Muscle failure and risk of adverse outcomes in older adults: a derivation and multicohort validation study

Hélio José Coelho-Júnior, Alejandro Álvarez-Bustos, Isabel Rodríguez-Sánchez, Francesco Landi, Emanuele Marzetti

Summary

Background Sarcopenia is inconsistently defined, and its definition might not adequately capture muscle failure, underlying disability, and other adverse outcomes in older adults. We aimed to develop and validate an evidence-based, outcome-driven model of muscle failure.

Methods In this derivation and multicohort validation study, we analysed data from the SPRINTT randomised controlled trial (1519 participants; mean age 78·9 years [SD 5·8]; 431 [28·4%] men, 1088 (71·6%) women) and validated findings in five independent cohorts (ilSIRENTE, NHANES, HRS, ELSA, and CHARLS). Neuromuscular domains including appendicular lean mass, muscle strength and power, mobility, and physical activity were assessed. Principal component analysis identified clusters of measures, which were tested for associations with disability in mobility and disability in activities of daily living, hospitalisation, and mortality using adjusted regression models. Predictive performance was evaluated with receiver operating characteristic curves. Area under the curve (AUC) values were defined as acceptable (0·7–0·8), excellent (0·8–0·9), or outstanding (>0·9).

Findings We identified two major muscle failure indexes: mobility (short-distance walking speed, 400 m walk) and physical activity (step counts, time spent standing, stepping, and sitting). In SPRINTT, the mobility model strongly predicted disability (AUC 0·721, 95% CI 0·698–0·795; p<0·0001), whereas the physical activity model was more closely associated with hospitalisation (0·646, 0·570–0·731; p=0·00040) and mortality (0·746, 0·612–0·926; p=0·0020). Combined models including mobility, handgrip strength, appendicular lean mass, and absolute muscle power improved prediction of mortality (0·746, 0·549–0·955; p=0·0076). External validation confirmed acceptable-to-strong discrimination for disability outcomes across cohorts (AUCs 0·7–0·9), while associations with hospitalisation and mortality were weaker and inconsistent (0·4–0·9). Compared with consensus-based sarcopenia definitions, the evidence-based model seemed to show better predictive ability for disability.

Interpretation Distinct mobility measures provide a robust framework for identifying muscle failure and predicting disability in older adults. Continuous activity monitoring might improve prediction of hospitalisation and mortality. This outcome-driven approach supports refinement of sarcopenia assessment and its implementation in clinical practice.

Funding None.

Copyright © 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Lancet Healthy Longev 2026

Published Online https://doi.org/10.1016/j.lanhl.2026.100843

Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy (H J Coelho-Júnior PhD, Prof F Landi MD, E Marzetti MD); **Department of Geriatrics, Orthopaedics and Rheumatology, Università Cattolica del Sacro Cuore, Rome, Italy** (H J Coelho-Júnior, Prof F Landi, E Marzetti); **Department of Health, Villanueva University, Madrid, Spain** (A Álvarez-Bustos PhD); **Geriatrics Department and Health Research Institute, Hospital Universitario Clínico San Carlos, Madrid, Spain** (I Rodríguez-Sánchez MD)

Correspondence to: Dr Emanuele Marzetti, Department of Geriatrics, Orthopaedics and Rheumatology, Università Cattolica del Sacro Cuore, Rome 00168, Italy **emanuele.marzetti@policlinicogemelli.it**

Introduction

The term sarcopenia was originally coined to describe the age-related decline in lean body mass affecting mobility and energy metabolism.[1] Initially regarded as a physiological corollary of ageing, sarcopenia is now recognised as a clinical condition, with standardised diagnostic criteria and formal classification as a disease.[2,3] While this transition has brought greater visibility to sarcopenia in clinical and research settings, important uncertainties remain.

Prevalence estimates and risk factors differ widely across studies and populations, reflecting heterogeneity in definitions, measurement approaches, and thresholds.[4–6] Current diagnostic frameworks rely largely on expert consensus, combining measures of muscle mass, strength, and physical performance. Although these assessments are clinically intuitive, they show restricted and inconsistent ability to predict disability, hospitalisation, or mortality. This limitation has led to criticism that existing definitions of sarcopenia might not adequately capture the broader construct of muscle failure, which underlies functional decline and adverse outcomes in older adults.[7,8]

To address these challenges, evidence-based approaches are needed that move beyond consensus definitions and that instead anchor diagnostic constructs in outcomes of clinical relevance. We have previously proposed that such an outcome-driven model should integrate multiple neuromuscular domains, assessed through robust methods and validated against hard endpoints.[7]

This premise recognises that impairments in muscle biology, as observed in sarcopenia, are unlikely to affect a single parameter in isolation. Rather, muscle failure spans a functional spectrum, from dynapenia (eg, low handgrip strength) to deficits in more complex behaviours requiring coordinated physical domains and physiological systems,

Research in context

Evidence before this study

We searched PubMed for English-language articles published between Jan 1, 1988, and Feb 11, 2026, including trials, observational studies, and meta-analyses examining associations between neuromuscular parameters and sarcopenia with adverse outcomes in older adults. Search terms combined synonyms of "sarcopenia", "neuromuscular function", "muscle strength", "muscle mass", and "muscle power" with "disability", "hospitalisation", or "mortality". Numerous studies have reported inconsistent findings regarding the predictive ability of current sarcopenia definitions. Estimates of prevalence and associated risk factors vary widely, largely due to heterogeneity in diagnostic criteria. Existing operational definitions have also been criticised for relying heavily on expert consensus rather than transparent, standardised methodologies, and for employing heterogeneous assessment tools that might not capture comparable aspects of muscle health. To overcome these limitations, we have recently proposed that a valid, evidence-based model of muscle failure should integrate multiple neuromuscular measures and be derived through robust methodological approaches, including advanced statistical modelling.

Added value of this study

This multicohort study provides validation of an evidence-based muscle failure model across independent, international datasets, demonstrating its ability to predict disability, hospitalisation, and mortality among community-dwelling older adults (aged ≥70 years). Unlike current sarcopenia definitions, which often rely on expert consensus and heterogeneous measures, our model incorporates multiple neuromuscular parameters identified through robust statistical modelling. We show that a construct combining short-distance walking speed and 400 m walk performance strongly predicts disability, whereas a construct reflecting habitual physical activity is more predictive of hospitalisation and mortality. Notably, these associations remained significant when the model was adapted to the available data in each validation cohort, suggesting that practical, scalable proxies, including self-reported measures, may be sufficient in certain contexts. Overall, our findings challenge the prevailing reliance on single performance measures for sarcopenia assessment and underscore the multidimensional nature of muscle function in relation to ageing outcomes.

Implications of all the available evidence

Current approaches to identifying sarcopenia demonstrate limited and inconsistent predictive ability for major adverse outcomes in older adults. Our findings support a shift towards multidimensional, evidence-based models of muscle failure that integrate complementary neuromuscular measures. For predicting disability, combining short-distance walking speed with 400 m walk performance captures distinct yet synergistic aspects of mobility that are essential for daily functioning. For hospitalisation and mortality, monitoring physical activity provides a more accurate reflection of overall risk. Adoption of such models in both research and clinical settings could enhance the identification of at-risk individuals, guide tailored interventions, and optimise the use of health-care resources. Wider implementation could be facilitated through wearable technologies or simple self-report tools, making these models adaptable across diverse populations, settings, and health-care systems. This approach offers the potential to redefine sarcopenia as a clinically actionable, prognostically meaningful construct, ultimately improving outcomes in ageing populations.

including mobility and habitual physical activity. Age-related declines across these inter-related dimensions are well documented,[9] and adverse outcomes such as disability typically reflect the cumulative effects of dysfunction across multiple neuromuscular domains.[10,11] Accordingly, models integrating complementary measures of muscle function might better capture clinically meaningful muscle failure and improve risk prediction in older adults compared with consensus-based frameworks. An alternative view is that selected single parameters may independently predict adverse outcomes, with some metrics demonstrating superior performance.[12–14] This perspective raises the question of whether muscle failure is best captured by individual neuromuscular markers or by integrated, multidimensional constructs.

In this study, we aimed to identify which individual or combined neuromuscular measures best predict disability, hospitalisation, and mortality in older adults using data from the Sarcopenia and Physical Frailty in Older People: Multi-Component Treatment Strategies (SPRINTT) trial[15] together with external validation cohorts. By linking diagnostic criteria to prospective outcome data, our objective was to advance a clinically meaningful framework for defining and managing muscle failure.

See Online for appendix

Methods

Study design and reference study

For this derivation and multicohort validation study, we undertook model development and validation in four sequential stages (figure). First, candidate models were identified within the reference SPRINTT database. Second, their associations with adverse outcomes (ie, mobility disability and activities of daily living [ADL] disability, hospitalisation, and death) were examined, and their predictive ability quantified. Third, candidate models were adapted using data from five external validation cohorts. Finally, the predictive performance of the adapted models was assessed in these external cohorts. The study was conducted in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (appendix p 5–7).[16]

The SPRINTT database was used as the basis for creating the muscle failure model.[15] SPRINTT was a multicentre, evaluator-blinded, randomised controlled trial designed to test whether a multicomponent intervention comprising structured physical activity with technological support and nutritional counselling reduced the risk of mobility disability in older adults with functional limitations, compared with a lifestyle education programme (ClinicalTrials.gov, NCT02582138).[15] The study took place in 16 clinical sites across 11 European countries, between January, 2016, and October, 2019. Eligible participants were community-dwelling individuals aged 70 years or older with mobility limitations. Inclusion criteria were: (1) a short physical performance battery (SPPB) score of 3 to 9; (2) low appendicular lean mass (ALM; absolute or BMI-adjusted), defined by sex-specific cutoffs; and (3) absence of mobility disability, operationalised as the ability to complete a 400 m walk in less than 15 min without sitting, stopping for more than 1 min, receiving assistance, or using a walker. Exclusion criteria included self-reported walking disability, cognitive impairment (Mini Mental State Examination [MMSE] score <24), terminal illness, participation in a structured physical activity programme, contraindications to safe participation as judged by local investigators, or anticipated relocation outside the study area within at least 2 years. Participants were randomly assigned to either the multicomponent intervention (n=760) or lifestyle education on healthy ageing (control group, n=759). Participants received the multicomponent intervention or lifestyle education for up to 36 months, depending on time of enrolment.

The trial protocol was approved by the ethics committee of the Università Cattolica del Sacro Cuore (Rome, Italy; coordinating centre) and by local ethics committees at each participating site. The trial was conducted in accordance with the Declaration of Helsinki, and all participants provided written informed consent before screening. This study was an analysis of the SPRINTT database and therefore no further ethical approval was required.

Candidate neuromuscular measures

Several neuromuscular domains were evaluated in SPRINTT as candidate markers of muscle failure. ALM was measured by whole-body dual-energy x-ray absorptiometry. Upper-limb and lower-limb strength were assessed with handgrip dynamometry and the five-time sit-to-stand (5STS) test, respectively, according to standardised protocols. Mobility was assessed using the 400 m walk test, in which participants completed ten laps on a 20 m linear course at their usual pace without overexertion (each lap consisted of a 40 m out-and-back distance). Muscle power was derived from 5STS using established equations, generating absolute (AMP), relative (RMP), allometric (ALMP), and specific muscle power (SMP) indices (appendix p 8). Lower-extremity function was assessed with the SPPB, comprising balance, 4-m gait speed, and 5STS;[17] each SPPB component was scored from 0 to 4, with higher scores indicating better performance, and summed to yield a total score ranging from 0 to 12. To incorporate muscle power, the conventional 5STS score was replaced with sex-specific quartiles of AMP, RMP, ALMP, or SMP (appendix p 9), yielding four indices: $SPPB_{AMP}$, $SPPB_{RMP}$, $SPPB_{ALMP}$, and $SPPB_{SMP}$. Physical activity was measured with the activPAL device (PAL Technologies, Glasgow, UK), worn on a thigh continuously for 7 consecutive days within 2 weeks of each clinical visit. Device outputs were processed to quantify daily step count and time spent sitting, standing, and stepping.

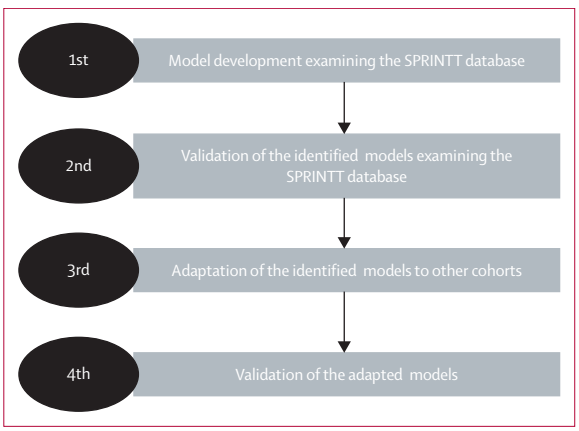

Figure: **Development and validation of the muscle failure model**

Adverse outcomes in SPRINTT

Follow-up visits were conducted at 3 months after randomisation and every 6 months from baseline. Mobility disability was defined as the inability to complete the 400 m walk test within 15 min without sitting, stopping for more than 1 min, requiring assistance from another person, or using a walker. Mobility disability was considered to have occurred at the earliest date participants (a) failed the 400 m walk test, (b) attended a scheduled visit and were unable to attempt the test, or (c) died. A standardised adjudication procedure was applied when the 400 m walk test was not performed. Disability in ADL was assessed using the six-item Katz index derived from the Pepper Assessment Tool for Disability.[18] Hospitalisation data were obtained at follow-up visits through participant self-report. Information on deaths was collected through proxy report, hospital records, and linkage with national registries.

Covariates

Age and sex were recorded for all participants. Global cognitive function was assessed using the MMSE. Nutritional status was screened using the Mini Nutritional Assessment Short Form (MNA–SF), which includes six questions (questions A–F2) covering appetite, weight loss, mobility, stress, neuropsychological problems, and BMI, scored 0–14.[19] Participants were classified as well nourished (score 12–14), at risk of malnutrition (8–11), or malnourished (0–7), according to standard cutoff values.

Multimorbidity was defined as the presence of three or more of the following conditions: diabetes, cardiovascular disease, osteoarthritis, or cancer.[20] Information on disease conditions was obtained through physician-administered interviews, including participant-reported or proxy-reported diagnoses. Medication inventories were reviewed to support diagnostic adjudication when prescriptions clearly indicated a specific condition. Bodyweight-adjusted protein intake was calculated from a 3-day dietary record collected annually in those allocated in the multicomponent intervention. Randomisation group (multicomponent intervention or lifestyle education) was also included as a covariate in the analyses.

Validation cohorts

Five external cohorts were used to validate muscle failure models (full cohort characteristics, assessment methods, and references are provided in appendix pp 10–18). The data collected spanned different time periods (iSIRENTE 2003–13, NHANES 1999–2000, HRS 1999–2022, ELSA 2004–09, and CHARLS 2011–13).

The Ageing and Longevity Study in the Sirente Geographic Area (ilSIRENTE) was a population-based longitudinal study in Italy examining ageing, frailty, and functional decline in community-dwelling adults aged 80 years and older.[21] Assessments included physical performance (ie, walking speed, 5STS, and balance), neuromuscular measures, cognitive and psychological evaluations, comorbidities, medications, and nutritional status, with longitudinal follow-up for disability, hospitalisation, and mortality.

The National Health and Nutrition Examination Survey (NHANES) is a cross-sectional survey of the non-institutionalised US population, collecting data on demographics, health behaviours, medical history, laboratory measures, and physical performance (handgrip strength and walking speed).[22]

The Health and Retirement Study (HRS) is a longitudinal study of US adults aged 50 years or older, capturing health, physical and cognitive function, and socioeconomic and psychosocial factors. Physical performance assessments include handgrip strength, walking speed, and the 5STS. Data from Wave 10 served as the baseline, while Wave 14 data were used for the follow-up assessment.[23]

The English Longitudinal Study of Ageing (ELSA) is a longitudinal cohort study of adults aged 50 years or older in England, with repeated assessments of physical performance (balance, walking speed, handgrip strength, and 5STS), cognition, health behaviours, chronic conditions, and biological markers. Data from Wave 2 served as the baseline, while Wave 4 data were used for the follow-up assessment.[24]

The China Health and Retirement Longitudinal Study (CHARLS) is a nationally representative longitudinal study of Chinese adults aged 45 years or older, designed to investigate ageing, health, and socioeconomic factors.[25] Baseline data were collected in 2011, with follow-up waves every 2–3 years, including assessments of demographics, health status, comorbidities, medications, socioeconomic and family characteristics, and detailed physical and cognitive measures.

All cohorts were granted relevant local ethics approval.

Statistical analysis

Data from SPRINTT participants were analysed for the overall sample. Data are presented as mean (SD) for continuous variables and as frequency (%) for categorical variables. We used principal component analysis (PCA) to identify latent constructs underlying the neuromuscular measures and to reduce data dimensionality while preserving the variance shared across variables. This approach enabled the derivation of data-driven neuromuscular constructs reflecting muscle failure as it manifests across multiple functional domains, rather than relying on predefined or consensus-based classifications. We examined the correlation structure among variables to confirm suitability for PCA, ensuring sufficient intercorrelation to support the presence of shared underlying dimensions.

We did the component extraction using principal axis factoring, which is appropriate for identifying latent constructs when the focus is on shared rather than total variance. We determined the number of components retained using a combination of Kaiser's criterion (eigenvalues >1) and visual inspection of the scree plot to identify inflection points. This combined approach balanced statistical criteria with interpretability and minimised the risk of overextraction.

To enhance interpretability and achieve a simpler structure, we applied orthogonal Varimax rotation. We examined rotated factor loadings to assess the contribution of each neuromuscular variable to the retained components. We used variables with the highest loadings to define the conceptual meaning of each component, which we labelled accordingly. Only components that were both statistically robust and clinically interpretable were retained for subsequent analyses.

We calculated component scores for each participant for the retained components. These scores were then combined into composite clusters, derived as weighted sums of the corresponding component scores, to represent multidimensional neuromuscular constructs. We then entered the PCA-derived components and clusters as independent variables in regression models to examine their associations with mobility disability, ADL disability, hospitalisation, and mortality.

We examined the associations between individual neuromuscular markers, PCA-derived clusters, and composite clusters using binary logistic regression for ADL disability and Cox proportional hazards models for mobility disability, hospitalisation, and mortality, with time-to-event as the time variable. Models were adjusted for age, sex, MMSE score, MNA–SF score, multimorbidity, and randomisation group. Model performance was compared using the Akaike information criterion

(AIC), with models showing a difference in AIC (ΔAIC) of 2 or lower considered to have comparable support.

For each outcome, we tested alternative representations of the same PCA-derived construct across different PCA specifications. Because these models captured the same latent construct and demonstrated similar model fit, only the best-fitting representation (lowest AIC) for each construct is presented in the main tables. Results from alternative specifications were highly consistent and are therefore not shown to improve clarity.

We performed receiver operating characteristic curve analyses to assess the discriminative performance of muscle failure models for time-to-event outcomes. Area under the curve (AUC) values were interpreted as acceptable (0·7–0·8), excellent (0·8–0·9), or outstanding (>0·9). In addition, we applied the European Working Group on Sarcopenia in Older People 2 (EWGSOP) paradigm to calculate AUC values for a consensus-based definition of sarcopenia, allowing comparison of the discriminative performance of our muscle failure models with an established and widely used sarcopenia framework. Time at risk was defined as the interval between baseline assessment and occurrence of the event of interest; participants who did not experience the event were censored at their last available observation. We evaluated the proportional hazards assumption using log–minus–log survival plots, which showed no substantial deviations from proportionality.

We applied the same analytical framework in the validation cohorts, with cohort-specific methods described in the appendix (pp 11–18). We calculated effect sizes, including odds ratios (ORs) and hazard ratios (HRs) with their respective 95% CIs for all analyses, and we considered p values lower than 0·05 significant. Cases with missing data were handled by listwise deletion, in accordance with SPSS procedures. Analyses were done using SPSS version 23 (IBM, New York, NY, USA).

Role of the funding source

There was no funding source for this study.

Results

We extracted data from the 1519 participants in the SPRINTT study (mean age 78·9 years [SD 5·8]; 431 [28·4%] men, 1088 (71·6%) women; table 1). Baseline data for ilSIRENTE, NHANES, HRS, and ELSA are reported in the appendix (pp 19–24). For the PCA, based on eigenvalues and visual inspection, two consistent clusters emerged: cluster 1 (physical activity), comprising total steps, standing time, stepping time, and sitting time; and cluster 2 (mobility), including 4 m gait speed and 400 m walk time (appendix pp 25–32). A third cluster, cluster 3 (multiple neuromuscular parameters), varied across analyses and included different combinations of cluster 1, handgrip strength, ALM, total steps, sitting time, stepping time, 5STS, and RMP. Variables derived from PCA for subsequent analyses are detailed in the appendix (pp 33–35).

	Values
Age, years	78·9 (5·8)
Sex	
Female	1088 (71·6%)
Male	431 (28·4%)
BMI, kg/m^2	28·5 (5·7)
Multimorbidity	325 (21·4%)
Mini Mental State Examination score	27·9 (1·8)
Mini Nutritional Assessment Short Form score	12·7 (1·6)
Isometric handgrip strength, kg	
Men	28·5 (9·2)
Women	16·9 (5·5)
Five-time sit-to-stand, s	19·2 (6·6)
Appendicular lean mass, kg	
Men	21·0 (3·4)
Women	14·7 (2·1)
4 m walking speed, s	5·8 (1·6)
Time to walk 400 m, s	522·3 (148·0)
Steps, n/day	6736·5 (3272·9)
Sitting and lying activities, h/day	17·1 (2·7)
Standing, h/day	5·2 (2·1)
Stepping, h/day	1·5 (0·6)
Mobility disability at follow-up	683 (4508%)
ADL disability at follow-up	36 (2·4%)
Hospitalisation at follow-up	483 (31·8%)
Death at follow-up	64 (4·2%)
Randomisation group	
Multicomponent intervention	760 (50·0%)
Lifestyle education	759 (50·0%)

Data are n (%) or mean (SD). ADL=activities of daily living.

Table 1: **Main characteristics of SPRINTT participants (n=1519)**

Unadjusted and adjusted Cox and binary regression analyses examining associations between significant principal components (PCs) and mobility disability, ADL disability, hospitalisation, and mortality outcomes are reported in the appendix (pp 36–51). In adjusted models, multiple physical performance measures, PCs, and combined clusters were significantly associated with mobility disability, hospitalisation, and mortality, whereas only PCs were associated with ADL disability. AIC comparisons indicated that models including cluster 2 (mobility) provided the best fit for predicting both mobility (AIC 2788·781) and ADL disability (230·3), whereas models including cluster 1 (physical activity) best predicted hospitalisation (2313·441) and death (334·309). For mortality, the optimal model was cluster 3 (multiple neuromuscular parameters), which included cluster 1 (physical activity), handgrip strength, AMP, and ALM (AIC= 333·695; table 2).

Predictive performance, adjusted for covariates, was evaluated for mobility disability, hospitalisation, and death using AUC analyses (table 3; appendix pp 52–54). The mobility model (cluster 2) demonstrated acceptable discrimination for disability (AUC 0·721, 95% CI 0·698–0·795; p<0·0001), while the physical activity model (cluster 1) and multiple neuromuscular parameters model (cluster 3)

	−2LL	AIC	ΔAIC
Mobility disability			
Cluster 2 (mobility)	2772·781	2788·781	0·000
Cluster 3 (multiple neuromuscular parameters)	2805·493	2821·493	37·702
Cluster 1 (physical activity)	2809·067	2825·067	38·766
ADL disability			
Cluster 2—Mobility	214·3	230·3	0·10
Hospitalisation			
Cluster 1 (physical activity)	2297·441	2313·441	0·000
Cluster 2 (mobility)	2305·200	2321·200	7·759
Cluster 3 (multiple neuromuscular parameters)	2305·636	2321·636	8·195
Stepping activity	2503·035	2519·035	205·594
Total steps	2505·411	2521·411	207·970
Standing activity	2505·517	2521·517	208·076
Sitting and lying activity	2507·631	2523·631	210·190
Handgrip strength	2776·212	2792·212	478·771
SPPB	2902·987	2918·987	605·546
400 m	2905·531	2921·531	608·090
Death			
Cluster 3 (multiple neuromuscular parameters)	317·695	333·695	0·000
Cluster 1 (physical activity)	318·309	334·309	0·614
Cluster 2 (mobility)	318·891	334·891	1·196
Total steps	332·135	348·135	14·440
Stepping activity	332·159	348·159	14·464
SPPB	397·687	413·687	79·992

ΔAIC ≤2 indicates comparable support. Only significant models in the Cox and binary regression were analysed. For each outcome, multiple representations of the same PCA-derived construct were tested across different PCA specifications. Because these models reflect the same latent construct and showed highly comparable model fit, only the best-fitting representation (lowest AIC) for each construct is reported in the main tables to improve clarity. Results for alternative PCA specifications were highly consistent and are therefore not shown. −2LL=minus twice the log-likelihood. ADL=activities of daily living. AIC=Akaike information criterion. PCA=principal component analysis. SPPB=Short Physical Performance Battery. ΔAIC=difference in AIC relative to the best-fitting model for each outcome (ΔAIC=0).

Table 2: **Akaike information criteria for significant muscle failure index**

	AUC (95% CI)	p value*
Mobility disability		
Cluster 2 (mobility)	0·721 (0·698–0·795)	<0·0001
Hospitalisation		
Cluster 1 (physical activity)	0·646 (0·570–0·731)	0·00040
Death		
Cluster 1 (physical activity)	0·746 (0·612–0·926)	0·0020
Cluster 3 (multiple neuromuscular parameters)	0·746 (0·594–0·955)	0·0076

AUC=area under the curve. *p values correspond to tests of the null hypothesis that the AUC equals 0·5.

Table 3: **AUC for study outcomes**

showed acceptable discrimination for mortality (cluster 1: 0·746, 0·612–0·926; p=0·0020; cluster 3: 0·746, 0·594–0·955; p=0·0076). The physical activity model (cluster 1) performed poorly for hospitalisation (0·646, 0·570–0·731, p=0·00040).

External validation was conducted for cluster 1 and cluster 2, as insufficient data were available to develop an adapted cluster 3 model. The mobility model (cluster 2) showed the most consistent associations with disability outcomes (appendix pp 55–71). In NHANES, adapted cluster 2 was significantly associated with ADL disability (OR 683·2, 95% CI 197·6–2361·9; p<0·0001) and mobility disability (8·6, 7·5–9·8; p<0·0001), and AUC analyses demonstrated acceptable discrimination for ADL disability (AUC 0·7, 95% CI 0·735–0·812; p<0·0001; appendix pp 59–60). In ilSIRENTE, adapted cluster 2 was associated with high odds of ADL disability (OR 43·6, 7·0–243·5; p<0·0001), and AUC analyses indicated strong predictive performance (AUC 0·9, 0·938–0·978; appendix pp 55–58). Similar findings were observed in HRS, in which adapted cluster 2 was significantly associated with ADL disability (HR 0·233, 95% CI 0·1–0·5; p<0·0001), and AUC analyses showed strong predictive performance (AUC 0·9, 0·840–1·000; appendix pp 61–64). In ELSA and CHARLS, adapted cluster 2, handgrip strength, and 5STS were significantly associated with mobility disability, with acceptable discrimination (appendix pp 65–71). On the other hand, only cluster 2 was significantly associated with ADL disability.

Associations involving the physical activity model (cluster 1) and hospitalisation or mortality were more heterogeneous across cohorts (appendix pp 56–71). In ilSIRENTE, adapted cluster 1 was associated with mortality (OR 1·8, 95% CI 1·6–2·0; p<0·0001; appendix pp 57–58), and in HRS, adapted cluster 1 was associated with hospitalisation (HR 84·8, 95% CI 1·2–589·8; p=0·040; appendix pp 62–64). In ELSA and CHARLS, no measures were significantly associated with hospitalisation or death (appendix pp 70).

In ilSIRENTE, probable sarcopenia as defined by EWGSOP was significantly associated with increased odds of ADL disability (OR 4·5, 95% CI 1·7 to 11·8; p<0·0001; AUC 0·9; p<0·0001) but not with hospitalisation (0·1, −0·7 to 0·8; p=87) or death (0·1, −0·9 to 1·2; p=0·79; appendix pp 55–58). In HRS, probable sarcopenia as defined by EWGSOP was not significantly associated with ADL disability, hospitalisation, or death (appendix pp 61–63). Data required to operationalise the EWGSOP and Asian Working Group for Sarcopenia definitions were not available in the ELSA and CHARLS cohorts, respectively.

Discussion

The present study supports a multicentric, outcome-driven model of muscle failure to predict adverse outcomes in community-dwelling older adults. Distinct mobility dimensions were the strongest predictors of ADL and mobility disability, whereas overall physical activity was more closely associated with hospitalisation and mortality. Across diverse cohorts, these models, even when based on self-reported proxies, performed comparably to or better than the EWGSOP paradigm. Collectively, these findings indicate that this outcome-based framework merits further evaluation and integration into both research and clinical practice.

The predictive strength of the mobility model probably reflects the multiple capacities captured by 4 m gait speed and the 400 m walk, both fundamental to ADL and mobility tasks.[26] Walking integrates cardiovascular, respiratory,

neuromuscular, and endocrine functions. From a kinesiological perspective, short-distance (4–10 m) and long-distance (400 m) walks provide complementary assessments. Short walks primarily capture strength, power, and balance,[12,27] whereas the 400 m walk imposes an aerobic challenge, typically lasting around 10 min in functionally limited older adults, with oxygen requirements similar to those of light-to-moderate intensity exercise.[28]

ADL performance similarly requires coordinated activation of multiple capacities.[10,11] Everyday tasks such as bathing, dressing, and transferring involve standing, walking, and sitting, all of which depend on muscle strength, power, and balance.[29] The relevance of the 400 m walk might stem from its ecological similarity to repeated functional demands of daily living. The principle of specificity, whereby training or testing adaptations are greatest in the tasks most closely mimicked,[30] further explains the strong association of the 400 m walk with disability. Therefore, short-distance walking speed and the 400 m walk capture complementary aspects of mobility, ranging from crossing a room to walking a block, making their combination highly effective for predicting disability across diverse populations.

These results align with evidence that other physical capacities, including muscle strength as defined in the EWGSOP paradigm,[3] are associated with disability. However, such measures capture only part of the multifaceted nature of functional performance. The stronger association observed with the mobility model suggests that combining short and long walking assessments provides a more holistic metric of coordinated task performance.

In contrast, the physical activity model showed stronger associations with hospitalisation and mortality, probably reflecting its broader capture of overall health status. However, variability across cohorts in these associations suggests caution in considering these endpoints primary for sarcopenia.

Although further validation is needed, our findings suggest that muscle failure is most effectively identified using two complementary mobility tests to predict disability. For outcomes beyond disability, integrating wearable or smartphone-based monitoring over a week could provide practical data on activity, walking speed, and distance, enabling feasible and routine assessment in clinical practice. Self-reported proxies of mobility performed reasonably well, supporting applicability where objective measures are unavailable.

A central contribution of our results is the shift from a consensus-based definition of sarcopenia to an outcome-driven construct of muscle failure. Rather than redefining mobility limitation, this framework positions impaired walking as one observable manifestation of compromised muscle function, anchored to prospective endpoints such as disability. It thereby provides an empirical foundation largely absent from previous definitions and aligns clinical priorities with research paradigms traditionally focused on muscle mass or isolated strength.

Building on this shift, the framework offers several opportunities for research. It supports harmonised outcome definitions across cohorts, justifies the use of simple walking measures as primary endpoints in trials, and enables integration of wearable technologies for longitudinal, real-world assessment. The clinical relevance of the framework is equally compelling. Short-distance walking speed and the 400 m walk test are low-cost, feasible, and are both (though the 400 m walk test is used to a lesser extent in practice) already embedded in geriatric practice, facilitating scalable implementation and earlier risk identification. Wearable-derived activity metrics can further enhance remote monitoring and personalised intervention. Collectively, these findings advance an outcome-driven, multidimensional conceptualisation of sarcopenia as muscle failure.

Notwithstanding its strengths, the study has several limitations. Participants in SPRINTT were preselected for mobility limitations, which might restrict generalisability, although this population does closely reflect older adults who typically seek medical attention for functional impairments in primary care. ADL outcomes were analysed as binary variables rather than time-to-event outcomes due to limited follow-up. Objective 400 m walk test data were unavailable in validation cohorts, and self-reported proxies might differ from measured outcomes. Adaptations were also required for physical activity, hospitalisation, and covariates. Although nutritional status was accounted for, daily dietary intake—and thus total energy and macronutrient composition—was not assessed. These factors play an important part in neuromuscular function and might influence the risk of adverse outcomes. Cohort heterogeneity, while supporting the robustness of the model, might limit extrapolation to healthier older populations. Data were also lacking from South America, central America, and Africa. Some external validation cohorts used complex survey sampling designs. Analyses were conducted without applying sampling weights because the primary objective was to assess associations and predictive performance rather than to obtain population-representative estimates. Although this approach enhanced methodological harmonisation and cross-cohort comparability, it might have limited the generalisability of effect estimates to national populations. In addition, missing data were handled using case-wise deletion, which might have introduced bias if the excluded observations differed systematically from those included in the analyses. Finally, the combined model (cluster 3) incorporating cluster 1, handgrip strength, AMP, and ALM could not be validated in most validation cohorts. Therefore, the possibility that this composite approach provides superior prediction of mortality cannot be excluded.

In summary, by empirically anchoring sarcopenia definitions in clinically relevant outcomes such as disability, this study provides a pragmatic redefinition of muscle failure. Its innovation lies in combining conventional walking assessments with wearable-based monitoring, yielding a multidimensional yet feasible framework. The

implications extend across trial methodology, clinical screening, and digital health integration, offering a pathway to transform how sarcopenia is identified, monitored, and managed in health-care systems.

Contributors

HJC-J and EM conceived the study, reviewed the literature, analysed and interpreted the data, and drafted the original manuscript. HJC-J and AÁ-B analysed and interpreted the data and contributed to manuscript revision. HJC-J and IR-S contributed to data analysis, interpretation, and critical revision of the manuscript. FL provided supported interpretation of findings and contributed to manuscript revision. HJC-J, AÁ-B, IR-S, and EM verified and reviewed the data to ensure consistency. All authors approved the final version of the manuscript and had full access to the data. HJC-J and EM had primary responsibility for the decision to submit for publication.

Declaration of interests

FL and EM declare funding from the Innovative Medicines Initiative Joint Undertaking (grant IMI-JU 115621). FL and EM received in-kind support from the European Federation of Pharmaceutical Industries and Associations as part of the Innovative Medicines Initiative Joint Undertaking for the submitted work. EM received personal fees from Pfizer outside the submitted work. All other authors declare no competing interests.

Data sharing

Anonymised raw data of the SPRINTT trial can be shared upon request to Luca Mariotti (luca.mariotti1@unicatt.it). A data access agreement must be signed. Data from the ilSIRENTE study can be obtained from FL upon reasonable request. Data from the ELSA, HRS, and CHARLS studies are publicly available and can be accessed through their respective data repositories subject to their data access policies.

Acknowledgments

The Article processing charge was funded by the Italian Ministry of Health (Ricerca Corrente 2026).

References

1 Rosenberg IH. Sarcopenia: origins and clinical relevance. *J Nutr* 1997; **127** (5 suppl): 990S–91S.

2 Chen LK, Woo J, Assantachai P, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. *J Am Med Dir Assoc* 2020; **21:** 300–07.e2.

3 Cruz-Jentoft AJ, Bahat G, Bauer J, et al, and the Writing Group for the European Working Group on Sarcopenia in Older People 2 (EWGSOP2), and the Extended Group for EWGSOP2. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing* 2019; **48:** 601.

4 Petermann-Rocha F, Balntzi V, Gray SR, et al. Global prevalence of sarcopenia and severe sarcopenia: a systematic review and meta-analysis. *J Cachexia Sarcopenia Muscle* 2022; **13:** 86–99.

5 Davies B, Walter S, Rodríguez-Laso A, et al. Differential association of frailty and sarcopenia with mortality and disability: insight supporting clinical subtypes of frailty. *J Am Med Dir Assoc* 2022; **23:** 1712–16.e3.

6 Schaap LA, van Schoor NM, Lips P, Visser M. Associations of sarcopenia definitions, and their components, with the incidence of recurrent falling and fractures: The Longitudinal Aging Study Amsterdam. *J Gerontol A Biol Sci Med Sci* 2018; **73:** 1199–204.

7 Coelho-Júnior HJ, Marzetti E. Capturing what counts in muscle failure: a critical appraisal of the current operational models of sarcopenia. *Lancet Healthy Longev* 2025; **6:** 100756.

8 Evans WJ, Guralnik J, Cawthon P, et al. Sarcopenia: no consensus, no diagnostic criteria, and no approved indication: how did we get here? *Geroscience* 2024; **46:** 183–90.

9 Coelho-Junior HJ, Uchida MC, Gonçalves IO, et al. Age- and gender-related changes in physical function in community-dwelling Brazilian adults aged 50 to 102 years. *J Geriatr Phys Ther* 2021; **44:** E123–31.

10 Manini T. Development of physical disability in older adults. *Curr Aging Sci* 2011; **4:** 184–91.

11 Guralnik JM, Fried LP, Salive ME. Disability as a public health outcome in the aging population. *Annu Rev Public Health* 1996; **17:** 25–46.

12 Cuoco A, Callahan DM, Sayers S, Frontera WR, Bean J, Fielding RA. Impact of muscle power and force on gait speed in disabled older men and women. *J Gerontol A Biol Sci Med Sci* 2004; **59:** 1200–06.

13 Bean JF, Leveille SG, Kiely DK, Bandinelli S, Guralnik JM, Ferrucci L. A comparison of leg power and leg strength within the InCHIANTI study: which influences mobility more? *J Gerontol A Biol Sci Med Sci* 2003; **58:** 728–33.

14 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* 2003; **95:** 1851–60.

15 Bernabei R, Landi F, Calvani R, et al, and the SPRINTT consortium. Multicomponent intervention to prevent mobility disability in frail older adults: randomised controlled trial (SPRINTT project). *BMJ* 2022; **377:** e068788.

16 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, for the STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; **147:** 573–77.

17 Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994; **49:** M85–94.

18 Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 1963; **185:** 914–19.

19 Vellas B, Guigoz Y, Garry PJ, et al. The Mini Nutritional Assessment (MNA) and its use in grading the nutritional state of elderly patients. *Nutrition* 1999; **15:** 116–22.

20 Vetrano DL, Roso-Llorach A, Fernández S, et al. Twelve-year clinical trajectories of multimorbidity in a population of older adults. *Nat Commun* 2020; **11:** 3223.

21 Landi F, Russo A, Cesari M, et al. The ilSIRENTE study: a prospective cohort study on persons aged 80 years and older living in a mountain community of Central Italy. *Aging Clin Exp Res* 2005; **17:** 486–93.

22 Liu X, Huang Y, Fu J, et al. Associations of arthritis with functional disability and depressive symptoms in general US adults: NHANES 1988-1994 and 1999-2018. *Aging Med (Milton)* 2024; **7:** 705–16.

23 Bing S, Chen Z, Wu D, et al. Evolution of sarcopenia status and risk of incident cardiovascular disease. *Eur J Prev Cardiol* 2025; published online Feb 27. https://doi.org/10.1093/eurjpc/zwaf115.

24 Maniscalco L, Veronese N, Ragusa FS, et al. Sarcopenia using muscle mass prediction model and cognitive impairment: A longitudinal analysis from the English longitudinal study on ageing. *Arch Gerontol Geriatr* 2024; **117:** 105160.

25 Tang W, Yu J, Yang K, Jia J, Mao X, Song F. Association between sarcopenia and components of metabolic syndrome among Chinese older adults: a population-based longitudinal study using CHARLS. *BMC Public Health* 2025; **25:** 2823.

26 Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. *JAMA* 2011; **305:** 50–58.

27 Suzuki T, Bean JF, Fielding RA. Muscle power of the ankle flexors predicts functional performance in community-dwelling older women. *J Am Geriatr Soc* 2001; **49:** 1161–67.

28 Garcia RE, Cawthon PM, Nicklas BJ, et al. Usual-paced 400 m long distance corridor walk estimates cardiorespiratory fitness among older adults: The Study of Muscle, Mobility and Aging. *J Am Geriatr Soc* 2024; **72:** 858–65.

29 den Ouden MEM, Schuurmans MJ, Brand JS, Arts IEMA, Mueller-Schotte S, van der Schouw YT. Physical functioning is related to both an impaired physical ability and ADL disability: a ten year follow-up study in middle-aged and older persons. *Maturitas* 2013; **74:** 89–94.

30 Kraemer WJ, Ratamess NA. Fundamentals of resistance training: progression and exercise prescription. *Med Sci Sports Exerc* 2004; **36:** 674–88.