predict fall risk in older adults (Montero-Odasso et al., 2023).

Research has shown that slow gait speed is associated with functional decline, morbidity, and mortality in older adults (Kawajiri et al., 2019). Although the cut-off points to identify individuals with slow gait speed is still under debate, the most frequently used values are 1 m/s and 0.8 m/s. Castell et al. (2013) reported that a gait speed of $\leq$ 0.8 m/s doubles the probability of frailty diagnosis. Similarly, the WFG Task Force recommends using gait speed with a cut-off value of $<$ 0.8 m/s to predict falls (Montero-Odasso et al., 2023). Further, 0.8 m/s has been recommended as an "easy-to-remember" cut-off point to predict adverse outcomes (Abellan van Kan et al., 2009). Gait speed cut-off points are important for developing a roadmap for prevention and treatment in the follow-up of older adults. For instance, according to the European consensus on sarcopenia, measuring muscle mass is recommended for anyone below a gait speed threshold of 0.8 m/s (Cruz-Jentoft et al., 2010). Therefore, assessment of gait speed using appropriate thresholds is an essential component of geriatric evaluation.

Quantifying biological molecules and their chemical kinetics in blood and tissues has been fundamental to research and medical diagnoses for decades. In particular, blood biomarkers comprehensively reflect physiological functionality and may help us understand the

* Corresponding author at: Laboratoire COMETE Mobilités: Vieillissement, Pathologie, Santé, UMRS 1075 INSERM, Université de Caen Normandie, 14032 Caen, France.
E-mail address: evrim.gokce@unicaen.fr (E. Gökçe).
[1] 0000–0003-1548–8785

https://doi.org/10.1016/j.bbi.2024.12.007
Received 21 June 2024; Received in revised form 5 December 2024; Accepted 9 December 2024
Available online 12 December 2024

molecular processes underlying aging and health deterioration (Piazza et al., 2010). Moreover, they can also serve as essential tools in translating scientific concepts into diagnostic and therapeutic technologies.

Several blood biomarkers related to the aging process have been identified, and their discovery continues to expand in aging studies (Moqri et al., 2024). A growing body of evidence highlights the potential role of blood biomarkers in relation to age-related adverse health outcomes such as multimorbidity and mortality (Fabbri et al., 2015; Tanaka et al., 2020). These biomarkers offer greater sensitivity in reflecting metabolic changes that occur with aging and may also provide valuable insights into the pathophysiology of age-related mobility loss. One of the most frequently investigated categories of blood biomarkers is inflammatory parameters in aging studies. Gait speed deterioration with aging is a process that has been reported to be related, among other factors, to increased levels of circulatory inflammatory biomarkers such as Interleukin-6 (IL-6), C-reactive protein (CRP), and Fibrinogen (Verghese et al., 2011; Ahmed-Yousef et al., 2023; Baptista et al., 2012). In contrast, other biomarkers detected in the circulation, such as antioxidants and Insulin-like Growth Factor-1 (IGF-1), have been positively associated with gait speed (Sahni et al., 2021; Córdova et al., 2015). Therefore, taking a holistic approach that includes various categories of biomarkers, such as inflammatory markers, antioxidants, metabolic markers, and lipid profiles, can provide a deeper understanding of age-related functional decline.

By examining blood biomarkers, it may be possible to identify patterns associated with an increased risk of mobility loss in older adults, thereby improving our understanding of the underlying physiological pathways and aiding in developing targeted interventions to improve health outcomes. Machine learning (ML), an innovative analytical approach, uses advanced artificial intelligence algorithms to detect patterns within data, aiming to create predictive models that consistently deliver accurate results (Bayliss & Jones, 2019). One advantage of using ML over traditional linear models is its ability to effectively capture complex, non-linear interactions between predictors, thus enhancing predictive performance. These tools have been invaluable in interpreting complex biological data and may hold the key to unlocking the complex underlying mechanisms of physical function in older adults. As mentioned above, while earlier studies have identified important relationships between specific blood biomarkers such as IL-6 and CRP and gait speed (Verghese et al., 2011, 2012), to the best of our knowledge, no study has yet investigated the relationship between gait speed and a broader set of blood biomarkers using an explainable ML methodology to capture non-linear and complex interactions.

In this study, we aimed to determine, among a large number of blood biomarkers, those associated with slow gait speed using a cohort, the MIDUS Biomarker dataset. We hypothesized that blood biomarkers would be useful to predict slow gait speed in older adults. Additionally, we aimed to enhance our understanding of the relationship between blood biomarkers and gait speed by employing distinct methodologies, including model feature importance scores and Shapley Additive Explanations (SHAP) values, and logistic regression associated with ROC curves to determine whether cut-off values of biomarkers exist in the prediction of slow gait speed. Through these approaches, we sought not only to predict gait speed but also to elucidate the impact of biomarkers on this outcome, thus providing insights into the physiological factors influencing gait speed in older individuals.

2. Methods

2.1. Design and sample

This study used the Midlife in the United States (MIDUS) dataset, a national longitudinal study conducted over 20 years. The MIDUS study began in 1995–1996 with 7,108 non-institutionalized adults selected randomly via phone dialing. Approximately 9.2 years later, 75 % of surviving participants took part in the follow-up study, MIDUS II (2004–2009). All participants underwent a medical interview, completed self-administered questionnaires, and underwent a physical examination that included a 50-foot timed walk. Additionally, as part of the refinement process for MIDUS II, a subsample of African Americans from the Milwaukee, Wisconsin region was recruited. They were invited to participate in all of the measures from MIDUS I and MIDUS II and the collection of biological samples. Thus, the Biomarker Project (Project 4) of MIDUS 2 comprised two subsamples, the longitudinal survey sample (n = 1,054) and the Milwaukee sample (n = 201), and collected data from 1,255 individuals, ranging in age from 34 to 84 years old. In this study, we analyzed a subsample of participants aged ≥ 60 (n = 405). All participants provided informed consent as approved by The University of Wisconsin-Madison Health Sciences Institutional Review Board. More details of the study are available on the MIDUS website (Available at: https://midus.wisc.edu/).

2.2. Blood biomarker assessment

Data was collected in three General Clinical Research Centers (the University of California Los Angeles, the University of Wisconsin, and Georgetown University). Eligible participants traveled and stayed overnight at one of these centers. The individuals' fasting blood samples were collected in the morning before breakfast.

Participants were requested to avoid strenuous activity before the blood draw. Venous blood samples were collected in 10 ml serum separator vacutainers from the non-dominant arm if possible. After drawing, the tubes were gently inverted 3–5 times and taken to the lab for immediate processing. A maximum of 2 h was allowed between blood draw and centrifuging. Fresh whole blood samples were refrigerated, shipped weekly to the MIDUS Biocore Lab, and assayed for glycosylated hemoglobin. Frozen serum and plasma in 1 ml aliquots were shipped to the MIDUS Biocore Lab monthly for the remaining biomarkers.

The MIDUS biomarkers dataset contains various biomarkers related to the functioning of several physiological systems, including the hypothalamic–pituitary–adrenal axis, the autonomic nervous system, the immune system, the cardiovascular system, the musculoskeletal system, antioxidants, and metabolic processes. The biomarkers assessed as follows: Total cholesterol, high-density lipoprotein (HDL), low-density lipoprotein (LDL), Triglycerides, Dehydroepiandrosterone (DHEA), Dehydroepiandrosterone-sulfate (DHEA-S), CRP, Intracellular Adhesion Molecule (ICAM), IL-6, soluble IL-6 receptor (sIL-6r), Fibrinogen, E-Selectin, IL-8, IL-10, Tumour Necrosis Factor alpha (TNF-α), Bone Specific Alkaline Phosphatase (BSAP), aminoterminal propeptide type 1 procollagen (P1NP), n-telopeptide type 1 collagen (NTx), *trans*-beta-carotene, 13-*cis*-beta-carotene, alpha-carotene, beta-cryptoxanthin, lutein, zeaxanthin, lycopene, retinol, alpha-tocopherol, gamma-tocopherol, creatinine, glycosylated hemoglobin (HbA1c), glucose, insulin, and IGF-1.

There were two assessments for IL-6, IL-8, IL-10, and TNF-α measurements, including enzyme-linked immunosorbent assay (ELISA) and Immunoelectrochemiluminescence (ECLIA) in MIDUS study. Since ECLIA provides superior assay performance than ELISA, we included only ECLIA assessments for these biomarkers in our analyses (Bolton et al., 2020). Assay details for each biomarker are available on the MIDUS Project 4 website (Available at: https://midus.wisc.edu/midus2/project4/).

2.3. Gait speed assessment

A standardized 50-foot timed walk procedure was used to measure gait speed. Participants were requested to walk at their usual speed to a 25-foot turnaround point and return to the starting point. Timing began upon the instruction to start and stopped as soon as the participant's foot crossed the starting point on their return. Each participant completed two trials, and the completion time (in seconds) for each trial was

recorded and then averaged. To calculate gait speed scores (in meters per second), we divided 15.24 m by the average trial time in seconds. A faster gait is indicated by higher scores.

2.4. Machine learning framework

ML model constructions and evaluations were performed using Python (version 3.7). The implemented Python packages were as follows: numpy, pandas, seaborn, scikit-learn, smote imbalanced-learn, yellowbrick, and OPTUNA. The ML model is available at the GitHub repository: https://github.com/evrimgokce/MIDUS.

2.4.1. Data handling and preprocessing

Data on demographics, blood biomarkers, and gait were initially compiled into a Microsoft Excel (Microsoft Inc., Redmond, WA, USA) database. As prevalent in the ML literature, we included age and sex as features in the model (Bozkurt et al., 2020). Thus, the features of slow gait speed prediction included numerical (age and blood biomarkers) and categorical (sex) data. We created two categories of gait speed, slow gait speed (< 0.8 m/s) and normal gait speed (≥ 0.8), based on the prior literature recommending this cut-off point (Abellan van Kan et al., 2009; Castell et al., 2013; Montero-Odasso et al., 2023). There were six samples with no gait speed information; these samples were dropped. In addition to this, 12 samples had various numbers of missing blood sample measurements. These missing values were imputed using mean values (Alasadi et al., 2017). After preprocessing, we had 399 samples with 35 independent variables and the target gait speed. In ML terminology, independent and dependent variables are referred to as features and target, respectively. We will utilize this terminology throughout the method section.We partitioned the original sample into a training set (80 %) and a test set (20 %) to evaluate predictive models. To ensure similar data distributions between training and test data, we utilized the adversarial validation technique that creates a combined dataset and trained a classifier to differentiate between training and test samples based on their features (Qian et al., 2022). According to this technique, if the classifier struggles to differentiate between the two datasets, it indicates that their distributions are similar, which is crucial for building reliable models. Typically, an area under curve (AUC) score of around 0.5 indicates that the model struggles to differentiate between the train and test samples. This suggests that their distributions are similar, making it challenging for the model to distinguish between them. We employed a Gaussian Naive Bayes (GaussianNB) model to verify this and proceeded to the modeling phase. Upon observing a close to 0.5 AUC score (0.493), indicating a challenge in differentiation, we proceeded with the modeling phase.

As the distribution of the target variable was quite imbalanced (slow gait speed, %18.3), we used the Synthetic Minority Over-sampling Technique (SMOTE) to address this issue. SMOTE generates synthetic samples for the minority class, thereby balancing the class distribution. By creating synthetic samples that resemble the minority class, SMOTE effectively increases the representation of the minority class in the dataset, improving the model's ability to learn from and correctly classify minority class instances (Chawla et al., 2002).

2.4.2. Training and hyperparameter tuning

We used extreme gradient boosting (XGBoost) as an ML algorithm to predict the gait speed category, which generally performs well on tabular data sets (Grinsztajn et al., 2022). It is a tree-based ensemble ML algorithm with a gradient-boosting framework. The modeling phase had three consecutive steps:

Step 1: Hyperparameter optimization with cross-validation using all features.

In step one, we trained a base model using all available features (blood biomarkers, age, and sex). We adopted a pipeline approach during the hyperparameter optimization process to avoid data leakage. Our modeling pipeline has three consecutive components: a preprocessor object responsible for missing values imputation, an oversampling object that performs SMOTE on minority class, and a predictive model, XGBoost.

Given the limited size of the dataset, we adopted a rigorous cross-validation approach to ensure reliable model evaluation. The selected method, Repeated Stratified K-fold cross-validation, was configured with parameters n_splits = 5, n_repeats = 10, and a predefined seed to ensure reproducibility (Pedregosa et al., 2011). This method divides the training data into five folds while ensuring a proportional representation of each class within each fold. During each iteration, the model is trained on four folds, and performance is evaluated on the remaining fold (referred to as hold-out data). The hold-out fold is varied iteratively to ensure a comprehensive evaluation. To address variability and ensure thoroughness, we shuffled the data and repeated the entire process ten times for each hyperparameter combination before reporting the final evaluation metric. The final performance metric is the average score across all iterations (5 x 10 iterations). We utilized the OPTUNA package for hyperparameter optimization, an open-source framework based on Bayesian optimization designed to automate the tuning process of ML models (Akiba et al., 2019). The optimization process involved evaluating 300 iterations, which corresponds to testing 300 different combinations of hyperparameters using the cross-validation strategy described previously.

Step 2: Feature selection based on the feature importance scores of the base model.

Feature importance plot shows the importance levels of features based on gain (See Supplementary Fig. 1). Gain is defined as the improvement in accuracy brought by a feature to the branches it is on. It is evident from the plot that certain features have a more significant impact on the algorithm's decision process compared to others. The data set has a relatively small sample size (399 samples) and a high number of features (35 columns). This scenario often leads to overfitting, where less informative features contribute to the model's complexity without adding meaningful predictive power or even worsening the model performance due to decreased model generalization performance. We utilized the SelectFromModel method from scikit-learn's feature selection module to mitigate this issue (Pedregosa et al., 2011). This method selects features based on their importance scores derived from the model, as depicted in the feature importance plot. During the feature selection process, we start with the most important features and iteratively add subsequent important ones. At each iteration, we perform cross-validation with the same settings in Step 1 using the base model's hyperparameters. Through this iterative process, we identify the optimal subset of features that maximize cross-validation performance.

Step 3: Repeat hyperparameter optimization with cross-validation using selected features.

In Step 3, we repeated the same methodology in Step 1 using selected features instead of the entire feature set.

2.4.3. Explainable ML approach

We used the Shapley Additive Explanations (SHAP) to further evaluate the importance of selected features in the XGBoost model, another useful tool for explaining ML models (Lundberg & Lee, 2017). Shapley's values, originating from game theory, offer valuable insights into the inner workings of complex machine-learning models. By providing Shapley-based explanations, these values shed light on the individual contributions of input features to model predictions, thus enhancing the interpretability and transparency of the model.

See Fig. 1. for a summary of the study workflow.

2.5. Statistical analysis

Continuous variables were presented as mean ± standard deviation (SD). Following ML feature selection, the receiver-operating characteristic (ROC) curve was generated for each selected blood biomarker to determine the cut-off values, and the area under the ROC curve was

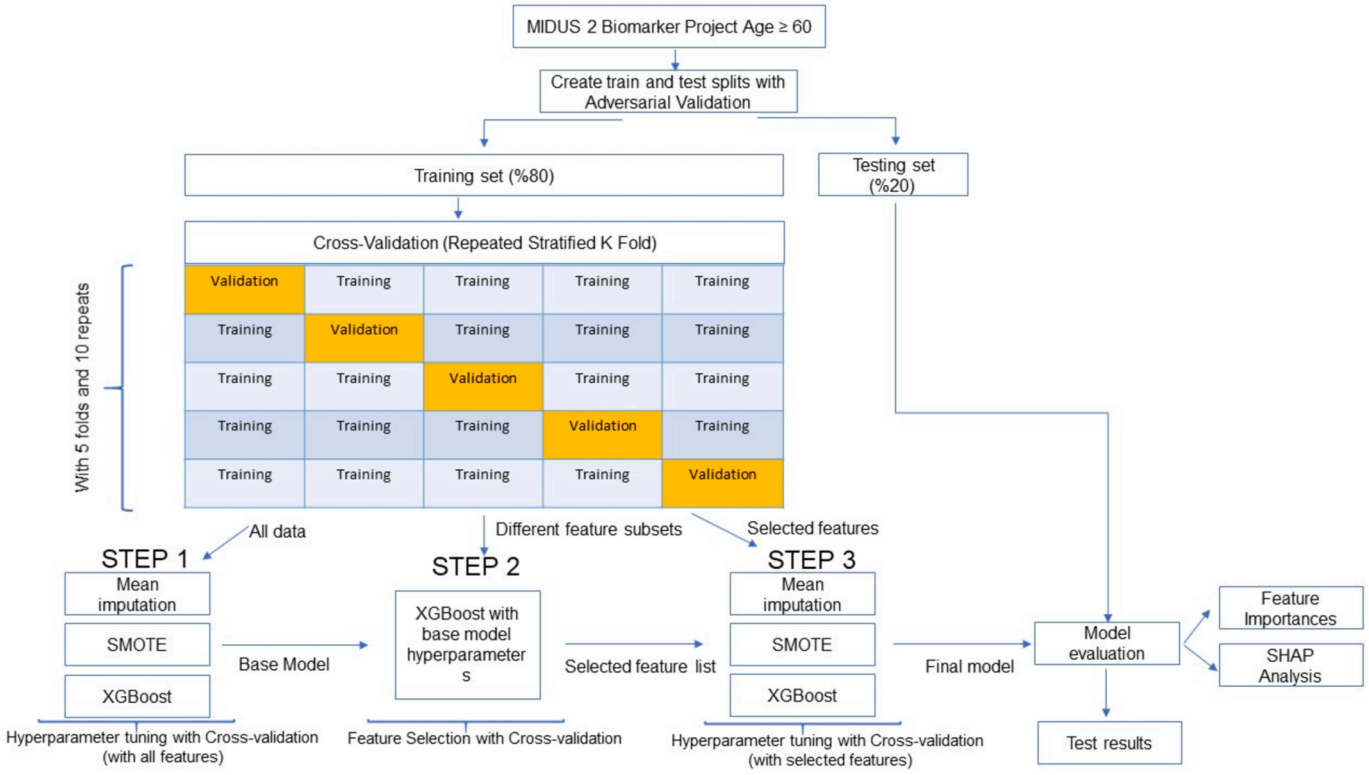

Fig. 1. Study workflow.

calculated. Optimal dichotomous cut-off values were established by maximizing Youden's index (Perkins & Schisterman, 2006). Statistical procedures were conducted using SPSS version 20.0 (IBM Corp, Armonk, NY, USA).

3. Results

A total of 405 participants were included in the study, with an average age of 68.29 (6.45). Among them, there were 217 females (53.6 %). Gait speed was above 0.8 m/s in 81.7 % and below 0.8 m/s in 18.3 % of the participants. Other baseline demographic and blood biomarker characteristics of 405 participants are summarized in Table 1. and Table 2, respectively.

Regarding the ML model, we evaluated the standard measures derived from the confusion matrix, including accuracy, sensitivity, specificity, and F1 score; however, we note that the F1 score is the recommended metric for imbalanced data sets instead of accuracy (Lones, 2021). The base model achieved a cross-validation F1 score of 0.399 and a test score 0.462 with the optimal hyperparameter combination (see Supplementary Material). During feature selection, our model demonstrated that the highest cross-validation score was achieved with the six most important features as follows: IL-6, sex, CRP, HbA1c, age, and IL-8 (Fig. 2a). Feature selection resulted in an improved cross-validation F1 score of 0.427 and a test score of 0.526 with the optimal hyperparameter combination for the base model (Table 3). In step 3, we performed the hyperparameter optimization with cross-validation with the selected six features. This refined model serves as our final predictive model with a cross-validation F1 score of 0.461 and a test score of 0.537, achieving an AUC of 0.75 (Fig. 2d). The performance summary of the final model is presented in Table 3. Additionally, we implemented the Nx Cross-Validation method to robustly assess the impact of sample size using different data percentages during cross-validation (Balki et al., 2019). Performance levels remained consistent after reaching 60 % of the data, indicating that the sample size was sufficient for the analysis.

Table 1
Participant demographics.

Characteristics	n = 405
Age (years)	68.29 (6.45)
Sex (%)	
FemaleMale	53.6 46.4
BMI (kg/m^2)	29.48 (0.32)
Gait speed (m/s)	0.98 (0.23)*
Marital status (%) (married)	65.4
Subjective well-being (M, SD)	5.05 (0.06)
(1 = strongly disagree, 7 = strongly agree)	
Data collection site (%)	
UCLA	38.3
UWGeorgetown	36.0
	25.7

Note. Continuous variables are presented as mean ± standard variation. BMI, body mass index; UCLA, University of California, Los Angeles; UW, University of Wisconsin.* n = 399.

We used the SHAP method to interpret XGBoost outputs in relation to their interactions and to reveal the importance of features. Fig. 3a demonstrates the average effect of each selected feature on the magnitude of the model output. Fig. 3b shows the most influential factors in descending order as follows: IL6, sex, CRP, HbA1c, Age, IL-8. A higher SHAP value indicates a higher slow gait speed risk. The results showed that IL-6 is the strongest factor in slow gait speed in older adults. Elevated IL-6, CRP, HbA1c, IL-8, increasing age, and female sex increased the risk of slow gait speed.

The ROC curve analysis of selected blood biomarkers is summarized in Fig. 4. The AUC of IL-6 was 0.691 ($p < 0.001$) with a 1.405 pg/mL cut-off value. The AUC of CRP was 0.612 ($p = 0.003$) with a cut-off value of 1.27 ug/mL. The AUC of HbA1c was 0.637 ($p < 0.001$) with a cut-off value of 6.221 %. The AUC of IL-8 was 0.614 ($p = 0.016$), with a cut-off value of 13.66 pg/mL. Since age was among the six selected features to predict slow gait speed, we also calculated the cut-off value for

Table 2
Blood biomarker levels of participants.

Blood biomarker	n	M (SD)	Range
Hemoglobin A1c %	400	6.22 (0.95)	4.00—12.01
Total cholesterol (mg/dL)	400	180.29 (40.49)	91—308
Triglycerides (mg/dL)	400	125.42 (70.50)	37—344
HDL cholesterol (mg/dL)	400	55.56 (17.21)	19—115
LDL cholesterol (mg/dL)	400	99.68 (35.04)	16—231
Creatinine (mg/dL)	400	0.83 (0.20)	0.5 – 2.1
DHEA-S (ug/dL)	397	105.83 (72.57)	2—495
DHEA (ng/mL)	398	6.03 (4.07)	0.4 – 24.0
Fasting glucose (mg/dL)	397	101.74 (26.51)	5—377
Fasting insulin (uIU/mL)	397	12.63 (10.19)	1—89
IGF1 (ng/mL)	397	125.45 (49.82)	29—291
IL-6 (pg/mL)	400	1.30 (1.30)	0.19 – 13.93
IL-8 (pg/mL)	400	15.93 (13.49)	4.96 – 235.14
IL-10 (pg/mL)	400	0.47 (2.26)	0.06 – 43.66
TNF alpha (pg/mL)	400	2.44 (0.95)	0.45 – 9.52
Fibrinogen (mg/dL)	399	362.25 (92.34)	123—857
C-reactive protein (ug/mL)	399	3.01 (5.32)	0.1 – 61.7
sE-selectin (ng/mL)	400	39.99 (18.86)	0.09 – 149.23
sICAM-1 (ng/mL)	400	303.89 (106.07)	4.00 – 896.47
Lutein (umol/L)	397	0.31 (0.21)	0.03 – 1.63
Zeaxanthin (umol/L)	397	0.06 (0.04)	0.01 – 0.37
Beta-cryptoxanthin (umol/L)	397	0.21 (0.19)	0.02 – 1.44
13-*cis*-beta-carotene (umol/L)	396	0.08 (0.07)	0 – 0.9
Alpha-carotene (umol/L)	397	0.09 (0.12)	0 – 1.74
Trans-beta-carotene (umol/L)	397	0.71 (0.88)	0.02 – 7.08
Lycopene (umol/L)	397	0.42 (0.23)	0.02 – 1.83
Gamma-tocopherol (umol/L)	397	3.32 (2.63)	0.44 – 20.82
Alpha-tocopherol (umol/L)	397	31.47 (13.10)	5.61 – 93.16
Retinol (umol/L)	397	1.91 (0.70)	0.71 – 6.90
n-Telopeptide type 1 collagen (nM BCE)	400	14.43 (7.02)	2.04 – 54.68
Bone-specific alkaline phosphatase (UL)	400	28.27 (12.20)	7.30 – 163.91
Amino-terminal propeptide type 1 procollagen (ugL)	400	54.68 (28.07)	7.51 – 182.79

Note. Data are presented as mean ± standard variation. IL, interleukin; sE-selectin, soluble E-selectin; sICAM-1, soluble intracellular adhesion molecule-1.

age. The AUC of age was 0.661 ($p < 0.001$), with a cut-off value of 74.5.

4. Discussion

In this study, we aimed to investigate the complex relationship between blood biomarkers and their predictive capacity for slower gait speed, a reliable proxy for assessing mobility and functional status in older adults. We have followed an explainable ML framework to predict slow gait speed based on a variety of blood biomarkers, including markers of inflammatory/metabolic/bone/renal functions, antioxidants, hormones, and lipids.

We used the feature selection method, a common data preprocessing method in ML modeling, to improve the discriminatory power (Cai et al., 2018). Our analysis initially considered 35 features, 33 of which were blood biomarkers; then, we reached the highest predictive power with six features, including IL-6, sex, CRP, HbA1c, age, and IL-8. We demonstrated that elevated IL-6, CRP, HbA1c, IL-8, and increasing age were positively associated with slow gait speed in older adults and determined a cut-off value for each of these selected blood biomarkers. Our results are compatible with previous literature reporting elevated inflammatory response as a contributing pathway to age-related decline in physical performance and mobility (Cesari et al., 2004; Penninx et al., 2004).

Our findings demonstrated that, among the extensive array of biomarkers examined, IL-6 is the strongest predictor of slower gait speed in older adults. IL-6 is a pro-inflammatory marker defined as the "cytokine of gerontologists" (Ershler, 1993). It has been reported that circulating levels of IL-6 increase with age, resulting in mobility restrictions (Wei et al., 1992; Custodero et al., 2023). In particular, elevated circulating levels of IL-6 have been associated with slower gait speed in older adults (Beavers et al., 2021; Kositsawat et al., 2020; Newman et al., 2016; Verghese et al., 2011). Although the causal relationship between the increased IL-6 and gait speed decline has not yet been fully established, one potential mechanism could involve IL-6 activating pathways that regulate muscle protein degradation and impede myogenesis, causing muscle atrophy (Belizário et al., 2016). Indeed, it has been suggested that elevated levels of IL-6 may contribute to a decline in mobility through a parallel decrease in knee extensor muscle strength (Ferrucci et al., 2002; Custodero et al., 2020). Confirmingly, a randomized controlled study of older adults reported that an age-related increase in circulating IL-6 is an important contributor to declines in skeletal muscle strength, quality, and function (Grosicki et al., 2020).

Our study is the first to report a cut-off value of 1.405 pg/mL for IL-6 as predictive of slow gait speed. However, prolonged exposure to high IL-6 levels, rather than isolated elevations, might be more strongly associated with slower gait speed. For instance, Nadkarni et al. (2016) have suggested that an average IL-6 level of 2.7 pg/ml better predicted worsening gait speed over a 10-year follow-up. Therefore, future longitudinal research should verify the IL-6 cut-off value we identified in this study.

According to our ML model, CRP, an acute-phase protein considered a general inflammatory biomarker, was the second most important blood biomarker to predict slow gait speed. In line with our results, Penninx et al. (2004) reported that elevated CRP levels have predicted mobility limitation in older adults. Likewise, CRP levels have been reported to be inversely associated with gait speed (Beavers et al.,2021; Sousa et al., 2016; Taaffe et al., 2000). Additionally, Beavers et al. (2021) reported that IL-6 appeared to be more strongly associated with slower gait speed than CRP when these two biomarkers were assessed in combination. Our findings align with this observation, yet it is important to emphasize that our ML model considered a far greater quantity of biomarkers, highlighting IL-6 as the most influential predictor. Kositsawat et al. (2020), however, reported that once other biomarkers involving IL-6, IGF-1, and vitamin D were taken into consideration, the previously established link between higher levels of CRP and slow gait speed (< 0.8 m/s) became insignificant. They suggested that the observed phenomenon can be attributed to the overlapping effects of IL-6, resulting in offsetting previous effects. However, although the precise mechanisms involved are not yet fully comprehended, it is worth noting that the machine learning model assessing numerous biomarkers interacting with each other is likely to offer a more dependable explanation of how CRP affects mobility.

Our ROC analysis provided a cut-off value of 1.27 ug/ml for CRP to predict slow gait speed. On the other hand, Verghese et al. (2012) reported that older adults with elevated high-sensitivity CRP levels (≥ 3 ug/ml) had 0.89 cm/s per year faster decline in gait speed. The discrepancies in cut-off values could be attributed to different study designs and techniques employed to assess gait speed and CRP levels. Larger sample sizes and longitudinal data sets are required to report more accurate cut-off values for the relationship between CRP levels and gait speed.

Our findings demonstrated that HbA1c, a measure to estimate the mean blood glucose levels in the previous three months, is one of the strongest predictors of slower gait speed in older adults. This finding is consistent with a longitudinal study on diabetic patients showing that gait speed has increased in those who have decreased HbA1c value by 1 % or more over a year (Sugimoto et al., 2021). Similarly, Azmon et al. (2018) reported a significant negative correlation between HbA1c and gait speed among older adults with type 2 diabetes. Considering that poor glycemic control in older adults with diabetes has been associated with lower muscle strength and muscle mass (Ogama et al., 2019; Yoon et al., 2016), one potential explanation for the contribution of HbA1c on

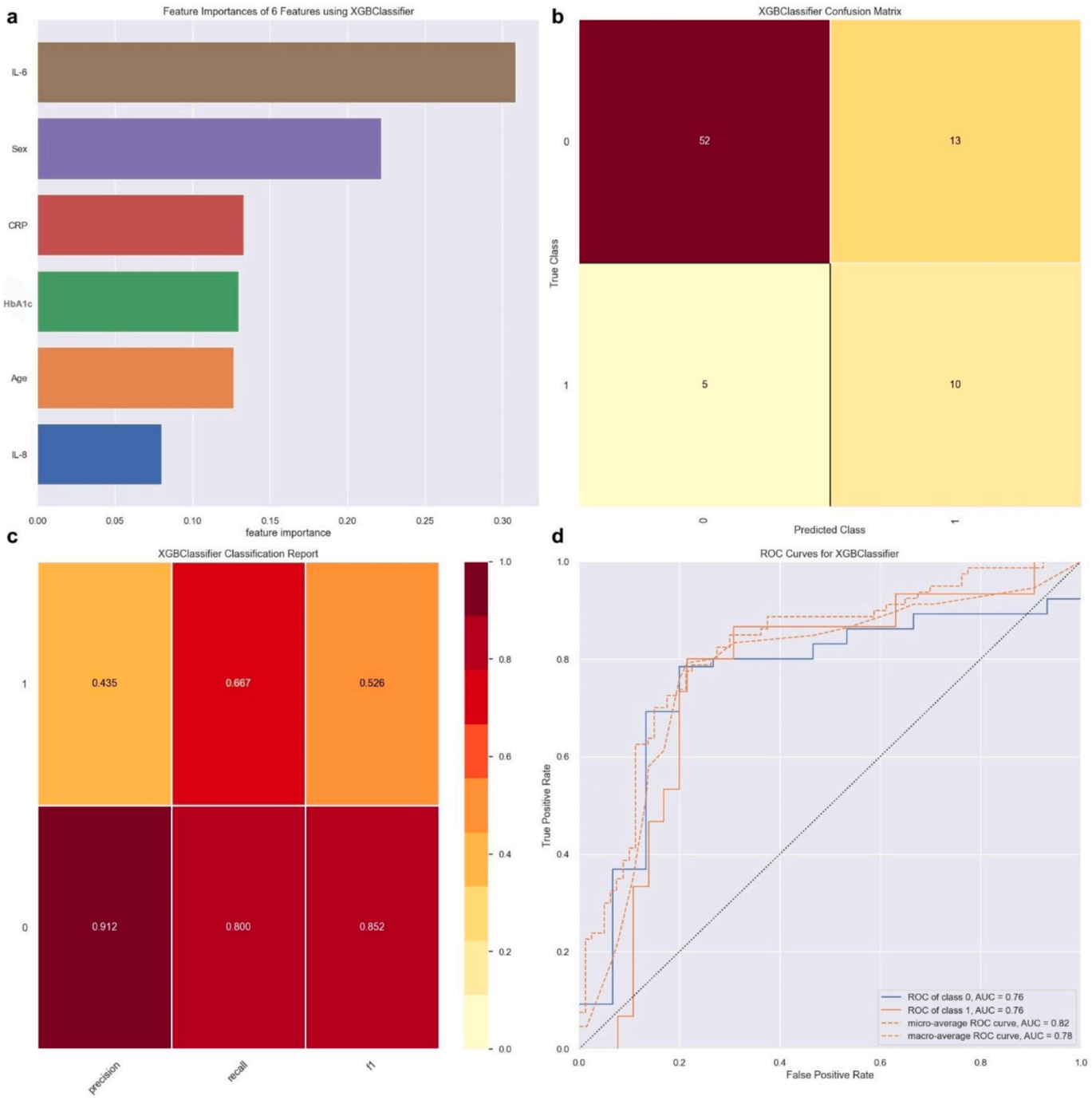

Fig. 2. **(a)** Feature importance in the XGBoost model. **(b)** Confusion matrix, presenting the summary of the actual class and predicated class. The diagonals indicate the correctly predicted classes (true positive or true negative), while the off-diagonal cells depict the incorrectly predicted classes (false positive or false negative). **(c)** Precision, sensitivity (recall), and F1 score of the model. **(d)** Receiver operating characteristic curves for the test set. AUC, area under curve; CRP, C-reactive protein; HbA1c, glycosylated hemoglobin, IL-6, interleukin 6; IL-8, interleukin 8; ROC, receiver operating characteristic; XGBoost, eXtreme Gradient Boosting.

Table 3
Performance metrics of the prediction model.

	Precision (%)	Recall (%)	Specificity (%)	AUC (95 % CI)	F1 Score
Cross-Validation	35.1	68.3	71.4	0.739	0.461
Test	42.3	73.3	76.9	0.754	0.537

Note. AUC, area under curve; CI, confidence interval.

slow gait speed may be its link to muscle metabolism. Additionally, HbA1c has been associated with inflammation (Wu et al., 2002; Gustavsson & Agardh, 2004). Relatedly, higher HbA1c levels have been associated with increased high-sensitivity CRP, suggesting the association of HbA1c with increased systemic inflammation (Ahmad et al., 2021). This appears to be consistent with our findings, indicating that inflammatory markers are the most important features in slow gait speed prediction.

Our ROC analysis provided a cut-off value of 6.22 % for HbA1c in predicting slow gait speed. Since this level is classified as pre-diabetes by the American Diabetes Association (2010), it is reasonable to suggest

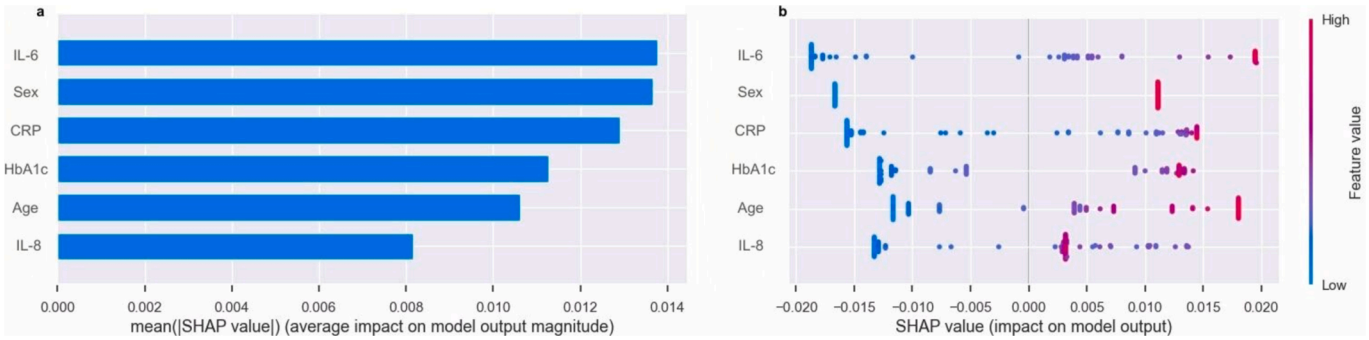

Fig. 3. SHAP plots. **(a)** Mean SHAP values **(b)** Relative effect of each feature, providing a comprehensive understanding of their respective influences. Feature importance is presented in increasing order (vertical axis), with each point representing a sample. The dot color gets redder when the feature value gets higher. For the binary feature vector (sex), red dots indicated the female sex. The relative impact of the feature on the model output is presented on the *x*-axis (the right of 0.0, increased slow gait speed risk; the left of 0.0, reduced slow gait speed risk). IL-6, interleukin 6; CRP, C-reactive protein; HbA1c, glycosylated hemoglobin; IL-8, interleukin 8; SHAP, Shapley Additive exPlanations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

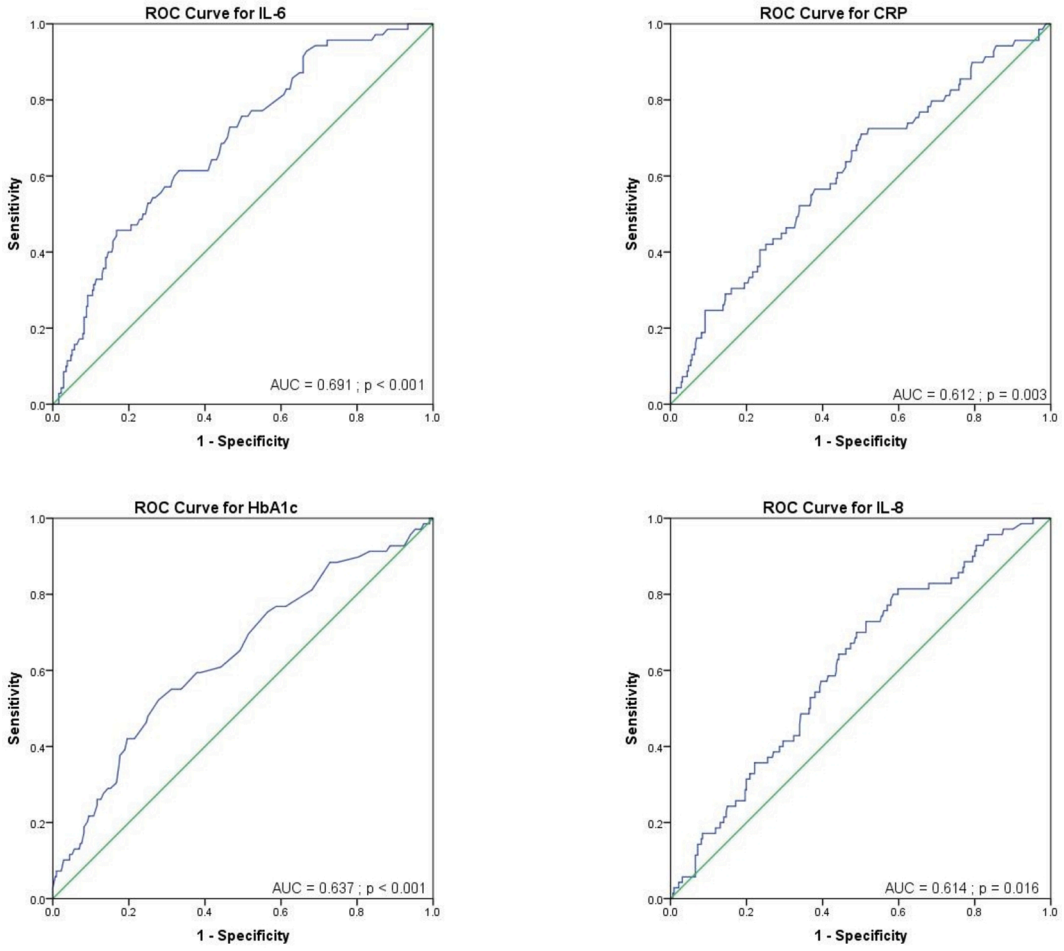

Fig. 4. The ROC analysis of selected biomarkers by ML model. AUC, area under curve; CRP, C-reactive protein; HbA1c, glycosylated hemoglobin, IL-6, interleukin 6; IL-8, interleukin 8; ROC, receiver operating characteristic.

that pre-diabetic older adults are at risk of mobility decline. Relatedly, a study including two cohorts of patients aged 50 years and older with type 2 diabetes has reported that HbA1c of 6.4 % was associated with a heightened risk of all-cause mortality (Currie et al., 2010). Although our study is the first report indicating an HbA1c cut-off value to predict slow gait speed, it is worth mentioning that our finding is quite close to the median value reported for all-cause mortality in patients with type 2 diabetes. Based on these observations, it is fair to suggest monitoring glycemic control status to prevent both mobility loss and all-cause mortality in older adults.

IL-8 is another inflammatory marker that stands out in our ML model. A longitudinal study of older adults found that higher levels of IL-

8 are associated with decreased lower appendicular lean mass and a higher risk of sarcopenia (Westbury et al., 2018). Also, it has been reported that IL-8 was inversely correlated with muscle strength (Dupont et al., 2023). These data suggest that IL-8, along with other inflammatory blood biomarkers, may negatively impact muscle mass and muscle strength, leading to slow gait speed. Our ROC analysis provided a cut-off value of 13.66 pg/mL for IL-8 to predict slow gait speed, which is higher than the mean values of 3.35 pg/ml for serum IL-8 and 3.24 pg/ml for plasma IL-8 reported in studies conducted on healthy individuals (González et al., 2001; Straczkowski et al., 2002). However, it should be noted that the participants in these studies were not older adults. Future research involving healthy older individuals could provide more comparable cut-off values to interpret our findings.

As we mentioned above, the effect of inflammation on skeletal muscle tissue may be one possible common mechanism to explain the inverse association between inflammatory markers and gait speed. Indeed, it has been reported that chronic inflammation promotes skeletal muscle protein breakdown (Wilson et al., 2017), which in turn may cause muscle weakness (Custodero et al., 2020). Similarly, inflammatory cytokines induce excessive generation of free-radical species in skeletal muscle, resulting in reduced muscle force generation (Supinski & Callahan, 2007). Thus, decreased muscle function may lead to slower gait speed in older adults.

Another potential explanation for the association between systemic inflammation and gait speed may be cognition. A *meta*-analysis of seven studies reported an association between high circulating levels of IL-6 and global cognitive decline in non-demented adults (Bradburn et al., 2017). Similarly, increased peripheral levels of IL-8 have been linked with poorer cognitive performance (Baune et al., 2008). Despite some inconsistencies in the literature (Gabin et al., 2018), a *meta*-analysis of four studies provides evidence of a weak but significant association between peripheral CRP level and global cognitive decline. Regarding HbA1c, although most of the previous studies considered samples of only diabetics, it has been reported that high levels of HbA1c are associated with poor cognition (Mimenza-Alvarado et al., 2020; Silverman et al., 2019). Since walking is a complex activity involving both motor skills and cognitive functions, increased levels of biomarkers related to inflammation are likely to negatively impact gait speed through deteriorated cognition.

Besides inflammatory markers, age and sex were within six of the most important features in the XGBoost algorithm, indicating that increasing age and being female contribute to slow gait speed, as previously reported (Andrews et al., 2023; Rössler et al., 2024). Regarding age, we established a cut-off value of 74.5 years of age to predict slow gait speed. Similarly, Castell et al. (2013) reported that a walking speed of less than 0.8 m/s was presented in 56.4 % of individuals $\geq$ 75 years of age. Relatedly, it has been reported that gait speed was especially important after age 75, predicting frailty and increased demand for more health resources (Studenski et al., 2011; Gómez Pavón et al., 2007). Regarding sex, SHAP analysis revealed that the female sex is associated with slower gait speed in older adults, consistent with existing literature reporting that older women typically walk more slowly than men (Wheaton & Crimmins, 2016). One possible explanation for this difference is the height difference between men and women, as taller individuals generally have faster gait speeds, with longer legs contributing to this advantage (Bohannon, 1997). On the other hand, prior research has demonstrated that the gait speed difference between women and men persists even after height adjustment (Sialino et al., 2019; Kasović et al., 2021). This suggests that other factors, including variations in muscle strength, exposure to sex hormones, lifestyle differences, and a greater burden of chronic diseases among women, may contribute to this observed sex-based difference (Tseng et al., 2014). Future research focusing on the effects of blood biomarkers on mobility should consider age subgroups and sex-specific pathways to better understand the complex mechanisms contributing to mobility loss.

Our ML model pointed out that inflammatory markers such as IL-6 and CRP are more important features contributing to slow gait speed than chronological age. Previous studies have reported that biological age, also known as physiological or functional age, depends on factors such as genetics, lifestyle, and medical history besides chronological age (Jazwinski et al., 2019). As a result, biological age can be a more precise indicator of aging than chronological age, offering a quantitative standard for tracking personalized aging trajectory. Our findings suggest that blood biomarkers, particularly inflammatory markers, may contribute to biological age and are promising indicators of monitoring physical functionality as people age.

This study offers several strengths. Firstly, we achieved significant predictive power by focusing on an extensive number of blood biomarkers that can be easily assessed in routine laboratory analyses, as well as simple sample characteristics such as age and sex. Secondly, we followed state-of-the-art practices for model development, including a rigorous cross-validation strategy, feature selection, dealing with imbalanced class distribution, hyperparameter tuning, and metric evaluation. Thirdly, we used two explainability methods (model feature importance scores and SHAP) to explain the relationship between the input features and the output of the XGBoost model, thereby helping to better understand the complex relationship between the blood biomarkers assessed and slow gait speed. Finally, we established the cut-off values of blood biomarker levels for the first time to predict slow gait speed.

A few limitations of this study are worth noting. First, our findings are based on a relatively small sample size and need to be confirmed through further research integrating multidimensional data. Second, since our model solely focuses on blood biomarkers, a more holistic approach considering multiple factors (e.g., body composition, cardiorespiratory fitness, comorbidities, age-associated diseases, medications, etc.) should corroborate our findings. Third, due to the cross-sectional design of this study, we cannot establish a causal relationship. Finally, it should be kept in mind that the model's performance may change with different datasets when applying ML algorithms; findings need to be corroborated in a broader range of population characteristics to generalize and verify our results.

5. Conclusions

To the best of our knowledge, this study is the first to address the complex relationships between numerous blood biomarkers and slow gait speed in older adults by applying an ML algorithm. Our findings indicate that low-grade inflammatory status plays a critical role in the underlying physiological mechanisms associated with mobility decline in older adults. Further research utilizing longitudinal mobility data is required to validate these findings and to explore how these biomarkers contribute to the progression of mobility impairment over time. These insights could form the basis for a deeper understanding of the biological signature of age-related mobility decline, helping to advance research in age-related functional impairments.

CRediT authorship contribution statement

Evrim Gökçe: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Thomas Freret:** Writing – review & editing, Methodology, Conceptualization. **Antoine Langeard:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbi.2024.12.007.

Data availability

Dataset is available at: https://github.com/evrimgokce/MIDUS.

References

Abellan van Kan, G., Rolland, Y., Andrieu, S., Bauer, J., Beauchet, O., Bonnefoy, M., Cesari, M., Donini, L.M., Gillette Guyonnet, S., Inzitari, M., Nourhashemi, F., Onder, G., Ritz, P., Salva, A., Visser, M., Vellas, B., 2009. Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) Task Force. *J. Nutr. Health Aging* 13 (10), 881–889. https://doi.org/10.1007/s12603-009-0246-z.

Ahmad, K., Verghese, D., Kinninger, A., Dahal, S., Ghanem, A., Rezvanizadeh, V., Manubolu, V.S., Javier, D.A., Vicuna Jr, W., Shafter, A.M., Alalawi, L., Aldana Bitar, J., 2021. Correlation Between Hemoglobin A1c and High Sensitivity C-Reactive Protein in Population with Type II Diabetes Mellitus in the Semaglutide Treatment on Coronary Plaque Progression (STOP) Trial. Am. Heart J. 242, 157. https://doi.org/10.1016/j.ahj.2021.10.031.

Ahmed-Yousef, N.S., Dilian, O., Iktilat, K., Agmon, M., 2023. CRP, but not fibrinogen, is associated with gait speed as early as middle age, in females but not males. Sci Rep 13 (1), 15571. https://doi.org/10.1038/s41598-023-42183-1.

Akiba, T., Sano, S., Yanase, T., Ohta, T., 2019. August). Optuna: A next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. https://doi.org/10.1145/3292500.3330701.

Alasadi, S.A., Bhaya, W.S., 2017. Review of Data Preprocessing Techniques in Data Mining. J Eng Appl Sci 12, 4102–4107.

Andrews, A.W., Vallabhajosula, S., Boise, S., Bohannon, R.W., 2023. Normal gait speed varies by age and sex but not by geographical region: a systematic review. J Physiother 69 (1), 47–52. https://doi.org/10.1016/j.jphys.2022.11.005.

American Diabetes Association (2010). Diagnosis and classification of diabetes mellitus. *Diabetes care*, *33 Suppl 1*(Suppl 1), S62–S69. https://doi.org/10.2337/dc10-S062.

Azmon, M., Gayus, N., Michal, H., Olmer, L., Cukierman-Yaffe, T., 2018. The Association between Glucose Control and Functional Indices in Older People with Diabetes. *Int. J. Endocrinol.* 2018, 1053815. https://doi.org/10.1155/2018/1053815.

Balki, I., Amirabadi, A., Levman, J., Martel, A.L., Althobaity, M., Farag, A., Gazarian, A., 2019. Sample-size determination methodologies for machine learning in medical imaging research: A systematic review. Can. Assoc. Radiol. J. 70 (4), 344–353. https://doi.org/10.1016/j.carj.2019.06.002.

Baptista, G., Dupuy, A.M., Jaussent, A., Durant, R., Ventura, E., Sauguet, P., Picot, M.C., Jeandel, C., Cristol, J.P., 2012. Low-grade chronic inflammation and superoxide anion production by NADPH oxidase are the main determinants of physical frailty in older adults. *Free Radic. Res.* 46 (9), 1108–1114. https://doi.org/10.3109/10715762.2012.692784.

Baune, B.T., Ponath, G., Golledge, J., Varga, G., Arolt, V., Rothermundt, M., Berger, K., 2008. *Neurobiol. Aging* 29 (6), 937–944. https://doi.org/10.1016/j.neurobiolaging.2006.12.003.

Bayliss, L., Jones, L.D., 2019. The role of artificial intelligence and machine learning in predicting orthopaedic outcomes. Bone & Joint Journal 101-B(12), 1476–1478. https://doi.org/10.1302/0301-620X.101B12.BJJ-2019-0850.R1.

Beavers, D.P., Kritchevsky, S.B., Gill, T.M., Ambrosius, W.T., Anton, S.D., Fielding, R.A., King, A.C., Rejeski, W.J., Lovato, L., McDermott, M.M., Newman, A.B., Pahor, M., Walkup, M.P., Tracy, R.P., Manini, T.M., 2021. Elevated IL-6 and CRP Levels Are Associated With Incident Self-Reported Major Mobility Disability: A Pooled Analysis of Older Adults With Slow Gait Speed. *J. Gerontol. A Biol. Sci. Med. Sci.* 76 (12), 2293–2299. https://doi.org/10.1093/gerona/glab093.

Belizário, J.E., Fontes-Oliveira, C.C., Borges, J.P., Vannier, E., 2016. Skeletal muscle wasting and renewal: A pivotal role of myokine IL-6. Springerplus 5, 619. https://doi.org/10.1186/s40064-016-2197-2.

Bohannon, R.W., 1997. Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants. *Age Ageing* 26 (1), 15–19. https://doi.org/10.1093/ageing/26.1.15.

Bolton, J.S., Chaudhury, S., Dutta, S., Gregory, S., Locke, E., Pierson, T., Bergmann-Leitner, E.S., 2020. Comparison of ELISA with electro-chemiluminescence technology for the qualitative and quantitative assessment of serological responses to vaccination. *Malar. J.* 19 (1), 159. https://doi.org/10.1186/s12936-020-03225-5.

Bozkurt, S., Cahan, E.M., Seneviratne, M.G., Sun, R., Lossio-Ventura, J.A., Ioannidis, J.P. A., Hernandez-Boussard, T., 2020. Reporting of demographic data and representativeness in machine learning models using electronic health records. *Journal of the American Medical Informatics Association : JAMIA* 27 (12), 1878–1884. https://doi.org/10.1093/jamia/ocaa164.

Bradburn, S., Sarginson, J., Murgatroyd, C.A., 2018. Association of peripheral interleukin-6 with global cognitive decline in non-demented adults: A meta-analysis of prospective studies. Front. Aging Neurosci. 9, 438. https://doi.org/10.3389/fnagi.2017.00438.

Cai, J., Luo, J., Wang, S., Yang, S., 2018. Feature selection in machine learning: A new perspective. Neurocomputing 300, 70–79. https://doi.org/10.1016/j.neucom.2017.11.077.

Castell, M.V., Sánchez, M., Julián, R., Queipo, R., Martín, S., Otero, Á., 2013. Frailty prevalence and slow walking speed in persons age 65 and older: Implications for primary care. BMC Fam. Pract. 14, 86. https://doi.org/10.1186/1471-2296-14-86.

Cesari, M., Penninx, B.W., Pahor, M., Lauretani, F., Corsi, A.M., Rhys Williams, G., Guralnik, J.M., Ferrucci, L., 2004. Inflammatory markers and physical performance in older persons: the InCHIANTI study. *J. Gerontol. A Biol. Sci. Med. Sci.* 59 (3), 242–248. https://doi.org/10.1093/gerona/59.3.m242.

Chawla, N.W., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: Synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357. https://doi.org/10.1613/jair.953.

Córdova, C., Boullosa, D. A., Custódio, M. R., Quaglia, L. A., Santos, S. N., Freitas, W. M., Sposito, A. C., Nóbrega, O. T., & Brazilian Study on Healthy Aging (2016). Atheroprotective Properties of Serum IGF-1 in the Carotid and Coronary Territories and Beneficial Role on the Physical Fitness of the Oldest Old. *The journals of gerontology. Series A, Biological sciences and medical sciences*, *71*(10), 1281–1288. https://doi.org/10.1093/gerona/glv216.

Cruz-Jentoft, A. J., Baeyens, J. P., Bauer, J. M., Boirie, Y., Cederholm, T., Landi, F., Martin, F. C., Michel, J. P., Rolland, Y., Schneider, S. M., Topinková, E., Vandewoude, M., Zamboni, M., & European Working Group on Sarcopenia in Older People, 2010. Sarcopenia: European consensus on definition and diagnosis: Report of the European Working Group on Sarcopenia in Older People. *Age Ageing* 39 (4), 412–423. https://doi.org/10.1093/ageing/afq034.

Currie, C.J., Peters, J.R., Tynan, A., Evans, M., Heine, R.J., Bracco, O.L., Zagar, T., Poole, C.D., 2010. Survival as a function of HbA(1c) in people with type 2 diabetes: a retrospective cohort study. *Lancet (london, England)* 375 (9713), 481–489. https://doi.org/10.1016/S0140-6736(09)61969-3.

Custodero, C., Anton, S.D., Beavers, D.P., Mankowski, R.T., Lee, S.A., McDermott, M.M., Fielding, R.A., Newman, A.B., Tracy, R.P., Kritchevsky, S.B., Ambrosius, W.T., Pahor, M., Manini, T.M., 2020. ENRGISE study investigators: The relationship between interleukin-6 levels and physical performance in mobility-limited older adults with chronic low-grade inflammation: The ENRGISE Pilot study. *Arch. Gerontol. Geriatr.* 90, 104131. https://doi.org/10.1016/j.archger.2020.104131.

Custodero, C., Pahor, M., Mazzoccoli, C., Manini, T.M., Anton, S.D., Mazzocca, A., Lozupone, M., Panza, F., Sabbà, C., Solfrizzi, V., 2023. Effect of change of interleukin-6 over time on gait speed response: Results from the lifestyle interventions and independence for elders study. *Mech. Ageing Dev.* 210, 111763. https://doi.org/10.1016/j.mad.2022.111763.

Dupont, J., Vercauteren, L., Amini, N., Lapauw, L., De Schaepdryver, M., Poesen, K., Dedeyne, L., Verschueren, S., Tournoy, J., Koppo, K., Gielen, E., 2023. Are inflammatory markers associated with sarcopenia-related traits in older adults with sarcopenia? - A cross-sectional analysis of the ENHANce study. *Exp. Gerontol.* 178, 112196. https://doi.org/10.1016/j.exger.2023.112196.

Ershler, W.B., 1993. Interleukin-6: a cytokine for gerontologists. J Am Geriatr Soc 41 (2), 176–181. https://doi.org/10.1111/j.1532-5415.1993.tb02054.x.

Fabbri, E., An, Y., Zoli, M., Simonsick, E.M., Guralnik, J.M., Bandinelli, S., Boyd, C.M., Ferrucci, L., 2015. Aging and the burden of multimorbidity: associations with inflammatory and anabolic hormonal biomarkers. *The Journals of Gerontology. Series a, Biological Sciences and Medical Sciences* 70 (1), 63–70. https://doi.org/10.1093/gerona/glu127.

Ferrucci, L., Penninx, B.W., Volpato, S., Harris, T.B., Bandeen-Roche, K., Balfour, J., Leveille, S.G., Fried, L.P., Md, J.M., 2002. Change in muscle strength explains accelerated decline of physical function in older women with high interleukin-6 serum levels. *Journal of the American Geriatrics Society* 50 (12), 1947–1954. https://doi.org/10.1046/j.1532-5415.2002.50605.x.

Gabin, J.M., Saltvedt, I., Tambs, K., Holmen, J., 2018. The association of high sensitivity C-reactive protein and incident Alzheimer disease in patients 60 years and older: The HUNT study, Norway. *Immunity & Ageing : I & A* 15, 4. https://doi.org/10.1186/s12979-017-0106-3.

Gómez Pavón, J., Martín Lesende, I., Baztán Cortés, J.J., Regato Pajares, P., Formiga Pérez, F., Segura Benedito, A., Abizanda Soler, P., de Pedro Cuesta, J., 2008. (Grupo Impulsor del documento), grupos de trabajo de expertos de las sociedades científicas, & Administraciones sanitarias: Prevención de la dependencia en las personas mayores [Preventing dependency in the elderly]. *Rev. Clin. Esp.* 208 (7), 361–362. https://doi.org/10.1157/13124318.

González, C., Cava, F., Ayllón, A., Guevara, P., Navajo, J.A., González-Buitrago, J.M., 2001. Biological variation of interleukin-1beta, interleukin-8 and tumor necrosis factor-alpha in serum of healthy individuals. *Clin. Chem. Lab. Med.* 39 (9), 836–841. https://doi.org/10.1515/CCLM.2001.139.

Grinsztajn, L., Edouard, O., Gaël, V. (2022) Why Do Tree-based Models Still Outperform Deep Learning on Tabular Data? arXiv.

Grosicki, G.J., Barrett, B.B., Englund, D.A., Liu, C., Travison, T.G., Cederholm, T., Koochek, A., von Berens, Å., Gustafsson, T., Benard, T., Reid, K.F., Fielding, R.A., 2020. Circulating Interleukin-6 Is Associated with Skeletal Muscle Strength, Quality, and Functional Adaptation with Exercise Training in Mobility-Limited Older Adults. *J. Frailty Aging* 9 (1), 57–63. https://doi.org/10.14283/jfa.2019.30.

Gustavsson, C.G., Agardh, C.D., 2004. Markers of inflammation in patients with coronary artery disease are also associated with glycosylated haemoglobin A1c within the normal range. *Eur. Heart J.* 25 (23), 2120–2124. https://doi.org/10.1016/j.ehj.2004.09.008.

Jazwinski, S.M., Kim, S., 2019. Examination of the Dimensions of Biological Age. *Front. Genet.* 10, 263. https://doi.org/10.3389/fgene.2019.00263.

Jung, H.W., Jang, I.Y., Lee, C.K., Yu, S.S., Hwang, J.K., Jeon, C., Lee, Y.S., Lee, E., 2018. Usual gait speed is associated with frailty status, institutionalization, and mortality in community-dwelling rural older adults: a longitudinal analysis of the Aging Study of Pyeongchang Rural Area. *Clin. Interv. Aging* 13, 1079–1089. https://doi.org/10.2147/CIA.S166863.

Kasović, M., Štefan, L., Štefan, A., 2021. Normative Data for Gait Speed and Height Norm Speed in ≥ 60-Year-Old Men and Women. *Clin. Interv. Aging* 16, 225–230. https://doi.org/10.2147/CIA.S290071.

Kawajiri, H., Mishina, H., Asano, S., Kono, Y., Hayashi, H., Niwa, J.I., Doyu, M., Kimura, S., Yamada, S., 2019. Maximum Walking Speed at Discharge Could Be a

Prognostic Factor for Vascular Events in Patients With Mild Stroke: A Cohort Study. *Arch. Phys. Med. Rehabil.* 100 (2), 230–238. https://doi.org/10.1016/j.apmr.2018.05.025.

Kim, H.J., Park, I., Lee, H.J., Lee, O., 2016. The reliability and validity of gait speed with different walking pace and distances against general health, physical function, and chronic disease in aged adults. *J. Exerc. Nutr. Biochem.* 20 (3), 46–50. https://doi.org/10.20463/jenb.2016.09.20.3.7.

Kositsawat, J., Kuo, C.L., Barry, L.C., Melzer, D., Bandinelli, S., Ferrucci, L., Wu, R., Kuchel, G.A., 2020. Interaction Between Vitamin D and Interleukin 6 on Slow Gait Speed: 6-Year Follow-up Data of Older Adults From InCHIANTI. *J. Gerontol. A Biol. Sci. Med. Sci.* 75 (6), 1161–1166. https://doi.org/10.1093/gerona/glz165.

Lones, M. (2024). How to avoid machine learning pitfalls: a guide for academic researchers. *arXiv*. https://doi.org/10.48550/arXiv.2108.02497make it APA reference.

Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems. *Arxiv*.

Middleton, A., Fritz, S.L., Lusardi, M., 2015. Walking speed: the functional vital sign. *J. Aging Phys. Act.* 23 (2), 314–322. https://doi.org/10.1123/japa.2013-0236.

Mimenza-Alvarado, A.J., Jiménez-Castillo, G.A., Yeverino-Castro, S.G., Barragán-Berlanga, A.J., Pérez-Zepeda, M.U., Ávila-Funes, J.A., Aguilar-Navarro, S.G., 2020. Effect of poor glycemic control in cognitive performance in the elderly with type 2 diabetes mellitus: The Mexican Health and Aging Study. *BMC Geriatr.* 20 (1), 424. https://doi.org/10.1186/s12877-020-01827-x.

Montero-Odasso, M., van der Velde, N., Martin, F. C., Petrovic, M., Tan, M. P., Ryg, J., Aguilar-Navarro, S., Alexander, N. B., Becker, C., Blain, H., Bourke, R., Cameron, I. D., Camicioli, R., Clemson, L., Close, J., Delbaere, K., Duan, L., Duque, G., Dyer, S. M., Freiberger, E., … Task Force on Global Guidelines for Falls in Older Adults (2022). World guidelines for falls prevention and management for older adults: a global initiative. *Age and ageing*, *51*(9), afac205. https://doi.org/10.1093/ageing/afac205.

Moqri, M., Herzog, C., Poganik, J.R., Ying, K., Justice, J.N., Belsky, D.W., Higgins-Chen, A.T., Chen, B.H., Cohen, A.A., Fuellen, G., Hägg, S., Marioni, R.E., Widschwendter, M., Fortney, K., Fedichev, P.O., Zhavoronkov, A., Barzilai, N., Lasky-Su, J., Kiel, D.P., Kennedy, B.K., Ferrucci, L., 2024. Validation of biomarkers of aging. *NatureMedicine* 30 (2), 360–372. https://doi.org/10.1038/s41591-023-02784-9.

Nadkarni, N.K., Boudreau, R.M., Studenski, S.A., Lopez, O.L., Liu, G., Kritchevsky, S., Yaffe, K., Newman, A.B., Rosano, C., 2016. Slow gait, white matter characteristics, and prior 10-year interleukin-6 levels in older adults. *Neurology* 87 (19), 1993–1999. https://doi.org/10.1212/WNL.0000000000003304.

Newman, A.B., Sanders, J.L., Kizer, J.R., Boudreau, R.M., Odden, M.C., Hazzouri A., Z.A., Arnold, A.M., 2016. Trajectories of function and biomarkers with age: the CHS All Stars Study. *Int. J. Epidemiol.* 45 (4), 1135–1145. https://doi.org/10.1093/ije/dyw092.

Ogama, N., Sakurai, T., Kawashima, S., Tanikawa, T., Tokuda, H., Satake, S., Miura, H., Shimizu, A., Kokubo, M., Niida, S., Toba, K., Umegaki, H., Kuzuya, M., 2019. Association of Glucose Fluctuations with Sarcopenia in Older Adults with Type 2 Diabetes Mellitus. *J. Clin. Med.* 8 (3), 319. https://doi.org/10.3390/jcm8030319.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, É., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12 (85), 2825–2830.

Penninx, B.W., Kritchevsky, S.B., Newman, A.B., Nicklas, B.J., Simonsick, E.M., Rubin, S., Nevitt, M., Visser, M., Harris, T., Pahor, M., 2004. Inflammatory markers and incident mobility limitation in the elderly. *Journal of the American Geriatrics Society* 52 (7), 1105–1113. https://doi.org/10.1111/j.1532-5415.2004.52308.x.

Perkins, N.J., Schisterman, E.F., 2006. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am. J. Epidemiol.* 163 (7), 670–675. https://doi.org/10.1093/aje/kwj063.

Piazza, J.R., Almeida, D.M., Dmitrieva, N.O., Klein, L.C., 2010. Frontiers in the use of biomarkers of health in research on stress and aging. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 65 (5), 513–525. https://doi.org/10.1093/geronb/gbq049.

Qian, H., Wang, B., Ma, P., Peng, L., Gao, S., Song, Y. (2022). Managing Dataset Shift by Adversarial Validation for Credit Scoring. In: Khanna S, Cao J, Bai Q, Xu G, eds. PRICAI 2022: Trends in Artificial Intelligence. PRICAI 2022. Lecture Notes in Computer Science, vol 13629. Cham: Springer; 2022. https://doi.org/10.1007/978-3-031-20862-1_35.

Rössler, R., Wagner, J., Knaier, R., Rommers, N., Kressig, R.W., Schmidt-Trucksäss, A., Hinrichs, T., 2024. Spatiotemporal gait characteristics across the adult lifespan: Reference values from a healthy population - Analysis of the COmPLETE cohort study. *Gait Posture* 109, 101–108. https://doi.org/10.1016/j.gaitpost.2024.01.005.

Rydwik, E., Bergland, A., Forsén, L., Frändin, K., 2012. Investigation into the reliability and validity of the measurement of elderly people's clinical walking speed: a systematic review. *Physiother. Theory Pract.* 28 (3), 238–256. https://doi.org/10.3109/09593985.2011.601804.

Sahni, S., Dufour, A.B., Fielding, R.A., Newman, A.B., Kiel, D.P., Hannan, M.T., Jacques, P.F., 2021. Total carotenoid intake is associated with reduced loss of grip strength and gait speed over time in adults: The Framingham Offspring Study. *Am. J. Clin. Nutr.* 113 (2), 437–445. https://doi.org/10.1093/ajcn/nqaa288.

Sialino, L.D., Schaap, L.A., van Oostrom, S.H., Nooyens, A.C.J., Picavet, H.S.J., Twisk, J. W.R., Verschuren, W.M.M., Visser, M., Wijnhoven, H.A.H., 2019. Sex differences in physical performance by age, educational level, ethnic groups and birth cohort: The Longitudinal Aging Study Amsterdam. *PLoS One* 14 (12), e0226342.

Silverman, J.M., Schmeidler, J., Lee, P.G., Alexander, N.B., Beeri, M.S., Guerrero-Berroa, E., West, R.K., Sano, M., Nabozny, M., Rodriguez Alvarez, C., 2019. Associations of hemoglobin A1c with cognition reduced for long diabetes duration. *Alzheimer's & Dementia (new York* N. Y.), 5, 926–932. https://doi.org/10.1016/j.trci.2019.11.009.

Sousa, A.C., Zunzunegui, M.V., Li, A., Phillips, S.P., Guralnik, J.M., Guerra, R.O., 2016. Association between C-reactive protein and physical performance in older populations: results from the International Mobility in Aging Study (IMIAS). *Age Ageing* 45 (2), 274–280. https://doi.org/10.1093/ageing/afv202.

Straczkowski, M., Dzienis-Straczkowska, S., Stêpieñ, A., Kowalska, I., Szelachowska, M., Kinalska, I., 2002. Plasma interleukin-8 concentrations are increased in obese subjects and related to fat mass and tumor necrosis factor-alpha system. *J. Clin. Endocrinol. Metab.* 87 (10), 4602–4606. https://doi.org/10.1210/jc.2002-020135.

Studenski, S., Perera, S., Patel, K., et al., 2011. Gait speed and survival in older adults. JAMA 305 (1), 50–58. https://doi.org/10.1001/jama.2010.192.

Sugimoto, K., Ikegami, H., Takata, Y., et al., 2021. Glycemic Control and Insulin Improve Muscle Mass and Gait Speed in Type 2 Diabetes: The MUSCLES-DM Study. J Am Med Dir Assoc 22 (4), 834–838.e1. https://doi.org/10.1016/j.jamda.2020.11.003.

Supinski, G.S., Callahan, L.A., 2007. Free radical-mediated skeletal muscle dysfunction in inflammatory conditions. J Appl Physiol 1985 (102), 2056–2063.

Taaffe, D.R., Harris, T.B., Ferrucci, L., Rowe, J., Seeman, T.E., 2000. Cross-sectional and prospective relationships of interleukin-6 and C-reactive protein with physical performance in elderly persons: MacArthur studies of successful aging. J. Gerontol. A Biol. Sci. Med. Sci. 55 (12), M709–M715. https://doi.org/10.1093/gerona/55.12.m709.

Tanaka, T., Basisty, N., Fantoni, G., Candia, J., Moore, A.Z., Biancotto, A., Schilling, B., Bandinelli, S., Ferrucci, L., 2020. Plasma Proteomic Biomarker Signature of Age Predicts Health and Life Span. eLife, 9, e61073.

Tseng, L.A., Delmonico, M.J., Visser, M., Boudreau, R.M., Goodpaster, B.H., Schwartz, A. V., Simonsick, E.M., Satterfield, S., Harris, T., Newman, A.B., 2014. Body composition explains sex differential in physical performance among older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* 69 (1), 93–100. https://doi.org/10.1093/gerona/glt027.

Verghese, J., Holtzer, R., Oh-Park, M., Derby, C.A., Lipton, R.B., Wang, C., 2011. Inflammatory markers and gait speed decline in older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* 66 (10), 1083–1089. https://doi.org/10.1093/gerona/glr099.

Verghese, J., Holtzer, R., Lipton, R.B., Wang, C., 2012. High-sensitivity C-reactive protein and mobility disability in older adults. *Age Ageing* 41 (4), 541–545. https://doi.org/10.1093/ageing/afs038.

Wei, J., Xu, H., Davies, J.L., Hemmings, G.P., 1992. Increase of plasma IL-6 concentration with age in healthy subjects. *Life Sci.* 51 (25), 1953–1956. https://doi.org/10.1016/0024-3205(92)90112-3.

Westbury, L.D., Fuggle, N.R., Syddall, H.E., Duggal, N.A., Shaw, S.C., Maslin, K., Dennison, E.M., Lord, J.M., Cooper, C., 2018. Relationships Between Markers of Inflammation and Muscle Mass, Strength and Function: Findings from the Hertfordshire Cohort Study. *Calcif. Tissue Int.* 102 (3), 287–295 https://doi.org/10.1007.

Wheaton, F.V., Crimmins, E.M., 2016. Female disability disadvantage: a global perspective on sex differences in physical function and disability. *Ageing Soc.* 36 (6), 1136–1156. https://doi.org/10.1017/S0144686X15000227.

Wilson, D., Jackson, T., Sapey, E., Lord, J.M., 2017. Frailty and sarcopenia: The potential role of an aged immune system. *Ageing Res. Rev.* 36, 1–10. https://doi.org/10.1016/j.arr.2017.01.006.

Wu, T., Dorn, J.P., Donahue, R.P., Sempos, C.T., Trevisan, M., 2002. Associations of serum C-reactive protein with fasting insulin, glucose, and glycosylated hemoglobin: the Third National Health and Nutrition Examination Survey, 1988-1994. *Am. J. Epidemiol.* 155 (1), 65–71. https://doi.org/10.1093/aje/155.1.65.

Yoon, J.W., Ha, Y.C., Kim, K.M., Moon, J.H., Choi, S.H., Lim, S., Park, Y.J., Lim, J.Y., Kim, K.W., Park, K.S., Jang, H.C., 2016. Hyperglycemia Is Associated with Impaired Muscle Quality in Older Men with Diabetes: The Korean Longitudinal Study on Health and Aging. *Diabetes Metab. J.* 40 (2), 140–146. https://doi.org/10.4093/dmj.2016.40.2.140.