Biomedical Engineering Letters (2025) 15:443–454
https://doi.org/10.1007/s13534-025-00461-z

ORIGINAL ARTICLE

A wearable approach for Sarcopenia diagnosis using stimulated muscle contraction signal

Jihoon Shin[1] · Kwangsub Song[1] · Sung-Woo Kim[2] · Sangui Choi[1] · Hooman Lee[1] · Il-Soo Kim[3] · Sun Im[4] · Min Seok Baek[2]

Received: 12 June 2024 / Revised: 14 January 2025 / Accepted: 17 January 2025 / Published online: 5 February 2025

Abstract

Sarcopenia is a rapidly rising health concern in the fast-aging countries, but its demanding diagnostic process is a hurdle for making timely responses and devising active strategies. To address this, our study developed and evaluated a novel sarcopenia diagnosis system using Stimulated Muscle Contraction Signals (SMCS), aiming to facilitate rapid and accessible diagnosis in community settings. We recruited 199 adults from Wonju Severance Christian Hospital between July 2022 and October 2023. SMCS data were collected using surface electromyography sensors with the wearable device exoPill. Their skeletal muscle mass index, handgrip strength, and gait speed were also measured as the reference. Binary classification models were trained to classify each criterion for diagnosing sarcopenia based on the AWGS cutoffs. The binary classification models achieved high discriminative abilities with an AUC score near 0.9 in each criterion. When combining these criteria evaluations, the proposed sarcopenia diagnosis system performance achieved an accuracy of 89.4% in males and 92.4% in females, sensitivities of 81.3% and 87.5%, and specificities of 91.0% and 93.8%, respectively. This system significantly enhances sarcopenia diagnostics by providing a quick, reliable, and non-invasive method, suitable for broad community use. The promising result indicates that SMCS contains extensive information about the neuromuscular system, which could be crucial for understanding and managing muscle health more effectively. The potential of SMCS in remote patient care and personal health management is significant, opening new avenues for non-invasive health monitoring and proactive management of sarcopenia and potentially other neuromuscular disorders.

Keywords Sarcopenia · Neuromuscular system · Deep learning · Electrical stimulation · Surface electromyography

Abbreviations

SMI	Skeletal muscle mass index
HG	Handgrip strength
GAIT	Gait speed
MLP	Multi-layer perceptron
SVM	Support vector machine
RF	Random forest

✉ Min Seok Baek
minbaek@yonsei.ac.kr

[1] EXOSYSTEMS AI Research, Seongnam, Gyeonggi-do 13449, Republic of Korea

[2] Department of Neurology, Wonju Severance Christian Hospital, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju, Gangwon-do 26426, Republic of Korea

[3] Biomechanics Research and Development Center, RHIN Rehabilitation Hospital, Yongin, Gyeonggi-do 16864, Republic of Korea

[4] Department of Rehabilitation Medicine, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

1 Introduction

Sarcopenia is a gradual musculoskeletal disease commonly occurring with aging, often coexists with various chronic conditions such as chronic obstructive pulmonary disease and cardiovascular disease [1–4]. The significance of sarcopenia diagnosis is closely associated with morbidity, mortality, and healthcare expenditure. Although consensus is still evolving, according to recent sarcopenia diagnostic guidelines, individuals presenting with "*low muscle mass accompanied by low muscle strength or low physical performance*" are diagnosed with sarcopenia [2, 5, 6]. In clinical practice, muscle mass is typically evaluated using

height-adjusted appendicular skeletal muscle mass (also known as Skeletal Muscle mass Index; SMI) measured through techniques like Dual-energy X-ray Absorptiometry (DXA) [7] or Bioelectrical Impedance Analysis (BIA) [8, 9], muscle strength is assessed through handgrip strength [10], and physical performance is quantitatively evaluated using tests such as the Short Physical Performance Battery (SPPB) [11, 12], usual gait speed, 6-minute walk test, or Timed Up and Go (TUG) test [2, 13]. However, a survey conducted in Australia and New Zealand revealed that in reality, the accessibility to such evaluation methods is often limited due to the restricted availability of expensive equipment [14, 15]. Another drawback can be people who cannot properly perform the tasks required for assessment because of immobilization, life-threatening illnesses or consciousness problems from sedative medications do not even get a chance to be diagnosed [16–18].

There are many studies to diagnose sarcopenia using various methods. A cross-sectional study that applied machine learning algorithms to 11 selected features, including handgrip strength and gait speed, demonstrated an accuracy exceeding 0.9 with a sample of 160 participants [19]. A study utilizing machine learning models to analyze Electronic Health Records data from 1304 participants identified key predictors of sarcopenia and achieved a high prediction accuracy, with logistic regression and support vector machine models reaching an Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) of 0.91 and 0.99 respectively [20]. Another study developed sex-specific sarcopenia identification tools for maintenance hemodialysis patients using machine learning, where the voting classifiers achieved an AUC of 0.87 for males and 0.78 for females. Other studies utilized computed tomography images [21, 22], chest X-ray [23], or ultrasound [24]. However, these methods still require skilled personnel for data collection, which can limit accessibility. While simpler diagnostic approaches like the SARC-F questionnaire [25–27] exist, they often show low sensitivity [28, 29], indicating the need for more reliable and innovative methods to accurately diagnose sarcopenia.

Sarcopenia is strongly associated with both muscle function and structure, as evidenced by diagnostic criteria that emphasize muscle mass, strength, and physical performance. This age-related sarcopenia leads to decrease the myelin's cross-sectional area (CSA) of motor neurons and muscle mass of muscle fiber type II [30–32], and this results in a decline in physical function [30, 33]. Therefore, analyzing motor neurons and muscle fibers is essential for diagnosing sarcopenia. The motor unit, as the fundamental unit responsible for muscle contraction, is highly correlated with muscle function [34–36]. Given the importance of muscle function in sarcopenia, we believe that electromyography (EMG), which captures motor unit action potentials, is well-suited for evaluating muscle functionality. Voluntary surface EMG (sEMG) provide valuable information on motor unit action potentials, including firing rate, recruitment pattern, latency, and amplitude. Studies have shown significant correlations between sEMG features obtained during maximal voluntary contractions (MVC) and muscle mass [37, 38], muscle strength [38], and as well as physical performance [39], highlighting sEMG's potential in assessing these aspects of muscle health. Furthermore, recent findings suggest association between sEMG and sarcopenia [40–43]. However, traditional MVC-based sEMG methods can be challenging to apply clinically due to the need for skilled personnel for data collection, active voluntary contraction, and precise testing conditions, which can be uncomfortable or difficult for certain populations, including those with sarcopenia [44, 45]. To address this challenge in accessibility, this study leverages Stimulated Muscle Contraction Signal (SMCS), a type of sEMG signal measured during electrical stimulation [46]. Specifically, the electrical stimulation initiates action potential, generating involuntary muscle contractions and eliminating the need for individuals to volitionally exert force. Previous research has developed digital biomarkers related to muscle strength and endurance using SMCS, which reflects the state of motor neurons and is expected to include information related to physical performance [46]. Thus, this paper attempts to evaluate the muscle mass, muscle strength, and physical performance of subjects through SMCS, which encompasses bio-information of the neuromuscular system as a whole.

This study presents two main contributions. Firstly, it simplifies and automates the diagnosis of sarcopenia using a wearable device that measures SMCS signals. This system allows for measurements to be taken without the need for expert assistance, making it feasible for future applications in home monitoring and other settings. Secondly, instead of directly diagnosing sarcopenia, the system predicts the 'low' states of muscle mass, muscle strength, and physical performance individually, in alignment with conventional clinical practice. These predictions are then combined to diagnose sarcopenia. This approach acknowledges that even if two individuals are diagnosed with sarcopenia, the underlying causes and necessary treatments may differ, and the proposed diagnostic system is designed to reflect this complexity.

2 Materials and methods

2.1 Diagnosis criteria for Sarcopenia

This study adheres to the diagnostic criteria outlined by the Asian Working Group for Sarcopenia (AWGS) [2]. According to AWGS guidelines, sarcopenia is diagnosed based on low muscle mass accompanied by either low muscle strength or low physical performance. To quantify each diagnostic component, SMI, handgrip strength, and usual gait speed were measured (eFig. 1 in Online Resource 1). SMI was measured through multifrequency BIA using the InBody 770 (InBody, Seoul, South Korea) [47]. An SMI less than 7.0 kg/m² for males and less than 5.7 kg/m² for females was considered 'low'. Handgrip strength was assessed using a digital handgrip dynamometer (T.K.K.5401; Takei Scientific Instruments Co., Ltd., Tokyo, Japan) [48]. The participants were asked to stand upright with their shoulders in a neutral position, arms at the sides with fully extended elbows. The handgrip strength at maximal power for both sides were averaged. A handgrip strength less than 28.0 kgf for males and less than 18.0 kgf for females was considered 'low'. Gait speed was measured on the GAITRite Electronic Walkway (CIR Systems Inc., Peekskill, NY), a six-meter-long and 0.6-meter-wide walkway with pressure-activated sensors [49]. Participants were asked to walk on the GAITRite Electronic Walkway ten times barefoot at their usual walking speed. A gait speed lower than 100 cm/s was considered 'low' for both males and females. In the rest of the text, we denote the binary assessment of whether each diagnostic component is 'low' as $\mathrm{SMI}_{\mathrm{low}}$, $\mathrm{HG}_{\mathrm{low}}$, $\mathrm{GAIT}_{\mathrm{low}}$, respectively. In other words, sarcopenia is expressed as

$$\mathrm{SMI}_{\mathrm{low}} \wedge (\mathrm{HG}_{\mathrm{low}} \vee \mathrm{GAIT}_{\mathrm{low}}), \tag{1}$$

where $\vee$ denotes "logical or" and $\wedge$ denotes "logical and".

2.2 Details of population demographics

This study recruited participants at the outpatient clinic of the Department of Neurology at Wonju Severance Christian Hospital between July 2022 and October 2023. Patients who were 40 years and older, with no difficulty in walking, and were willing to participate in the study were recruited. Exclusion criteria included a previous diagnosis of idiopathic Parkinson's disease, atypical parkinsonism, normal pressure hydrocephalus, severe cardiopulmonary diseases, cancer, osteoarthritis, or orthopedic diseases. A total of 199 patients, comprising 94 males and 105 females, were enrolled. Participants were categorized into two groups: the sarcopenia group ($n=40$) and the healthy control group ($n=159$), following the AWGS guidelines. Before the collection of data, all participants provided written informed consent. This study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of the Yonsei University Wonju Severance Hospital (Ref# 2022-0234-001).

2.3 SMCS data collection protocol

Figure 1 illustrates the exoPill, a wearable device developed by EXOSYSTEMS (Seongnam, Gyeonggi-do, Republic of Korea), capable of delivering electrical stimulation and simultaneously measuring sEMG. To facilitate the collection of SMCS, hydrogel electrodes were positioned over each of the left and right quadriceps femoris muscles, at mid-point of the rectus femoris muscle. Participants were instructed to sit upright in a chair with both feet flat on the floor with knees set at a 90-degree angle. They were also instructed to relax their muscles during the measurement as shown in eFig. 2 in Online Resource 1. In accordance with the protocol established in previous research [46], we adapted and modified the electrical stimulation parameters to enhance the analytical feature space. To be specific, the frequency range of electrical stimulation was extended from 10 Hz to 30 Hz, as previously used, to 5 Hz to 30 Hz in this study, aiming to capture a broader spectrum of muscle response data. We also explored a range of electrical stimulation intensities from 0.5 mA to 44.7 mA and determined that an intensity of 18.0 mA produced visible muscle twitches without causing discomfort to participants; thus, this setting was selected. Each electrical stimulation phase was applied for 8 s, followed by a 2-second rest period, resulting in total measurement time lasting 60 s.

2.4 Data collection hardware specification

The wearable device, exoPill, comprises one cradle and four modules. The cradle recharges the batteries in the modules and provides information on the availability of the modules. Each module is equipped with an EMG sensor and an integrated electrical stimulator. The modules are responsible for emitting electrical stimulation signals, which are generated based on stimulation parameters received from the user application. Each module contains custom-built circuits, four electrodes, a bi-colored light-emitting diode (LED), a battery, and charging pins. The entire circuitry is enclosed within a polycarbonate case with dimensions of $50 \times 59 \times 23$ mm. Four magnetic electrodes, which couple to a hydrogel electrodes, are positioned beneath the module, slightly protruding from the case. One module weighs approximately 40 g with its charging pins exposed on the back and has one microcontroller in it.

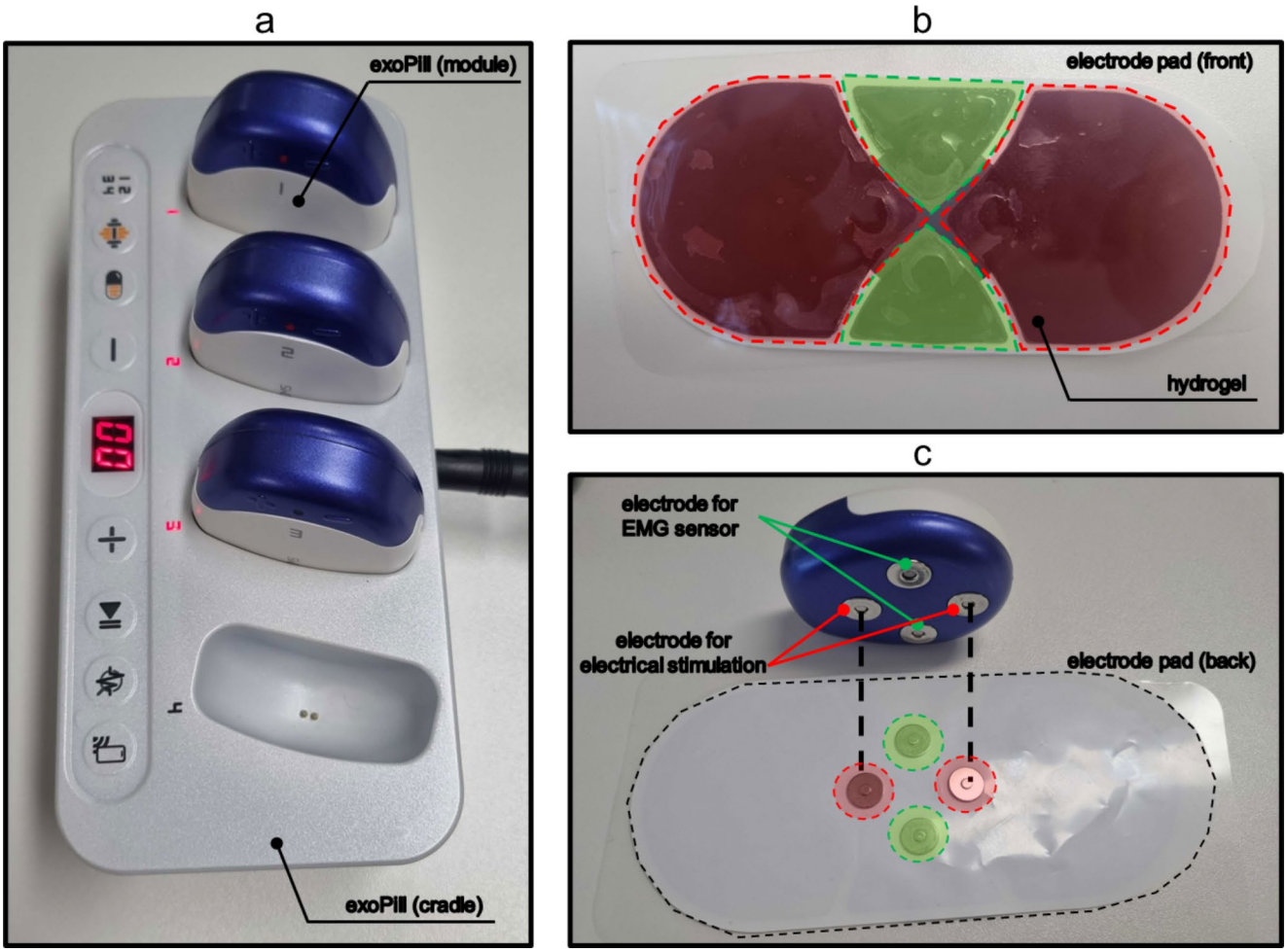

Fig. 1 Experimental equipment for collecting SMCS: **a** wearable device (exoPill) for the electrical stimulation and **b** hydrogel electrodes, **c** magnetically coupled

The microcontroller (STM32L451CE, STMicroelectronics) in the modules communicates with a Bluetooth low-energy module (NRF52840, Nordic Semiconductor) through UART communication at a rate of 115,200 bps. This communication allows the module to receive commands from the user application, which is connected via Bluetooth 5.0 communication and operates the circuitry of each module. The circuitry can generate a pulse wave with a frequency ranging from 5 Hz to 100 Hz and a maximum peak-to-peak amplitude of 100 V (measured at a non-inductive resistor of 500 ohms). The EMG recording system with band-pass filtering from 17 Hz to 450 Hz and amplification with a gain of 1,100 were used to detect the sEMG. The microcontroller then measures the filtered signal at a rate of 1,000 samples per second through 12-bit analog-to-digital conversion, of which only the upper 8 bits are utilized. This approach helps to minimize the impact of minor EMG fluctuations due to unwanted voluntary contractions, allowing the focus to remain on the EMG signals generated by involuntary muscle contractions in response to electrical stimulation (SMCS). Details of hardware specifications of EMG sensors are provided in Table 1. Healthy individual and sarcopenia patient SMCS samples are provided in eFig. 3 in Online Resource 1. After processing, the signal gets transmitted to the user application. The device is powered by a lithium-ion polymer battery with a voltage of 3.7 V and a capacity of 250 mAh. The battery is charged at a DC voltage of 4.2 V with a charging current of 450 mA.

Table 1 Hardware specification of EMG sensor in exoPill module

Hardware specification	
Input impedance	10GΩ
Input referred noise level	0.1μV(1 kHz)
Notch filter	50, 60 Hz
EMG filter	17 Hz~450 Hz
Common mode rejection ratio(dB)	86dB
Sampling rate	1000 Hz
Sensitivity (resolution)	12bit

The hydrogel electrodes for collecting SMCS were StiMus Electrode developed by HUREV Corp (Wonju,

Gangwon-do, Republic of Korea). Each electrode features a hydrogel contact surface component, measured 145 mm wide and 65 mm long, exhibiting a resistance of 50 Ω per 20 mm.

2.5 Proposed Sarcopenia diagnosis system

Figure 2 presents an overview of the proposed sarcopenia diagnosis system, which encompasses three main phases: data preprocessing, feature extraction, and binary classification model training. During the data preprocessing phase (Fig. 2a), SMCS data is segmented by each electrical stimulation frequency ranging from 5 Hz to 30 Hz. The raw sEMG signals captured by the wearable device were first amplified 1100 times using an amplifier configured with a 3.3 V single power supply, and a bias was applied before feeding the signal into the analog-to-digital converter. This step enabled the capture of motor unit responses with a potential difference of -1.5 mV to 1.5 mV. In this study, the 8-bit digitized data, with a resolution of 12 μV, was further normalized to a range of -1 to 1 for analysis, ensuring the baseline set to 0. This standardization ensures that subsequent analyses are conducted on data with consistent baselines and amplitudes, which is crucial for reliable feature extraction.

In the feature extraction phase (Fig. 2b), Muscle Contraction Pattern (MCP) features are systematically extracted from each segmented SMCS [46, 50]. This involves converting SMCS data into spectrograms using the Short Time Fourier Transform (STFT), with a window size of 64. Each electrical stimulation phase lasts for 8 s, and with a sampling rate of 1000 samples/sec, this results in a segment length of 8000 samples. Applying the STFT with the aforementioned parameters yields a spectrogram of size 32×125, composed of 32 frequency bins. To capture the essential dynamics of muscle activity, two types of envelope signals are computed for each frequency bin by summing the positive peaks with their adjacent negative peaks and interpolating the gaps [50]. A bias elimination method is then applied to one type of the envelope signals to enhance the accuracy of the features. Autocorrelation is performed on these envelopes to derive the final MCP features, which effectively capture the repetitive patterns of muscle activation and are crucial for accurate sarcopenia diagnosis. Since this peak-based computation can shorten the envelope signal's length

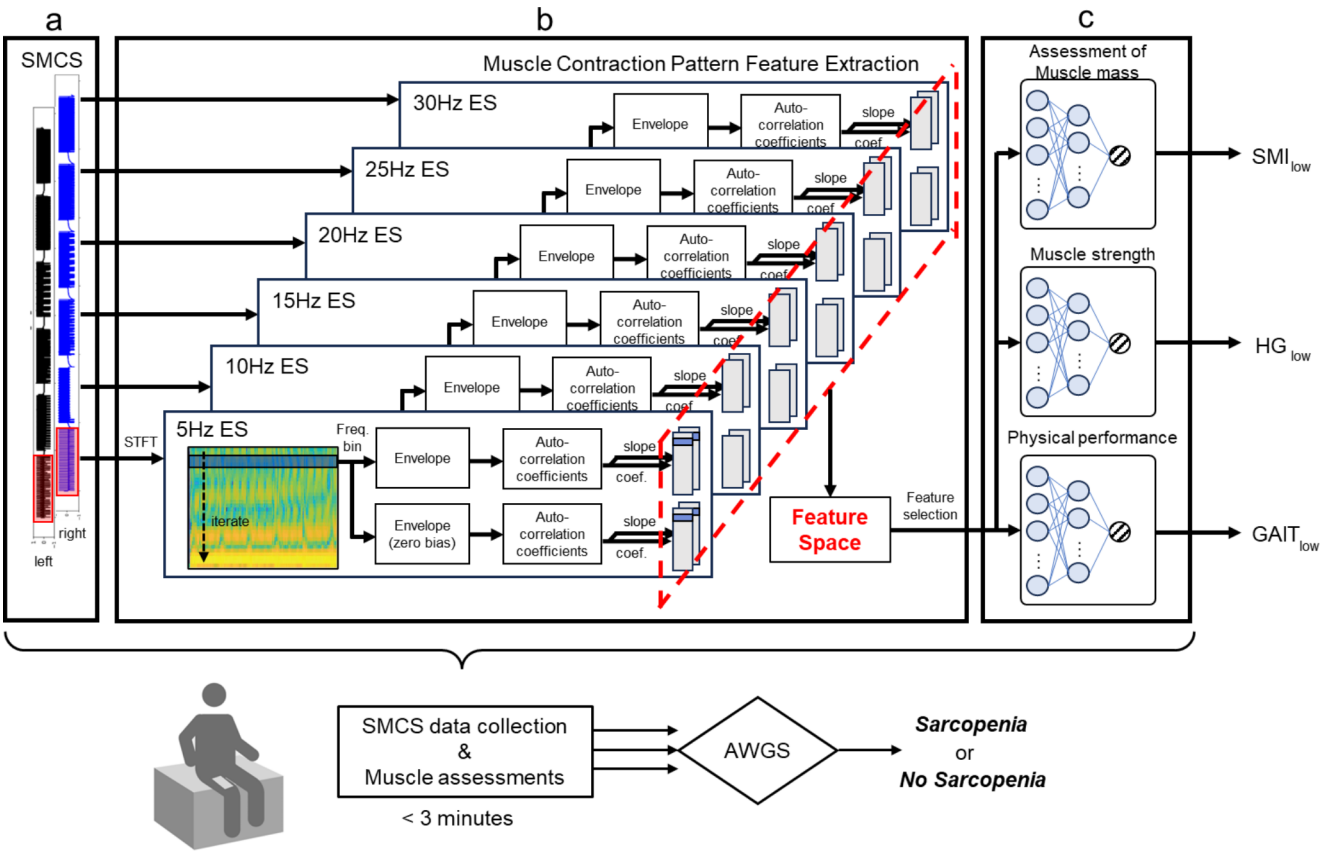

Fig. 2 Overview of the proposed sarcopenia diagnosis system. **a** collected SMCSs are segmented by each electrical stimulation period, **b** Muscle Contraction Pattern (MCP) features are extracted. **c** Features are further selected to use as input for the binary classification model, each target criterion. *Abbreviation*: ES, electrical stimulation; SMCS, Stimulated Muscle Contraction Signal; STFT, Short-Time Fourier Transform; coef., coefficients

to varying degrees, we selected a fixed length of 100 to ensure consistent operation across all collected data. Consequently, a 32×100 autocorrelation features and 32×99 slope features are created from each envelope signal. Given that there are two types of envelope for each electrical stimulation segment, combining across all 6 stimulation segments results in a total of 76,416 MCP features.

The model training phase (Fig. 2c) involves selecting features that demonstrate high relevance to the target, indicated by an AUC greater than 0.7. To reduce redundancy and enhance the model's predictive power, the Minimum Redundancy Maximum Relevance (mRMR) algorithm is applied to these selected features [51]. A grid search is conducted to determine the optimal number of selected features, which was set at 64 from options of 32, 64, and 128 [52]. The refined features serve as inputs for three different models: a Multi-Layer Perceptron (MLP), a Support Vector Machine (SVM), and a Random Forest (RF).

To rigorously evaluate the performance of each classification model for $\mathrm{SMI_{low}}$, $\mathrm{HG_{low}}$, and $\mathrm{GAIT_{low}}$, a nested 3×2 cross-validation method is implemented [53, 54]. This method ensures that each fold of the SMCS dataset is utilized as training, validation, and test sets in rotation, promoting an even distribution across references. The models are trained using the training set, with the validation set employed for early stopping, and the test set providing the final verification. Performance metrics such as the AUC, accuracy, precision, sensitivity, specificity, and F1 score are calculated to assess the efficacy of the models across the different sarcopenia criteria. The final system's sarcopenia diagnosis performance is evaluated using the outputs of these models, analyzing accuracy, precision, sensitivity, specificity, and F1 score.

Each MLP model, trained separately for each gender with specific diagnostic cutoffs for SMI and HG, comprises two hidden layers with 256 and 32 neurons respectively, and utilizes a dropout ratio of 0.5 to prevent overfitting [55]. The MLP models were trained using the Adam optimizer with a learning rate of 0.001, a batch size of 512, and up to 1500 epochs. Early stopping with a patience of 200 epochs was applied. The binary cross-entropy loss function was used for training, as it is well-suited for binary classification tasks. The activation function used in all hidden layers was Exponential Linear Unit (ELU).

For the SVM and RF models, the optimal parameter setting was selected based on those that empirically provided the best performance in our experiments. For each fold in the inner loop, SVM hyperparameters were selected from the following ranges: {0.1, 0.2, 0.5, 1.0} for regularization parameters, $\{1/(N_{feat} \cdot \sigma_X^2), 1/N_{feat}\}$, for kernel coefficient, and {'linear', 'poly', 'rbf'} for kernel type. N_{feat} denotes number of selected features, and σ_X^2 is the variance of selected features. For 'poly' kernels, the polynomial degree was chosen from {1, 2, 3}. For RF, hyperparameters were selected from: maximum depth {None, 10, 20, 50, 100}, minimum samples per leaf {1, 2, 4, 8}, minimum samples per split {2, 5, 10, 25}, and number of estimators {5, 10, 25, 50, 100, 200}. The selected optimal hyperparameters for each model are detailed in eTable 1 and eTable 2 in Online Resource 1.

3 Results

The following section summarizes the experimental results for the proposed methodology. We provide the classification results for the models targeting $\mathrm{SMI_{low}}$, $\mathrm{HG_{low}}$, and $\mathrm{GAIT_{low}}$. All computations were performed using Python 3.9.12 and TensorFlow 2.10.0.

3.1 Participants characteristics

Table 2 summarizes the demographic and sarcopenia-related characteristics of our study population, comprising 199 Korean participants. The table provides detailed comparisons across both genders within the healthy control group and sarcopenia group. The average age and weight of male participants were 75.0 years and 65.1 kg, respectively, while female participants averaged 74.5 years and 55.8 kg. Sarcopenia was identified in 40 participants according to the AWGS guidelines. Among these, the sarcopenia group included 16 males and 24 females. While there was little difference in mean age between genders within the sarcopenia group, female participants were slightly younger on average in the healthy control group. Overall, participants with sarcopenia were significantly older than those in the healthy control group, with statistical significance noted ($p < 0.001$ for females and $p < 0.05$ for males). In terms of height, female participants in the sarcopenia group were slightly shorter compared to those in the control group, a difference that reached statistical significance ($p < 0.05$). No significant height difference was observed among male participants.

3.2 Classification performance for $\mathrm{SMI_{low}}$, $\mathrm{HG_{low}}$, and $\mathrm{GAIT_{low}}$ criteria

Table 3 summarizes the performance metrics of each trained classification model, including MLP, SVM, and RF, compared to the ground-truth reference of diagnostic components. This comprehensive evaluation demonstrates the system capability to effectively classify $\mathrm{SMI_{low}}$, $\mathrm{HG_{low}}$, and $\mathrm{GAIT_{low}}$. Both the MLP and SVM models demonstrated similar performance, while the RF model shows comparatively lower performance. The experimental results

Table 2 Demographic and sarcopenia-related characteristics of the study population

	Male			Female		
	Total	Healthy control	Sarcopenia	Total	Healthy control	Sarcopenia
N	94	78	16	105	81	24
Age (years)	75.0±8.9	74.1±8.3	79.2±10.9[a]	74.5±7.4	73.2±7.2	79.0±6.2[d]
Height (cm)	165.6±5.8	166.1±5.9	163.1±4.8	153.3±5.1	154.0±4.9[b]	151.2±5.2[cd]
Body Weight (kg)	65.1±8.1	66.4±7.9	58.7±6.2[a]	55.8±7.8	58.2±6.6[b]	47.5±5.4[cd]
SMI (kg/m^2)	7.4±0.7	7.6±0.5	6.5±0.4[a]	5.9±0.7	6.2±0.4[b]	5.0±0.5[cd]
Handgrip strength (kgf)	29.1±7.6	29.1±7.2	21.6±6.3[a]	17.2±4.4	18.0±4.1[b]	14.3±4.5[cd]
Gait speed (cm/s)	82.6±23.4	87.1±21.9	60.8±17.8[a]	77.7±22.3	80.4±22.6	68.5±18.9[d]
Calf circumference (cm)	34.4±2.7	34.9±2.6	32.1±1.9[a]	32.7±2.9	33.7±2.2[b]	29.2±2.1[cd]
SARC-F score	1.6±1.8	1.2±1.5	3.3±2.2[a]	2.5±2.5	2.4±2.6[b]	2.8±2.3

Data are given as the mean and standard deviation. [a]$p < 0.05$ for the comparisons between male healthy control group and male sarcopenia group. [b]$p < 0.05$ for the comparisons between male healthy control group and female healthy control group. [c]$p < 0.05$ for the comparisons between male sarcopenia group and female sarcopenia group. [d]$p < 0.05$ for the comparisons between female healthy control group and female sarcopenia group.

Abbreviations: SMI, skeletal muscle mass index; SARC-F, Strength, assistance with walking, rising from a chair, climbing stairs, and falls questionnaire.

Table 3 The performance metric of the trained models in predicting three diagnostic components

		Model type	Accuracy	Sensitivity	Specificity	Precision	F1-score
Male	SMI_{low}	MLP	90.4%	79.0%	93.3%	75.0%	76.9%
		SVM	91.5%	84.2%	93.3%	76.2%	**80.0%**
		RF	87.2%	68.4%	92.0%	68.4%	68.4%
	HG_{low}	MLP	96.8%	93.8%	100.0%	100.0%	96.8%
		SVM	97.9%	95.8%	100.0%	100.0%	**97.9%**
		RF	96.8%	93.8%	100.0%	100.0%	96.8%
	$GAIT_{low}$	MLP	90.4%	97.2%	69.6%	90.8%	**93.9%**
		SVM	88.3%	91.5%	78.3%	92.9%	92.2%
		RF	90.4%	94.4%	78.3%	93.1%	93.7%
Female	SMI_{low}	MLP	92.4%	88.5%	93.7%	82.1%	85.2%
		SVM	92.4%	96.2%	91.1%	78.1%	**86.2%**
		RF	88.6%	73.1%	93.7%	79.2%	76.0%
	HG_{low}	MLP	93.3%	94.6%	92.0%	92.9%	93.7%
		SVM	96.2%	96.4%	96.0%	96.4%	**96.4%**
		RF	91.4%	90.9%	92.0%	92.6%	91.7%
	$GAIT_{low}$	MLP	87.6%	92.0%	66.7%	93.0%	**92.5%**
		SVM	83.8%	85.1%	77.8%	94.9%	89.7%
		RF	81.0%	89.7%	38.9%	87.6%	88.6%

Abbreviations: SMI, skeletal muscle mass index; HG, handgrip strength; GAIT, gait speed; MLP, multi-layer perceptron; SVM, support vector machine; RF, random forest.

showed that strong classification performance was achieved across various models, regardless of the specific model used. Unless otherwise specified, subsequent mentions of models and performance metrics refer to the MLP model for simplicity.

In males, the models exhibited strong performance across all metrics. For SMI_{low}, the accuracy was notably high at 90.4%, with a specificity of 93.3%, reflecting the model's ability to correctly identify normal cases. The precision was somewhat lower at 75.0%, suggesting some limitations in identifying low SMI cases without false positives. The HG_{low} model showed exceptional results, with an accuracy of 96.8% and a perfect specificity and precision of 100.0%, indicating its high reliability. $GAIT_{low}$ model also showed good sensitivity at 97.2%, although the specificity was lower at 69.6%, which may indicate challenges in distinguishing $GAIT_{low}$ accurately. In females, the system maintained high diagnostic standards as well, achieving an accuracy of 92.4% and a sensitivity of 88.5% for SMI_{low}. The HG_{low} model again performed excellently, mirroring the high precision seen in males. $GAIT_{low}$ model in females had an accuracy of 87.6% with a high sensitivity of 92.0%, similar to males, but with slightly lower specificity at 66.7%.

Figure 3 presents the Receiver Operating Characteristic (ROC) curves for the MLP classification models developed

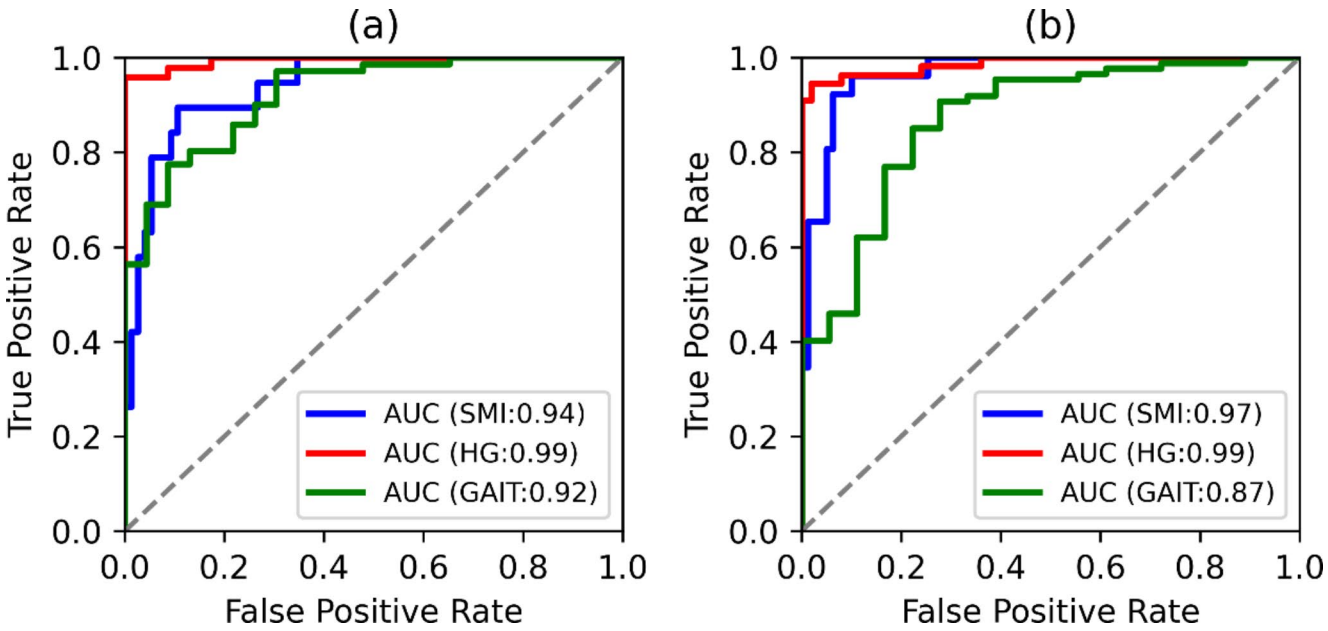

Fig. 3 Receiver Operating Characteristic (ROC) curves for MLP classification models: **a** in male groups and **b** female groups. *Abbreviation*: SMI, skeletal muscle mass index; HG, handgrip strength; GAIT, gait speed

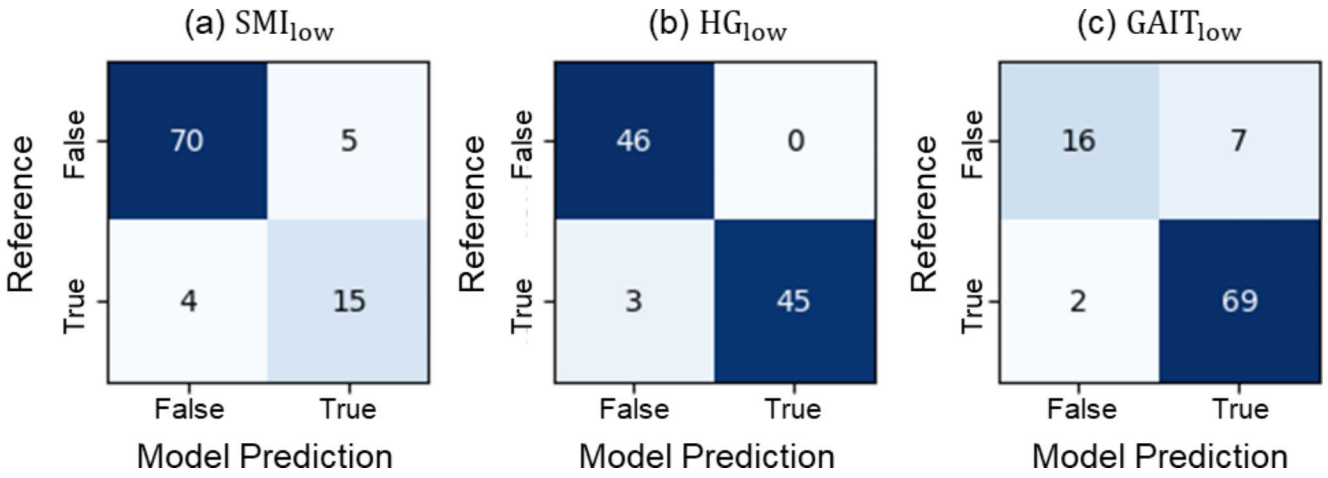

Fig. 4 Confusion matrix of the trained male classification MLP models in each criterion: **a** $\mathbf{SMI_{low}}$, **b** $\mathbf{HG_{low}}$, **c** $\mathbf{GAIT_{low}}$

to assess SMI, HG, and GAIT across genders. The ROC curves demonstrate the model ability to accurately classify these measures as either high or low according to the AWGS cutoffs. For males, the models achieved AUC scores of 0.94, 0.99, and 0.92 for SMI, HG, and GAIT, respectively, highlighting their strong discriminative capabilities, especially for HG which approaches perfect classification with an AUC of 0.99. The models for females recorded AUC scores of 0.97, 0.99, and 0.87 for SMI, HG, and GAIT, respectively, indicating high efficacy in classification accuracy.

Figures 4 and 5 display the confusion matrices for the MLP models assessing SMI_{low}, HG_{low}, and $GAIT_{low}$, using the AWGS sarcopenia cutoffs as reference for males and females, respectively. For SMI, the majority of participants were classified as normal based on the AWGS cutoff. Despite the higher prevalence of normal cases in the participants, the models effectively distinguished SMI_{low} cases, demonstrating their robustness. For instance, in the male group, the model correctly identified 15 out of 19 SMI_{low} subjects. In contrast, the gait speed classification faced a different challenge, as there was a higher proportion of $GAIT_{low}$ cases in the participants. For females, the GAIT model correctly predicted 12 normal and 80 low conditions, demonstrating high sensitivity in a scenario where low cases were more prevalent.

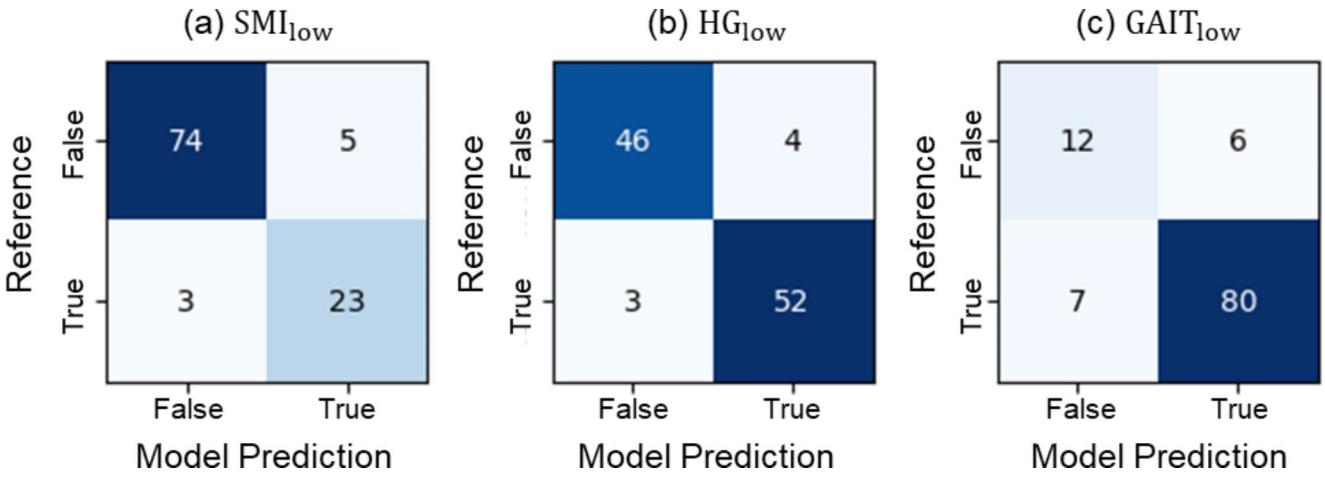

Fig. 5 Confusion matrix of the trained female classification MLP models in each criterion: **a** $\mathbf{SMI_{low}}$, **b** $\mathbf{HG_{low}}$, **c** $\mathbf{GAIT_{low}}$

Table 4 The performance metric of the trained models in diagnosing Sarcopenia

	Sex	Model type	Accuracy	Sensitivity	Specificity	Precision	F1-score
Sarcopenia[a]	Male	MLP	89.4%	81.3%	91.0%	65.0%	72.2%
		SVM	88.3%	75.0%	91.0%	63.2%	68.6%
		RF	87.2%	62.5%	92.3%	62.5%	62.5%
	Female	MLP	92.4%	87.5%	93.8%	80.8%	84.0%
		SVM	94.3%	95.8%	93.8%	82.1%	88.5%
		RF	88.6%	70.8%	93.8%	77.3%	73.9%

[a]Sarcopenia diagnosis performance metric is evaluated by combining three criteria evaluations of models, $\mathrm{SMI_{low}}$, $\mathrm{HG_{low}}$, $\mathrm{GAIT_{low}}$, according to AWGS.

3.3 Overall performance of the developed sarcopenia diagnosis system

The overall diagnostic efficacy of our developed sarcopenia diagnosis system is summarized in Table 4, which presents the accuracy, sensitivity, specificity, precision, and F1-score for each criterion across genders, and Fig. 6, which presents confusion matrices for detecting sarcopenia. When evaluating sarcopenia with these SMCS-based trained MLP models, the proposed system achieved an accuracy of over 89% in both genders, with males at 89.4% and females at 92.4%. The sensitivity was higher in females (87.5%) compared to males (81.3%), showing a slight gender discrepancy in detecting sarcopenia conditions.

4 Discussion

This study developed sarcopenia diagnosis system using deep learning and SMCS. Our approach introduces SMCS obtained via a wearable device, allowing for a simpler, consistent, and non-voluntary assessment of muscle response to electrical stimulation. By capturing the motor unit's response to a controlled stimulus, SMCS can offer insights into muscle function that are relevant for sarcopenia. Features related to muscle mass, muscle strength, and physical performance were extracted from 199 community-dwelling participants and used for training models. The developed sarcopenia diagnosis system achieved a fair performance with an accuracy of 89.4% among males, and that of 92.4% among females. With the aid of a specialized app, the wearable device autonomously conducts electrical stimulation, collects SMCS data, and diagnoses sarcopenia, facilitating the assessment process without requiring expert intervention. Furthermore, assessment using SMCS takes about one to two minutes at a time for each leg, so the 20 min of time required for conventional assessment can be reduced to about three minutes. It has been reported that diagnosing sarcopenia in critically ill patients is rather challenging since the patients cannot carry out tasks related to the assessment and whether their assessment results represent the muscle health remains in question [16, 17]. SMCS can serve as a fair alternative when the physical tasks required for the assessment are not allowed, or in situations where use of dynamometer, access to DXA or alternative muscle mass measurement techniques is limited. To sum up, diagnosing sarcopenia using the SMCS breaks away from the conventional way of having patients come to a medical facility and go through a complicated measuring process.

The approach presented in this study offers a more 'explainable' method for diagnosing sarcopenia by

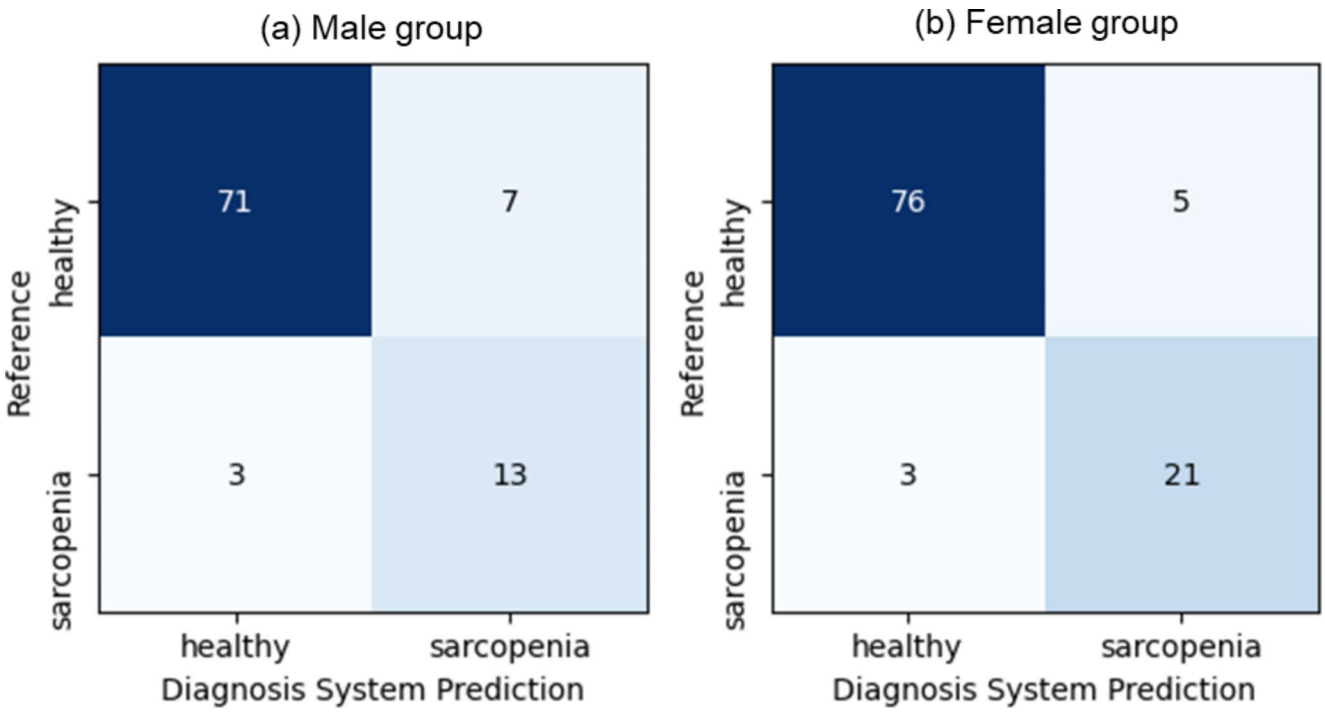

Fig. 6 Confusion matrix in sarcopenia diagnosis using developed diagnosis system with MLP models: **a** male group and **b** female group

evaluating three diagnostic components and integrating these assessments to determine the presence of sarcopenia. Unlike many existing machine learning-based sarcopenia diagnostics, which often deliver diagnostic results without showing the intermediary steps, our method provides clear insights into which specific indicators contribute to the diagnosis. This transparency is crucial because it not only helps in understanding the underlying causes of sarcopenia in individual patients but also aids healthcare professionals in tailoring specific interventions based on the deficient indicator. To the best of our knowledge, this is the first study to develop system to assess all three indicators simultaneously as a standalone diagnostic tool. Making the diagnosis more accessible, the proposed approach of diagnosing sarcopenia will help make timely medical interventions via early diagnosis and improve the prognosis in turn.

Due to the nature of the study and availability of data, participants were limited to Korean ethnicity for now, extending this work with other ethnicities remains a future task. Despite having a patient sample relatively small, we conducted thoughtfully tests using cross-validation to ameliorate this problem. Additionally, in this study, SMCS was measured only once per participant, which does not allow for reliability analysis such as test-retest. In future studies, we plan to revise the SMCS collection protocol to measure SMCS at least twice per participant to provide data on its reliability. This study's experiment does not include test-retest results; however, the reliability of the SMCS-based digital muscle marker technique was sufficiently verified in a separate ongoing clinical study [56]. SMCS was measured in a seated posture with both feet flat on the floor and knees bent at a 90-degree angle to minimize variability related to voluntary contraction and ensure participant comfort. However, different body postures and knee angles, or variations in body weight may influence muscle responses to electrical stimulation, potentially affecting SMCS measurements. In future research, we plan to further address variations in body weight by evaluating their impact on SMCS. Additionally, we will consider assessing SMCS in various resting postures, such as a reclined position, to determine the impact of posture on SMCS reproducibility and reliability. Our system classified each criterion to binary ($\mathrm{SMI_{low}}$, $\mathrm{HG_{low}}$, $\mathrm{GAIT_{low}}$), which may not be sufficient to capture the gradation of severity. Therefore, we consider extending this work adopting a regression model that could provide a continuous output, instead of binary, offering a more detailed assessment of disease severity. All these improvements would enable more precise monitoring and management of sarcopenia progression, tailoring interventions more effectively to the patient condition.

In conclusion, this study developed a sarcopenia diagnosis system to reduce the burden on patient diagnosis without going through multiple assessments at medical facilities. A total of six classification models (three models per sex) were developed from SMCS data, demonstrating the capability of SMCS to facilitate various muscle assessments. Of these models, two were for muscle mass (i.e., $\mathrm{SMI_{low}}$), two were for muscle strength (i.e., $\mathrm{HG_{low}}$), and two were for physical

performance (i.e., $GAIT_{low}$). Sarcopenia diagnosis made using these models' outputs and AWGS achieved an accuracy of 89.36% in males and that of 92.38% in females. The proposed approach can be used as a standalone diagnostic tool instead of the common approach of the dynamometer, TUG and DXA, making the diagnosis more accessible. It will make early diagnosis possible bringing timely medical intervention and treatment options can be more personalized for each person. This approach consumes less time and resources without the prerequisite of a trained medical staff or visiting a medical facility. The proposed approach has significant potential for remote patient care and personal health management.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13534-025-00461-z.

Acknowledgements This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government Ministry of Science and Information and Communication Technology (MSIT) (2022R1C1C1012535), the Technology Innovation Program(20018182) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea), and the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and Information and Communication Technology, the Ministry of Trade, Industry and Energy, the Ministry of Health Welfare, the Ministry of Food and Drug Safety) (Project Number: KMDF PR 20200901 0101).
We gratefully thank Hyun Ji Lee, Ji Young Lee, and Seo Hee Kim at Wonju Severance Christian Hospital for their technical assistance in the participants' enrollment, gait analysis, and SMCSs recording.

Author contributions All authors contributed to *Interpretation* of the study. *Conceptualization* was performed by Jihoon Shin, Kwangsub Song, Sun Im, Sangui Choi, Hooman Lee, and Min Seok Baek. *Formal analysis* was performed by Jihoon Shin, Kwangsub Song, Sung-Woo Kim, and Sangui Choi. *Participant recruitment and data curation* were performed by Min Seok Baek. *Original draft preparation* was performed by Jihoon Shin and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government Ministry of Science and ICT (2022R1C1C1012535), the Technology Innovation Program(20018182) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea), and the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health Welfare, the Ministry of Food and Drug Safety) (Project Number: KMDF PR 20200901 0101).

Declarations

Ethics approval This study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of the Yonsei University Wonju Severance Hospital (Ref# 2022-0234-001).

Consent to participate Before the collection of data, all participants provided written informed consent.

Consent to publish The authors affirm that human research participants provided informed consent for publication of the images in eFigures 1a, 1b, 1c, and 2 in online resource.

Competing interests All authors claim no conflicts of interest.

References

1. Cruz-Jentoft AJ, Sayer AA. Sarcopenia Lancet. 2019;393(10191):2636–46.
2. Chen LK, et al. Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia diagnosis and treatment. J Am Med Dir Assoc. 2020;21(3):300–e3072.
3. Ma K, et al. Pathogenesis of Sarcopenia in chronic obstructive pulmonary disease. Front Physiol. 2022;13:850964.
4. Sasaki KI, Fukumoto Y. Sarcopenia as a comorbidity of cardiovascular disease. J Cardiol. 2022;79(5):596–604.
5. Sayer AA, Cruz-Jentoft A. Sarcopenia definition, diagnosis and treatment: consensus is growing. Age Ageing, 2022. 51(10).
6. Cruz-Jentoft AJ, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing. 2019;48(4):601.
7. Mazess RB, et al. Dual-energy x-ray absorptiometry for total-body and regional bone-mineral and soft-tissue composition. Am J Clin Nutr. 1990;51(6):1106–12.
8. Kyle UG, et al. Bioelectrical impedance analysis–part I: review of principles and methods. Clin Nutr. 2004;23(5):1226–43.
9. Kyle UG, et al. Bioelectrical impedance analysis-part II: utilization in clinical practice. Clin Nutr. 2004;23(6):1430–53.
10. Bohannon RW. Muscle strength: clinical and prognostic value of hand-grip dynamometry. Curr Opin Clin Nutr Metab Care. 2015;18(5):465–70.
11. Pavasini R, et al. Short physical performance battery and all-cause mortality: systematic review and meta-analysis. BMC Med. 2016;14:1–9.
12. Fisher S, et al. Short physical performance battery in hospitalized older adults. Aging Clin Exp Res. 2009;21:445–52.
13. Podsiadlo D, Richardson S. The timed up & go: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc. 1991;39(2):142–8.
14. Yeung SSY, et al. Current knowledge and practice of Australian and New Zealand health-care professionals in Sarcopenia diagnosis and treatment: time to move forward! Australas J Ageing. 2020;39(2):e185–93.

15. Reijnierse EM, et al. Lack of knowledge and availability of diagnostic equipment could hinder the diagnosis of Sarcopenia and its management. PLoS ONE. 2017;12(10):e0185837.
16. Kizilarslanoglu MC, et al. Sarcopenia in critically ill patients. J Anesth. 2016;30(5):884–90.
17. Montero-Errasquín B, Cruz-Jentoft AJ. Acute Sarcopenia Gerontol. 2023;69(5):519–25.
18. Barry E, et al. Is the timed up and go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. BMC Geriatr. 2014;14(1):1–14.
19. Ozgur S et al. Performance Evaluation of Machine Learning Algorithms for Sarcopenia Diagnosis in Older Adults. Healthcare (Basel), 2023. 11(19).
20. Luo X, et al. Using machine learning to detect Sarcopenia from electronic health records. Digit Health. 2023;9:20552076231197098.
21. Dong X, et al. Identifying Sarcopenia in advanced non-small cell lung cancer patients using skeletal muscle CT radiomics and machine learning. Thorac Cancer. 2020;11(9):2650–9.
22. Burns JE, et al. A machine learning algorithm to Estimate Sarcopenia on abdominal CT. Acad Radiol. 2020;27(3):311–20.
23. Ryu J, et al. Chest x-ray-based opportunistic screening of Sarcopenia using deep learning. J Cachexia Sarcopenia Muscle. 2023;14(1):418–28.
24. Fu H, et al. Diagnostic test accuracy of ultrasound for Sarcopenia diagnosis: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle. 2023;14(1):57–70.
25. Malmstrom TK, Morley JE. SARC-F: a simple questionnaire to rapidly diagnose Sarcopenia. J Am Med Dir Assoc. 2013;14(8):531–2.
26. Malmstrom TK, et al. SARC-F: a symptom score to predict persons with Sarcopenia at risk for poor functional outcomes. J Cachexia Sarcopenia Muscle. 2016;7(1):28–36.
27. Woo J, Leung J, Morley JE. Validating the SARC-F: a suitable community screening tool for Sarcopenia? J Am Med Dir Assoc. 2014;15(9):630–4.
28. Kera T, et al. Limitations of SARC-F in the diagnosis of Sarcopenia in community-dwelling older adults. Arch Gerontol Geriatr. 2020;87:103959.
29. Kera T, et al. Utility of SARC-F in daycare facilities for older people. Geriatr Gerontol Int. 2022;22(10):889–93.
30. Aagaard P, et al. Role of the nervous system in Sarcopenia and muscle atrophy with aging: strength training as a countermeasure. Scand J Med Sci Sports. 2010;20(1):49–64.
31. Nilwik R, et al. The decline in skeletal muscle mass with aging is mainly attributed to a reduction in type II muscle fiber size. Exp Gerontol. 2013;48(5):492–8.
32. Scherbakov N, Sandek A, Doehner W. Stroke-related Sarcopenia: specific characteristics. J Am Med Dir Assoc. 2015;16(4):272–6.
33. Verdijk LB, et al. Characteristics of muscle fiber type are predictive of skeletal muscle mass and strength in elderly men. J Am Geriatr Soc. 2010;58(11):2069–75.
34. Heckman CJ, Enoka RM. Motor unit. Compr Physiol. 2012;2(4):2629–82.
35. Duchateau J, Enoka RM. Human motor unit recordings: origins and insight into the integrated motor system. Brain Res. 2011;1409:42–61.
36. Manini TM, Clark BC. Dynapenia and aging: an update. J Gerontol Biol Sci Med Sci. 2012;67(1):28–40.
37. Sung JH et al. Surface electromyography-driven parameters for representing muscle Mass and Strength. Sens (Basel), 2023. 23(12).
38. Watanabe K, et al. Relationships between muscle strength and multi-channel surface EMG parameters in eighty-eight elderly. Eur Rev Aging Phys Act. 2018;15:3.
39. Liang T et al. Surface electromyography-based analysis of the lower limb muscle network and muscle synergies at various gait speeds. IEEE Trans Neural Syst Rehabil Eng, 2023. PP.
40. Jian Y, et al. An Electromyographic signal acquisition system for Sarcopenia. Cham: Springer Nature Switzerland; 2024.
41. Li N, et al. Exploration of a machine learning approach for diagnosing Sarcopenia among Chinese community-dwelling older adults using sEMG-based data. J Neuroeng Rehabil. 2024;21(1):69.
42. Leone A et al. Comparative analysis of supervised classifiers for the evaluation of Sarcopenia using a sEMG-Based platform. Sens (Basel), 2022. 22(7).
43. Kumar KS, et al. sEMG-based Sarcopenia risk classification using empirical mode decomposition and machine learning algorithms. Math Biosci Eng. 2024;21(2):2901–21.
44. De Luca CJ. The use of surface electromyography in biomechanics. J Appl Biomech. 1997;13(2):135–63.
45. Mathur S, Eng JJ, MacIntyre DL. Reliability of surface EMG during sustained contractions of the quadriceps. J Electromyogr Kinesiol. 2005;15(1):102–10.
46. Song K et al. Digital biomarkers for diagnosis of muscle disorders using stimulated muscle contraction signal. IEEE Trans Neural Syst Rehabil Eng, 2023.
47. McLester CN, et al. Reliability and agreement of various InBody body composition analyzers as compared to dual-energy X-Ray absorptiometry in healthy men and women. J Clin Densitom. 2020;23(3):443–50.
48. Cadenas-Sanchez C, et al. Reliability and validity of different models of TKK hand dynamometers. Am J Occup Ther. 2016;70(4):7004300010.
49. Brown KC, et al. Gait speed and variability for usual pace and pedestrian crossing conditions in older adults using the GAITRite walkway. Gerontol Geriatr Med. 2015;1:2333721415618858.
50. Song K et al. Digital biomarker for muscle function assessment using surface electromyography with electrical stimulation and a non-invasive wearable device. IEEE Trans Neural Syst Rehabil Eng, 2024.
51. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. 2005;27(8):1226–38.
52. Fuadah YN, Pramudito MA, Lim KM. An optimal approach for heart sound classification using grid search in hyperparameter optimization of machine learning. Bioeng (Basel), 2022. 10(1).
53. Lewis MJ, et al. Nestedcv: an R package for fast implementation of nested cross-validation with embedded feature selection designed for transcriptomics and high-dimensional data. Bioinform Adv. 2023;3(1):vbad048.
54. Varma S, Simon R. Bias in error` estimation when using cross-validation for model selection. BMC Bioinformatics. 2006;7:91.
55. Srivastava N, et al. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res. 2014;15(1):1929–58.
56. Yoon M-J et al. Evaluating Precision in Muscle Quality Prediction Using Stimulated Muscle Contraction Signal. The 18th Annual Conference of Korean Society for NeuroRehabilitation, 2024.