High Accuracy Machine Learning Model for Sarcopenia Severity Diagnosis based on Sit-to-stand Motion Measured by Two Micro Motion Sensors

Keer WANG#, Hongyu ZHANG#, Clio Yuen Man CHENG#, Meng CHEN, King Wai Chiu LAI, Calvin Kalun OR, Yonghua CHEN, Yong HU, Arul Lenus Roy VELLAISAMY, Cindy Lo Kuen LAM, Ning XI, Vivian W. Q. LOU*, and Wen Jung LI*

***Abstract*— In this ageing society, sarcopenia as a geriatric condition that can have significant negative impacts on an individual's quality of life. Sarcopenia is a kind of aged syndrome associated with loss of muscle mass and function, which may lead to falls, fractures, gait disorders or even mortality. There are multiple ways to diagnose sarcopenia, such as using Magnetic resonance imaging (MRI), Dual-energy X-ray absorptiometry (DEXA) and Bioelectrical impedance analysis (BIA) etc. to calculate muscle mass; using handgrip or sit-to-stand to measure muscle strength; using short physical performance battery (SPPB), gait, and 5-time sit-to-stand to evaluate physical performance.**

In this work, we use two μIMUs worn on subjects to record their sit-to-stand motion, and then used several machine learning models to diagnose the severity of sarcopenia of the subjects. We recruited 53 elderly subjects in total for this work. The youngest subject is 65 years old and the oldest is 84 years old. Their average age is 70 years old. Among these 53 subjects, there are 12 healthy ones and 41 sarcopenia patients with different severity. The subject is instructed to do the single sit-to-stand (STS) three times, and two μIMUs attached to the subject's waist and thigh transfer the data to a computer by Bluetooth. We separated the STS motion process into 4 phases based on the angle and angular velocity, extracted a total of 510 features for motion analytics. These features were futher analyzed by sequential feature selection with 5 different machine learning models (SVM, KNN, decision tree, LDA, and multilayer perceptron). With our proposed methodology, all 53 subjects could be classified as healthy or having sarcopenia with risk level 1, 2, or 3. The best accuracy to distinguish the healthy or sarcopenia subjects is 98.32%, and the best results to distinguish sarcopenia risk levels from 0 (healthy) to 3 (most severe) is 90.44%.

Keywords—sarcopenia diagnosis, sarcopenia risk levels, AI-based diagnosis, sit-to-stand motion, micro-IMU

I. Introduction

Sarcopenia is a progressive and generalized loss of skeletal muscle mass and strength syndrome due to ageing [1]-[3]. Both muscle mass and muscle strength will begin to decrease after a person reaching a certain age. Typically, muscle mass decreases at a rate of 3-5% per decade after 30 years old, and the speed of decrease accelerates a lot after age 60 [4]. Risks of adverse outcomes like frailty, fall, physical disability, and mortality also increase due to sarcopenia. The prevalence of sarcopenia rises with ageing and differs across different settings and clinical conditions. Diagnosing sarcopenia early and getting treatments as early as possible is important.

Different research groups have different definitions of sarcopenia [5]-[11]. There are three main consideration components of the sarcopenia diagnosis in summary: muscle mass, muscle strength and physical performance.

To measure muscle mass, there are mainly six methods: Anthropometry, Bioelectrical impedance analysis (BIA), Dual-energy X-ray absorptiometry (DEXA), Computer tomography (CT), Magnetic resonance imaging (MRI) and Ultrasound [1], [10]-[15]. Anthropometry is a simple and low-cost technique easily applied in clinical practice and large population-based surveys. Yet, the sensitivity of changes may be limited due to different observers. BIA is a method for estimating body composition, in particular, body fat and muscle mass. It is safe and has no radiation exposure [17]-[19]. However, it has a significant disadvantage in that the muscle mass measurements can be distorted by hydration status and the presence of oedema. DEXA is usually used as a means of measuring bone mineral density (BMD), but also sometimes used to assess body

This work was supported by the Theme-based Research Scheme under The University Grants Committee [#T42-717/20-R].

Keer WANG, Hongyu ZHANG, Meng CHEN and Wen Jung LI are with the Department of Mechanical Engineering, City University of Hong Kong, Hong Kong, China (e-mail: keerwang3-c@my.cityu.edu.hk; hyzhang42-c@my.cityu.edu.hk; menchen@cityu.edu.hk; wenjli@cityu.edu.hk).

Clio Yuen Man CHENG and Vivian W. Q. LOU are with the Department of Social Work and Social Administration; Sau Po Centre on Ageing, The University of Hong Kong, Hong Kong, China (e-mail: cliocym@connect.hku.hk; wlou@hku.hk).

King Wai Chiu LAI is with the Department of Biomedical Engineering, Centre for Robotics and Automation, City University of Hong Kong, Hong Kong, China (e-mail: kinglai@cityu.edu.hk).

Calvin Kalun OR, and Ning XI are with the Department of Industrial and Manufacturing System Engineering, The University of Hong Kong, Hong Kong, China (e-mail: klor@hku.hk; xining@hku.hk).

Yonghua CHEN is with the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong, China (e-mail: yhchen@hku.hk).

Yong HU is with the Department of Orthpedics & Traumatology, Li Ka-Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (e-mail: yhud@hku.hk).

Arul Lenus Roy VELLAISAMY is with the Hong Kong Metropolitan University, Hong Kong, China (e-mail: vroy@hkmu.edu.hk).

Cindy Lo Kuen LAM is with the Department of Family Medicine and Primary Care, Li Ka-Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (e-mail: clklam@hku.hk).

Keer WANG, Hongyu ZHANG, and Clio Yuen Man CHENG have equally contributed to this work.

Vivian W. Q. LOU and Wen Jung LI are co-contact authors of this work.

composition and fat content [20][21]. Even though the radiation exposure is small, it is expensive, time-consuming and has poor accessibility. CT scan is a medical imaging technique that can obtain detailed internal body images [22]. It can review the images after scanning. Since CT scan mainly utilizes X-ray, it is radiation exposure and requires interpretation by a radiologist. It also needs a confined space for the scanner. MRI is similar to CT but uses strong magnetic fields, magnetic field gradients and radio waves to generate organ images. It cannot use if the patient has metalwork/some pacemakers [23]. Ultrasound is a real-time medical imaging technique to visualize muscles, tendons and many internal organs [24][25]. It has no ionizing radiation and is highly safe. However, there are no criteria for diagnosis of low muscle mass with ultrasound.

To measure muscle strength, researchers mainly use handgrip strength, knee flexion/extension and peak expiratory flow [7]. Handgrip strength is a simple and convenient measure of muscle strength, and it is widely used in clinical practice. It strongly correlates with lower extremity muscle power, knee extension torque and calf cross-sectional muscle area [26]-[28]. To measure the knee flexion/extension, the subjects are instructed to sit in an adjustable straight-back chair, the lower leg unsupported and the knee flexed to 90 degrees and the force applied to the ankle will be recorded [29]. The knee flexible/extension is suitable for research but not very convenient in clinical practice for the need for special equipment and training. Peak expiratory flow measures the strength of respiratory muscles, but it is not recommended as an isolated measure of muscle strength for sarcopenia diagnosis [7].

To measure physical performance, Short Physical Performance Battery (SPPB), Gait, and Sit-to-stand (STS) are usually instructed [30]-[32]. SPPB is an objective assessment tool for evaluating lower extremity function in the elderly [33]. It involves a series of physical performance tests, including balance tests (stand side by side, semi-tandem and tandem), gait, and STS [33][34]. It is a standard physical performance measurement for both research and clinical practice. Both the gait and STS tests are part of the SPPB, but they can also be used individually as a single parameter for clinical practice and research. The STS test varies with different research, such as single sit-to-stand (STS), the 5-time sit-to-stand (5TSTS), 30s chair stand and timed get-up-and-go [35]-[39].

A micro inertial measurement unit (μIMU) is an electronic sensing device that can be used to measure the body's specific acceleration, angular rate, and orientation [41]. Because the μIMUs are sensitive, low cost and portable [44]-[46], they are attracting increased scientific attention for motion analysis on disease diagnosis. Cuzzolin et al. utilized an IMU to diagnose Parkinson's based on gait analysis and achieved an accuracy of 85.51% over Parkinson's patients and healthy subjects [47]. Pham and his group mainly detect Parkinson's through the sit-to-stand and stand-to-sit transitions with one IMU attached to the lower back [48]. Witchel et al. also assess the sit-to-stand and stand-to-sit transition with IMUs but in multiple sclerosis to measure the patient's postural control [49]. Ko et al. in 2021 evaluate the physical performance of the elderly by time-up-and-go test and gait to diagnose sarcopenia [50]. Kim's group analyzed the gait data with 2 IMUs on feet and can receive an accuracy of 88.69% to identify osteopenia and 93.75% for sarcopenia [51]. This work uses a μIMU to record sit-to-stand motion of subjects and applied machine learning related algorithms to different patients with different sarcopenia severity.

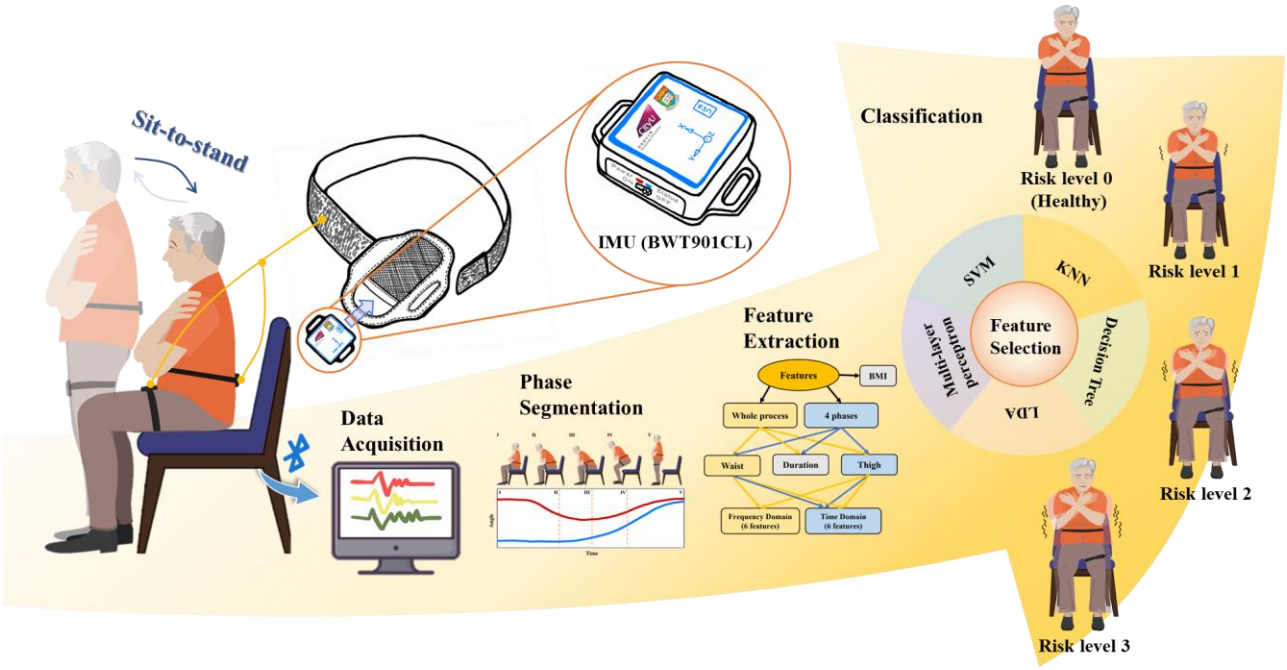

Fig. 1 The experimental process of sarcopenia diagnosis in this work.

II. Severities of Sarcopenia

There are many different groups to diagnose sarcopenia, but most of them just diagnose whether the subject has sarcopenia or not. However, some organizations also proposed different classifications of the severity of sarcopenia. EWGSOP introduced three conceptual stages of sarcopenia: pre-sarcopenia (low muscle mass), sarcopenia (low muscle mass and low muscle strength/low physical performance) and severe sarcopenia (low muscle mass, low muscle strength and low physical performance) [7]. EWGSOP2 in 2018 updated the operational definition of sarcopenia: probable sarcopenia (low muscle strength), sarcopenia (low muscle strength and low muscle quantity or quality), and severe sarcopenia (low muscle strength, low muscle quantity or quality and low physical performance) [5]. AWGS 2019 extended the diagnosis method from AWGS 2014 following extensive deliberations. They defined possible sarcopenia by low muscle strength with or without reduced physical performance; sarcopenia by low muscle mass, low muscle strength or low physical performance; severe sarcopenia by low muscle mass, low muscle strength and low physical performance.

The inclusion criteria are from the diagnosis approach of AWGS 2019. The cutoffs from the AWGS 2019 are shown below. Low muscle strength is defined as handgrip strength <28 kg for men and <18 kg for women; criteria for low physical performance are 6-m walk <1.0 m/s, Short Physical Performance Battery score ⩽9, or 5-time chair stand test ⩾12 seconds. The cutoffs for height-adjusted muscle mass are dual-energy X-ray absorptiometry, <7.0 kg/m2 in men and <5.4 kg/m2 in women, and bioimpedance, <7.0 kg/m2 in men and <5.7 kg/m2 in women [40].

The AWGS 2019 method is mainly used to diagnose whether the subject is healthy or not. A new definition of different levels of severities of sarcopenia is proposed based on AWGS 2019 method. There are three tasks from the physical performance part in the AWGS 2019 method: 6-metre walk, 5-time chair stand test and short physical performance battery (SPPB). Each task has a criterion: 6-metre walk < 1.0 m/s; 5-time chair stand test ⩾ 12s; SPPB ⩽ 9. Each criterion is a risk. If the subject meets one of these three, he/she is defined as risk level 1; If the subject meets any two of these three, he/she is defined as risk level 2; If the subject meets all these three, he/she is defined as risk level 3.

III. Experiment Design

To record the subjects' motion sit-to-stand, we utilize two μIMU sensors. The μIMU sensors we selected are commercial devices from WitMotion Co. Ltd (China). Its product model is BWT901CL. This kind of μIMU is small size, low-cost, portable, wireless and easy to operate. The IMU's dimension is 51 × 36 mm and 15 mm thick. The weight of the μIMU is only 20 grams. The parameters of the μIMU are shown in Table 1. We utilize the two 9-axis μIMUs to record the movement of sit-to-stand. The full-scale ranges of the accelerometer, gyroscope and magnetometer are ± 16g, $\pm 2000\ °/s$ and ± 2 Gauss, respectively. The sampling rate is 200 Hz. This kind of μIMU is chargeable and once it is fully charged, it can last for 4 hours. The data can be transferred to the computer by Bluetooth 2.0.

Table 1 Basic parameters of the μIMU (BWT901CL)

Dimension:	$51 \times 36 \times 15\ mm$
Weight:	$20g$
Measurement:	$9 - Axes$ (3-axes accelerometer, 3-axes gyroscope, and 3-axes magnetometer)
Sampling rate:	$200\ Hz$
Accelerometer parameters:	Range: $\pm 16g$; Revolution: $0.0005(g/LSB)$
Gyroscope parameters:	Range: $\pm 2000°/s$; Revolution: $0.061(°/s)/(LSB)$
Magnetometer parameters:	Range: $\pm 2\ Gauss$; Revolution: 0.0667 (mGauss/LSB)

We recruited 53 elderly subjects in total for this work. The youngest subject is 65 years old and the oldest is 84 years old. The average age is 70 years old. Among these 53 subjects, there are 12 healthy ones and 41 sarcopenia patients with different severity. According to the risk level definition mentioned previously, there are 12 patients in risk level 1, 6 patients in risk level 2 and 23 patients in risk level 3, making up the total of 41 sarcopenia patients. The subject demographics are shown in Table 2.

Table 2 Summary of participant demographics.

		n	Age (Years)	BMI (kg/m^2)
Healthy (Risk level 0)		12	67.4 ± 3.1	23.8 ± 2.8
Sarcopenia	**Risk level 1**	12	69.8 ± 5.0	23.9 ± 2.4
	Risk level 2	6	67.7 ± 1.9	25.1 ± 2.1
	Risk level 3	23	70.8 ± 4.2	24.0 ± 3.7
Total		53	70.0 ± 4.4	24.1 ± 3.2

The two μIMUs were attached to the subject's waist and thigh with belts separately, as shown in Fig. 2. The subjects were instructed to sit on an armless chair crossing their arms in front of the chest and then stand up. The height of the chair is 45cm. The subjects were not allowed to lean back on the chair or to assume a perched position. If the subject could complete the motion without any help, they could finish the motion with armrests. Each subject was asked to conduct the sit-to-stand process three times.

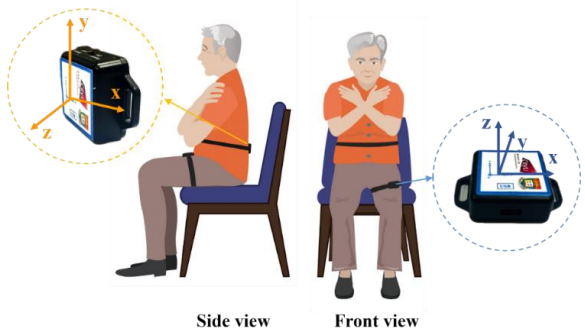

Fig. 2 Waist senor orientation has the x-axis put horizontally and pointing to the subject's left, the y-axis is vertical and points up, and the z-axis is in the sagittal plane with the direction parallel to the horizontal axis and points frontally. The thigh sensor's orientation is the same as the waist sensor.

IV. Data Analysis

A. Preprocessing

Since there are two μIMUs here, it's crucial to synchronize. The thigh data is recorded first and then the waist data. When the subject finishes the motion, the waist data is stopped first and then the waist data. The total data recorded from the thigh μIMU is more than that from the waist μIMU. In order to synchronize the two series of data, we cut the thigh data according to the absolute time.

To detect the starting moment, the sliding window is utilized. The length of the window is 40. OW is the standard deviation of the sliding window. "$OW > (0.05 * \textit{the mean value of } OW)$" can be viewed as a criterion. The first point to reach this criterion is defined as the starting point of sit-to-stand. The last point to reach this criterion is defined as the ending point of sit-to-stand. Fig. 3 shows the curve changes from 2 μIMUs during the single sit-to-stand.

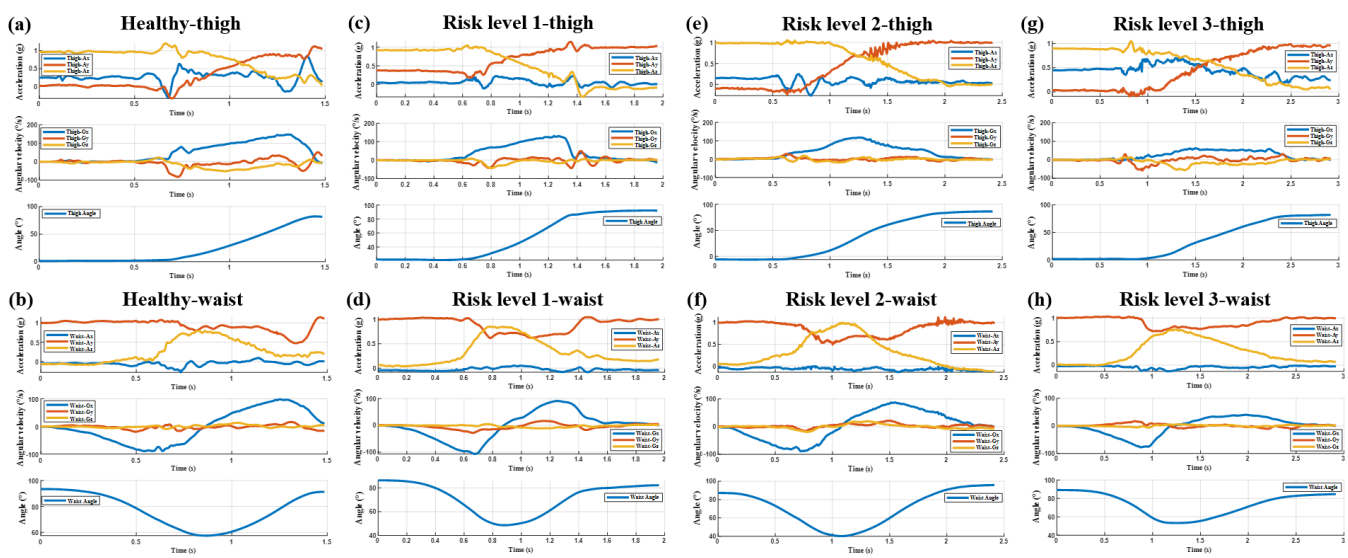

Fig. 3 (a) The sit-to-stand data of a healthy subject from the μIMU positioned on waist; (b) The sit-to-stand data of a healthy subject from the μIMU positioned on the thigh; (c) The sit-to-stand data of the subject with Risk level 1 from the μIMU positioned on waist; (d) The sit-to-stand data of the subject with Risk level 1 from the μIMU positioned on the thigh; (e) The sit-to-stand data of the subject with Risk level 2 from the μIMU positioned on waist; (f) The sit-to-stand data of the subject with Risk level 2 from the μIMU positioned on the thigh; (g) The sit-to-stand data of the subject with Risk level 3 from the μIMU positioned on waist; (h) The sit-to-stand data of the subject with Risk level 3 from the μIMU positioned on the thigh.

B. Phase Segmentation

Sit-to-stand is fundamental and essential in daily life. Many researchers have analyzed this motion and divided it into different phases. Among all these phase division methods, there are three fixed positions: Sitting, Seat-off and Standing. The Sitting position refers to the initial position of the motion of sit-to-stand. The Seat-off position refers to the position that the subject's body leaves the chair (there is no extra force applied to the chair). The Standing position refers to the end position of the motion of sit-to-stand.

In this paper, we divide the sit-to-stand motion into 5 critical points (4 phases) based on the collected μIMU data, as shown in Fig. 4. The main reference of the division method is the angular velocity and angle curves. These five points are identified as

1) STS initiation;
2) Peak flexion angular velocity;
3) Seat off;
4) Peak extension angular velocity;
5) STS termination.

STS initiation is defined by the starting point of sit-to-stand introduced in Preprocessing. Peak flexion angular velocity is the trough of the X-axis of angular velocity from the waist μIMU. Seat off is the point that the X-axis of angular velocity

from the waist μIMU equals zero or the peak of the waist angle. Peak extension angular velocity is the crest of the X-axis of angular velocity from the waist μIMU. STS termination is defined by the ending point of sit-to-stand introduced in preprocessing.

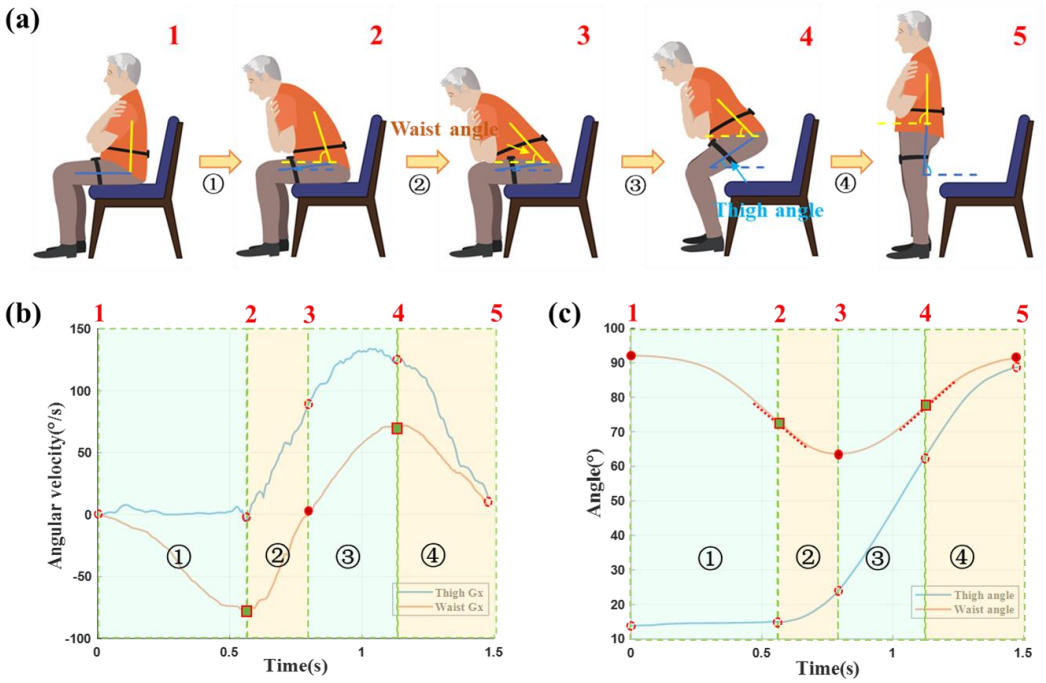

Fig. 4 The sit-to-stand motion can be segmented into 5 critical points (4 phases). (a) Shows five segmentation points of sit-to-stand; (b) shows the corresponding points on the curves of the x-axis of angular velocity of waist and thigh; (c) shows the corresponding points on the curves of the waist angle and thigh angle.

C. Feature Extraction

The features we extracted are from the sit-to-stand whole motion, 4 phases and personal information, as shown in Fig. 5. In the sit-to-stand part, we extract features from both the frequency domain and time domain. In the 4 phases part, we extract features from the time domain. Both the sit-to-stand part and 4 phases part include the features from the 7-axis of 2 μIMUs and their corresponding durations. For personal information, we calculate the BMI of each subject and take it as a feature.

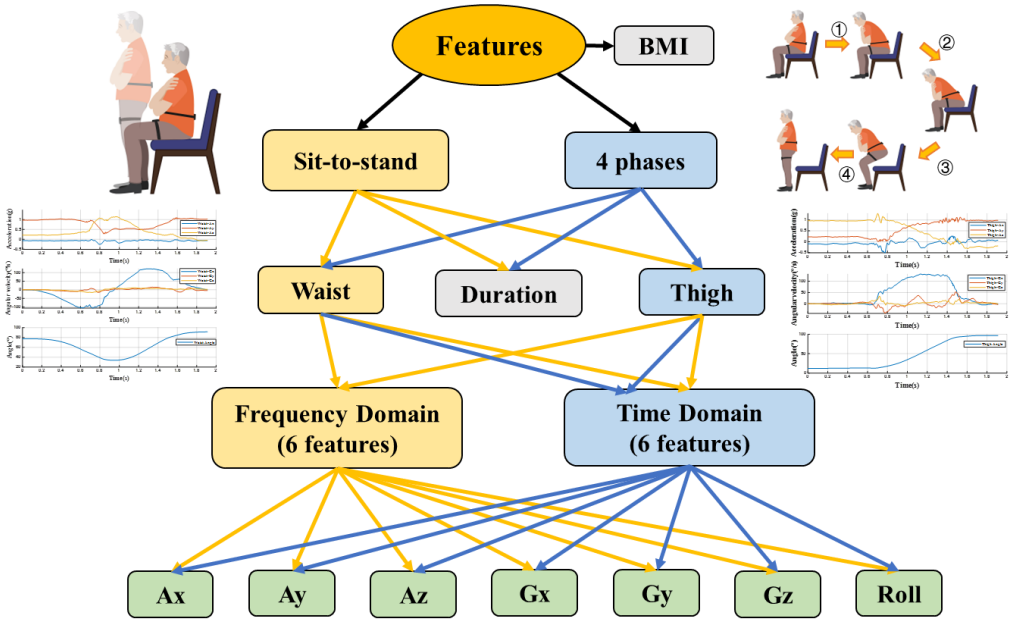

Fig. 5 Feature extraction from the STS and 4 phases.

The features extracted from the time domain and frequency domain are shown in Table 3. The features in the time domain include max, min, mean, peak to peak, standard deviation and root mean square. The features in the frequency domain include average frequency, Barycenter frequency, mean square frequency, frequency root mean square, frequency variance and frequency standard deviation.

Table 3 Statistical and morphological features.

Domain	Feature	Formula
Time Domain	Max	$x_{max} = \max_{n=1,2,\ldots N}\{x_n\}$
	Min	$x_{min} = \min_{n=1,2,\ldots,N}\{x_n\}$
	Mean	$\overline{x} = \frac{1}{N}\sum_{n=1}^{N} x_n$
	Peak to peak	$x_{p-p} = x_{max} - x_{min}$
	Standard deviation	$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}[x_n - \overline{x}]^2}$
	Root mean square	$x_{rms} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} {x_n}^2}$
Frequency Domain	Average frequency	$FA = \frac{1}{K}\sum_{k=1}^{K} S_k$
	Barycenter frequency	$FC = \frac{\sum_{k=1}^{K} f_k S_k}{\sum_{k=1}^{K} S_k}$
	Mean square frequency	$MSF = \frac{\sum_{k=1}^{K} {f_k}^2 S_k}{\sum_{k=1}^{K} S_k}$
	Frequency root mean square	$RMSF = \sqrt{MSF}$
	Frequency variance	$VF = \frac{\sum_{k=1}^{K}(f_k - FC)^2 S_k}{\sum_{k=1}^{K} S_k}$
	Frequency standard deviation	$RVF = \sqrt{VF}$

We separate these features into two feature sets: STS and STS+4 phases. The features of STS are 170 in total (2 μIMU × (6 Time Domain + 6 Frequency Domain) × 7 Axes + BMI + Sit-to-stand Duration = 170). The features of STS + 4 phases are 510 in total ([2 μIMU × (6 Time Domain + 6 Frequency Domain) × 7 Axes + Sit-to-stand Duration] + [2 μIMU × 4 Phases × 6 Time Domain + 4 × Phase Duration] + BMI = 510).

D. Machine Learning

1) Support Vector Machines (SVM)

SVM is a supervised two-class classifier. It is a general feed-forward neural network of traditional machine learning. The basic model of SVM is to find the best-separating hyperplane in the feature space to maximize the interval between positive and negative samples on the training set.

The interval of the function represents the certainty that the feature is positive or negative, given the training samples $(x^{(i)}, y^{(i)})$, then we can get

$$\gamma_i' = y_i(\omega^T x + b) \tag{1}$$

The geometric interval is the distance from the vector point to the overclocking surface. The equation is as following:

$$\gamma_i = \frac{y_i(\omega^T x + b)}{\|w\|} \tag{2}$$

Suppose there is a set of training samples with two classifications (+1 and -1), the decision surface is

$$\omega^T x + b = 0 \tag{3}$$

The x is the input vector, w is the adjustable weight vector, and b is the bias.

Suppose the model is linear, then

$$\begin{aligned} \omega^T x + b \geq 0 \quad if: d_i = +1 \\ \omega^T x + b \leq 0 \quad if: d_i = -1 \end{aligned} \tag{4}$$

Margin is twice the interval γ from the hyperplane to the nearest data point. If a plane can maximize γ, it is called the optimal hyperplane.

Suppose the weight and bias have optimal solution, the optimal hyperplane function and the corresponding discriminant function are

$$\begin{aligned} w_0^T x + b_0 = 0 \\ g(x) = w_0^T x + b_0 \end{aligned} \tag{5}$$

Then the (w_0, b_0) in the sample set $x^{(i)}$ must satisfy:

$$\begin{aligned} w_0^T x + b_0 \geq 0 \quad if: d_i = +1 \\ w_0^T x + b_0 \leq 0 \quad if: d_i = -1 \end{aligned} \tag{6}$$

The (x_i, d_i) is the support vector. These points are closest to the hyperplane. The distance between the points x and the optimal hyperplane is

$$\gamma = \frac{g(x)}{\|w\|} \tag{7}$$

In the positive and negative hyperplanes, any support vectors satisfying the following equation:

$$\gamma = \frac{g(x_p)}{\|w\|} = \frac{\pm 1}{\|w\|} \tag{8}$$

Margin can be defined as:

$$Margin = 2\gamma = \frac{2}{\|w\|} \tag{9}$$

In order to maximize γ, we need to minimize $\|w\|$. Maximize the margin between two class equals to minimize the Euclidean norm of weight vector w.

2) K-nearest Neighbor (KNN)

KNN is a machine learning technique and algorithm that can be used for both regression and classification tasks. KNN examines the labels of a chosen number of data points surrounding a target data point, in order to make a prediction about the class that the data point falls into. A KNN model calculates similarity using the distance between two points on a graph. The greater the distance between the points, the less similar they are. A KNN algorithm can be illustrate as following:

a) Setting K to the chosen number of neighbors.
b) Calculating the distance between a provided/test example and the dataset examples.
c) Sorting the calculated distances.
d) Getting the labels of the top K entries.
e) Returning a prediction about the test example.

In the first step, K is chosen individually and it tells the algorithm how many neighbors (how many surrounding data points) should be considered when rendering a judgment about the group the target example belongs to. In the second step, the model checks the distance between the target example and every example in the dataset. The distances are then added into a list and sorted. Afterward, the sorted list is checked and the labels for the top K elements are returned. When rendering a prediction about the target data point, it matters if the task is a regression or classification task. For a regression task, the mean of the top K labels is used, while the mode of the top K labels is used in the case of classification.

3) Decision Tree (DT)

A decision tree is a decision that goes from the root node to the leaf node step by step. It can be applied to both classification and regression. The leaf node is the final decision result.

The metric for feature segmentation is entropy. Entropy is the measure of how disorder the random variables is. Entropy is proportional to uncertainty. The equation of entropy is

$$H(x) = -\sum_{i=1}^{n} p_i \times \log p_i \tag{10}$$

Where p_i is the probability of case i. When the probability is 0 or 1, the random variable has an uncertainty relation of 1. When the probability is 0.5, the random variable is max.

The information gain indicates how the feature reduces the uncertainty relationship of the class, and the information gain index is usually used as the selection criterion of the node decision feature.

4) Linear Discriminate Analysis (LDA)

LDA is a supervised machine learning that project the data into a low-dimensional space, making the same type of data as compact as possible. The algorithm of LDA is as following.

Considering a set of data $D = (x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, x_i is a n-dimensional vector, class y_i belongs to $C_1, C_2, \dots, C_k$. We define the $N_{j\in1,2,\dots,k}$ the number of sample class j, $X_{j(j\in1,2,\dots,k)}$ the collection of sample class j, $\mu_{j(j\in1,2,\dots,k)}$ the average of sample class j.

$$\mu_j = \frac{1}{N_j}\sum_{x\in X_j} x \ \ (j \in 1,2,\dots,k) \tag{11}$$

$$\sigma_i = \sum_{x\in X_j} (x-\mu_j)(x-\mu_j)^T \ \ (j \in 1,2,\dots,k) \tag{12}$$

The between-class scatter matrix is

$$S_b = \sum_{j=1}^{k} N_j(\mu_j - u)\,(\mu_j - u)^T \tag{13}$$

Where u is the average value of all the data.

The Intra-class scatter matrix is

$$S_W = \sum_{j=1}^{k}\sum_{x\in X_j}(x-\mu_j)(x-\mu_j)^T \tag{14}$$

We project the low-dimensional space dimension into d-dimension. The corresponding basis vector is $\omega_1, \omega_2, \dots, \omega_d$, , forming a matrix W. Function J can be defined as below:

$$J = \frac{W^T S_b W}{W^T S_\omega W} = \prod_{i=1}^{d} \frac{{\omega_i}^T S_b \omega_i}{{\omega_i}^T S_\omega \omega_i} \tag{15}$$

$$S_\omega^{-1} S_b \omega_i = \lambda_i \omega_i \tag{16}$$

When equation (31) is existed, the function has maximum value. We take the eigenvectors corresponding to the first d largest eigenvalues to form matrix W. Since S_b is the addition of k matrices with a rank of 1, its rank is less than or equal to k. And because we know that after the first $(k-1)$ of μ_j, the last μ_k can be represented by the first $(k-1)$. The highest dimension after the dimensionality reduction of the LDA algorithm is $(k-1)$.

5) Multilayer perceptron (MLP)

A multilayer perceptron (MLP) is a fully connected class of feedforward artificial neural network (ANN). An MLP can be viewed as a directed graph consisting of multiple layers of nodes, each fully connected to the next layer. Except for the input node, each node is a neuron (or processing unit) with a nonlinear activation function. MLPs are usually trained by a supervised learning method known as the backpropagation algorithm. It generally consists of at least three layers of nodes: an input layer, a hidden layer and an output layer.

The algorithm for the MLP can be described as follows:

a) As with the perceptron, the inputs are passed through the MLP by taking the dot product of the input with the weights between the input layer and the hidden layer (WH). This dot product gives a value in the hidden layer.
b) MLPs make use of activation functions at each of their computed layers. There are many activation functions to discuss: rectified linear units (ReLU), sigmoid function, and tanh. Pass the computed output at the current layer through one of these activation functions.
c) Once the calculated output at the hidden layer has been pushed through the activation function, push it to the next layer in the MLP by taking the dot product with the corresponding weights.
d) Repeat step b and c until the output layer is reached.
e) At the output layer, the calculations are either used for a backpropagation algorithm according to the activation function selected for the MLP (in the case of training), or a decision is made based on the output (in the case of testing).

MLPs form the basis of all neural networks and have greatly increased the power of computers when applied to classification and regression problems. Thanks to the multilayer perceptron, computers are no longer limited by XOR cases and can learn rich and complex models.

E. Feature Selection

Sequential feature selection (SFS) can be used to identify the most problem-relevant features from all features. With SFS, we can reduce an initial a-dimensional feature space to a b-dimensional feature subspace (b<a) by using the greedy search algorithm. Fig. 6 shows the flow chart of SFS. We use 5 different machine learning models (SVM, KNN, DT, LDA, MLP) to select the 510 features. The selection criterion is the accuracy of the classification. The steps of SFS are as follows:

a) At the first time, one feature is put in the machine learning model, and it comes out as the best 1 feature for the classification.
b) Hold this feature, and group it with the other 509

features respectively as the input of the second time machine learning.

c) Repeat the steps until the input is 510 features.

d) Compare the accuracies of the 510 groups of features, the group with the highest accuracy is chosen as the best subset.

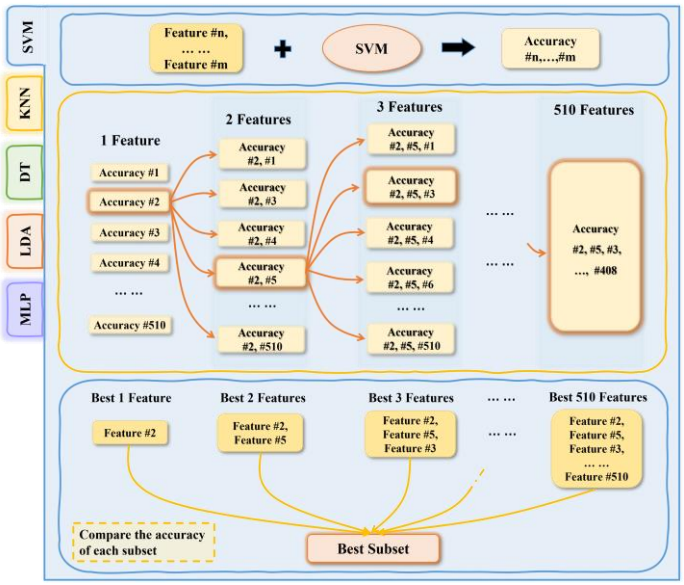

Fig. 6 Sequential Feature Selection (SFS).

We define the two classifications of healthy and sarcopenia subjects as Group 1, and the 4 classifications of risk level 0-3 subjects as Group 2. Table 4 shows the results of feature selection by SFS. We can see that in the comparison of Group 1, the DT method selects the fewest features (8 features), and the LDA method selects the most features (269 features). The SVM, KNN and MLP methods select 20 features, 11 features and 28 features respectively. To classify risk level 0-3, the method of DT, SVM and MLP choose 11 features, 20 features and 36 features. The features selected by the LDA and KNN methods for the comparison of Group 2 are a bit more, 65 features and 84 features respectively.

Table 4 Results of feature selection

Machine Learning	Feature Amount	
	Group 1 (Healthy Vs Sarcopenia)	Group 2 (Risk level 0-3)
SVM	20	20
KNN	11	84
DT	8	11
LDA	269	65
MLP	28	39

V. Results

Sarcopenia is a kind of disease that is quite prevalent for the elderly population beyond 60 years of age. The decreases in muscle mass and muscle strength easily lead to falls and bone fractures which are quite dangerous to the elderly. So, it is essential to diagnose sarcopenia and its severities early. In this paper, we utilize 2 μIMUs to analyze the movement of sit-to-stand for diagnosing different severity of sarcopenia.

We proposed a novel sit-to-stand phase segmentation method based on μIMU data. By dividing the process of sit-to-stand, we could extract more features and better analyze the motion.

We put the 2 feature sets (STS and STS+4 phases) into five classification models (SVM, LDA, DT, KNN and MLP) separately for different group classifications. The feature set STS has 170 features in total mainly from the motion of sit-to-stand. The feature set of STS+4 phases has 510 features in total from both the motion of sit-to-stand and 4 phases of the motion. For the classification models, the rate of the training set and testing set is 7:3 and the cross-validation is 10 folds.

The results shown are the average accuracies of 5 times classifications.

Fig. 7 shows the results of feature extraction. In the classification of healthy and sarcopenia subjects (Group 1), feature set STS can get the highest accuracy of 93.34% by KNN among five classification models; The feature set STS+4 phases can get the highest accuracy of 96.66% by MLP among five classification models. In the classification of risk level 0-3 subjects (Group 2), both the STS and STS+4 phases feature sets get the highest accuracy 90.44% and 85.82% respectively by using the classification model KNN.

(a) Group 1: Healthy Vs Sarcopenia

Classification model	STS	STS+4 phases
SVM	92.92 %	95.66 %
LDA	86.24 %	95.42 %
DT	90 %	89.12 %
KNN	93.34 %	94.64 %
MLP	91.26 %	96.66 %

(b) Group 2: Risk level 0-3

Classification model	STS	STS+4 phases
SVM	84.18 %	82.48 %
LDA	54.16 %	77.9 %
DT	76.68 %	68.36 %
KNN	90.44 %	85.82 %
MLP	86.66 %	84.16 %

Fig. 7 The results of the feature extraction. (a) shows the accuracies of STS and STS+4 phases with 5 classification models (SVM, LDA, DT. KNN and MLP) in the comparison of healthy and sarcopenia subjects. (b) shows the accuracies of STS and STS+4 phases with 5 classification models in the comparison of risk level 0-3 subjects.

We select features from 510 features by sequential feature selection. Fig. 8 shows the results of feature selection and classifications. The first blue boxes in Fig. 7 (a) and (b) show the classification results with features selected out by SFS with SVM. The yellow boxes show the classification results with features selected out by SFS with KNN. The green boxes show

the classification results with features selected out by SFS with DT. The red boxes show the classification results with features selected out by SFS with LDA. The purple boxes show the classification results with features selected out by SFS with MLP. In each box, there are five bars which refer to the accuracies obtained from the five classification models (SVM, KNN, DT, LDA and MLP).

In Group 1, to classify the healthy and sarcopenia subjects, we can see that feature set selected out by MLP (28 features) leads to the highest accuracy 98.32% using the classification SVM model. In Group 2, to classify the risk level 0-3 subjects, the feature set selected out by MLP (39 features) leads to the highest accuracy 84.58% calculated by the classification model KNN.

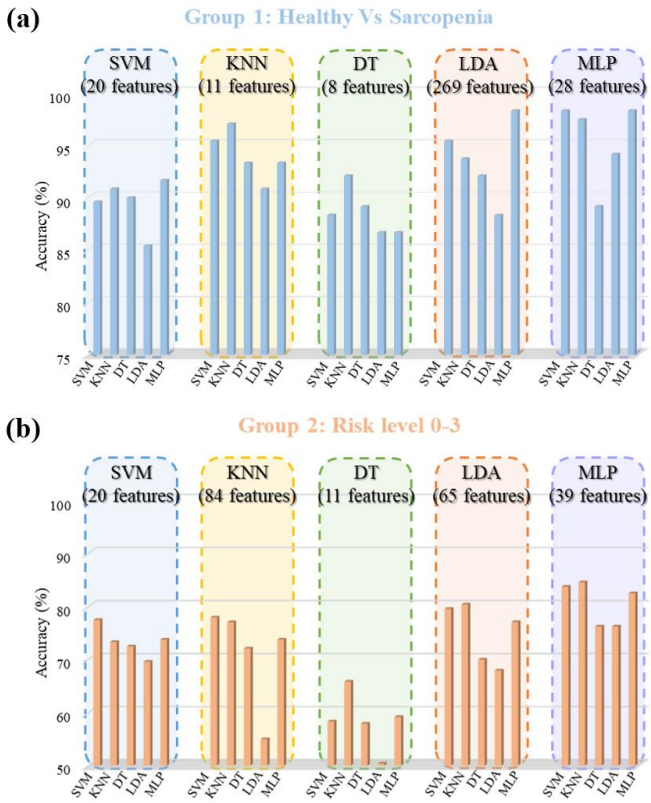

Fig. 8 The results of feature selection. (a) Shows the accuracies calculated by 5 feature sets selected out by feature selection with 5 classification models (SVM, LDA, DT. KNN and MLP) in the comparison of healthy and sarcopenia subjects. (b) shows the accuracies calculated by 5 feature sets selected out by feature selection with 5 classification models in the comparison of risk level 0-3 subjects.

In summary, the highest accuracy obtained in classifying the healthy and sarcopenia subjects is 98.32%, which was obtained by using the SVM classification model with 28 features selected out by SFS; the highest accuracy obtained in classifying the risk level 0-3 subjects is 90.44% using the KNN classification model feature set STS (170 features).

VI. Conclusion

In this work, we presented our results on diagnosing the severity of sarcopenia of 41 subjects by sit-to-stand motion with two wearable μIMUs. We proposed a novel way to analyze the motion of sit-to-stand using 2 μIMUs attached on the subjects' thigh and waist and separating the motion into 4 phases. In total, 510 features can be extracted from the entire sit-to-stand motion and the 4 phases. With the help of machine learning models, we can classify the healthy and sarcopenia subjects with a high accuracy of 98.32%. Moreover, we have obtained an accuracy of 90.44% in distinguishing the sarcopenia subjects of three different risk levels 1, 2, and 3.

References

[1] Ata, A. M., Kara, M., Kaymak, B., & Özçakar, L. (2020). Sarcopenia Is Not "love": You Have to Look Where You Lost it!. American Journal of Physical Medicine & Rehabilitation, 99(10), e119-e120.

[2] Beaudart, C., Zaaria, M., Pasleau, F., Reginster, J. Y., & Bruyère, O. (2017). Health outcomes of sarcopenia: a systematic review and meta-analysis. PloS one, 12(1), e0169548.

[3] Lauretani, F., Russo, C. R., Bandinelli, S., Bartali, B., Cavazzini, C., Di Iorio, A., ... & Ferrucci, L. (2003). Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. Journal of applied physiology, 95(5), 1851-1860.

[4] Rubbieri, G., Mossello, E., & Di Bari, M. (2014). Techniques for the diagnosis of sarcopenia. Clinical Cases in Mineral and Bone Metabolism, 11(3), 181.

[5] Cruz-Jentoft, A. J., Bahat, G., Bauer, J., Boirie, Y., Bruyère, O., Cederholm, T., ... & Zamboni, M. (2019). Sarcopenia: revised European consensus on definition and diagnosis. Age and ageing, 48(1), 16-31.

[6] Muscaritoli, M., Anker, S. D., Argilés, J., Aversa, Z., Bauer, J. M., Biolo, G. I. A. N. N. I., ... & Sieber, C. C. (2010). Consensus definition of sarcopenia, cachexia and pre-cachexia: joint document elaborated by Special Interest Groups (SIG)"cachexia-anorexia in chronic wasting diseases" and "nutrition in geriatrics". Clinical nutrition, 29(2), 154-159.

[7] Cruz-Jentoft, A. J., Baeyens, J. P., Bauer, J. M., Boirie, Y., Cederholm, T., Landi, F., ... & Zamboni, M. (2010). Sarcopenia: European consensus on definition and diagnosisReport of the European Working Group on Sarcopenia in Older PeopleA. J. Cruz-Gentoft et al. Age and ageing, 39(4), 412-423.

[8] Chen, L. K., Liu, L. K., Woo, J., Assantachai, P., Auyeung, T. W., Bahyah, K. S., ... & Arai, H. (2014). Sarcopenia in Asia: consensus report of the Asian Working Group for Sarcopenia. Journal of the American Medical Directors Association, 15(2), 95-101.

[9] McLean, R. R., Shardell, M. D., Alley, D. E., Cawthon, P. M., Fragala, M. S., Harris, T. B., ... & Dam, T. T. L. (2014). Criteria for clinically relevant weakness and low lean mass and their longitudinal association with incident mobility impairment and mortality: the foundation for the National Institutes of Health (FNIH) sarcopenia project. Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 69(5), 576-583.

[10] Morley, J. E. (2008). Sarcopenia: diagnosis and treatment. The Journal of Nutrition Health and Aging, 12(7), 452-456.

[11] Beaudart, C., Reginster, J. Y., Slomian, J., Buckinx, F., Dardenne, N., Quabron, A., ... & Bruyere, O. (2015). Estimation of sarcopenia prevalence using various assessment tools. Experimental gerontology, 61, 31-37.

[12] Cooper, C., Fielding, R., Visser, M. V., Van Loon, L. J., Rolland, Y., Orwoll, E., ... & Kanis, J. A. (2013). Tools in the assessment of sarcopenia. Calcified tissue international, 93(3), 201-210.

[13] Cruz-Jentoft, A. J., & Sayer, A. A. (2019). Sarcopenia. The Lancet, 393(10191), 2636-2646.

[14] Edwards, M. H., & Buehring, B. (2015). Novel approaches to the diagnosis of sarcopenia. Journal of Clinical Densitometry, 18(4), 472-477.

[15] Fielding, R. A., Vellas, B., Evans, W. J., Bhasin, S., Morley, J. E., Newman, A. B., ... & Zamboni, M. (2011). Sarcopenia: an undiagnosed

condition in older adults. Current consensus definition: prevalence, etiology, and consequences. International working group on sarcopenia. Journal of the American Medical Directors Association, 12(4), 249-256.

[16] Wang, J., Thornton, J. C., Kolesnik, S., & Pierson Jr, R. N. (2000). Anthropometry in body composition: an overview. Annals of the New York Academy of Sciences, 904(1), 317-326.

[17] Abu Khaled, M., McCutcheon, M. J., Reddy, S., Pearman, P. L., Hunter, G. R., & Weinsier, R. L. (1988). Electrical impedance in assessing human body composition: the BIA method. The American journal of clinical nutrition, 47(5), 789-792.

[18] Karlsson, R., Roos, H., Fägerstam, L., & Persson, B. (1994). Kinetic and concentration analysis using BIA technology. Methods, 6(2), 99-110.

[19] Szabo, A., Stolz, L., & Granzow, R. (1995). Surface plasmon resonance and its use in biomolecular interaction analysis (BIA). Current opinion in structural biology, 5(5), 699-705.

[20] Haarbo, J., Gotfredsen, A., Hassager, C., & Christiansen, C. J. C. P. (1991). Validation of body composition by dual energy X - ray absorptiometry (DEXA). Clinical Physiology, 11(4), 331-341.

[21] Kohrt, W. M. (1998). Preliminary evidence that DEXA provides an accurate assessment of body composition. Journal of applied physiology, 84(1), 372-377.

[22] Goldman, L. W. (2007). Principles of CT and CT technology. Journal of nuclear medicine technology, 35(3), 115-128.

[23] Hashemi, R. H., Bradley, W. G., & Lisanti, C. J. (2012). MRI: the basics: The Basics. Lippincott Williams & Wilkins.

[24] Dalecki, D. (2004). Mechanical bioeffects of ultrasound. Annu. Rev. Biomed. Eng., 6, 229-248.

[25] Leighton, T. G. (2007). What is ultrasound?. Progress in biophysics and molecular biology, 93(1-3), 3-83.

[26] Norman, K., Stobäus, N., Gonzalez, M. C., Schulzke, J. D., & Pirlich, M. (2011). Hand grip strength: outcome predictor and marker of nutritional status. Clinical nutrition, 30(2), 135-142.

[27] Innes, E. V. (1999). Handgrip strength testing: a review of the literature. Australian Occupational Therapy Journal, 46(3), 120-140.

[28] McGrath, R. P., Kraemer, W. J., Snih, S. A., & Peterson, M. D. (2018). Handgrip strength and health in aging adults. Sports medicine, 48(9), 1993-2000.

[29] Amis, A. A., Senavongse, W., & Bull, A. M. (2006). Patellofemoral kinematics during knee flexion - extension: An in vitro study. Journal of orthopaedic research, 24(12), 2201-2211.

[30] Guralnik, J. M., Branch, L. G., Cummings, S. R., & Curb, J. D. (1989). Physical performance measures in aging research. Journal of gerontology, 44(5), M141-M146.

[31] Studenski, S., Perera, S., Wallace, D., Chandler, J. M., Duncan, P. W., Rooney, E., ... & Guralnik, J. M. (2003). Physical performance measures in the clinical setting. Journal of the American Geriatrics Society, 51(3), 314-322.

[32] Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., ... & Wallace, R. B. (1994). A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. Journal of gerontology, 49(2), M85-M94.

[33] Pavasini, R., Guralnik, J., Brown, J. C., Di Bari, M., Cesari, M., Landi, F., ... & Campo, G. (2016). Short physical performance battery and all-cause mortality: systematic review and meta-analysis. BMC medicine, 14(1), 1-9.

[34] de Fátima Ribeiro Silva, C., Ohara, D. G., Matos, A. P., Pinto, A. C. P. N., & Pegorari, M. S. (2021). Short physical performance battery as a measure of physical performance and mortality predictor in older adults: A comprehensive literature review. International Journal of Environmental Research and Public Health, 18(20), 10612.

[35] Cruz-Jentoft, A. J., & Morley, J. E. (Eds.). (2021). Sarcopenia. John Wiley & Sons.

[36] Alcazar, J., Losa-Reyna, J., Rodriguez-Lopez, C., Alfaro-Acha, A., Rodriguez-Mañas, L., Ara, I., ... & Alegre, L. M. (2018). The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. Experimental gerontology, 112, 38-43.

[37] Balachandran, A. T., Vigotsky, A. D., Quiles, N., Mokkink, L. B., Belio, M. A., & Glenn, J. M. (2021). Validity, reliability, and measurement error of a sit-to-stand power test in older adults: A pre-registered study. Experimental Gerontology, 145, 111202.

[38] Lee, S. Y., Choo, P. L., Pang, B. W. J., Lau, L. K., Jabbar, K. A., Seah, W. T., ... & Wee, S. L. (2021). SPPB reference values and performance in assessing sarcopenia in community-dwelling Singaporeans–Yishun study. BMC geriatrics, 21(1), 1-11.

[39] Yoshiko, A., Ogawa, M., Shimizu, K., Radaelli, R., Neske, R., Maeda, H., ... & Akima, H. (2021). Chair sit-to-stand performance is associated with diagnostic features of sarcopenia in older men and women. Archives of gerontology and geriatrics, 96, 104463.

[40] Chen, L. K., Woo, J., Assantachai, P., Auyeung, T. W., Chou, M. Y., Iijima, K., ... & Arai, H. (2020). Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. Journal of the American Medical Directors Association, 21(3), 300-307.

[41] Iosa, M., Picerno, P., Paolucci, S., & Morone, G. (2016). Wearable inertial sensors for human movement analysis. Expert review of medical devices, 13(7), 641-659.

[42] Leroy, E., Fisette, P., & Abedrabbo Ode, G. (2019). Analysis of the spine through a multibody model and IMU technology (Doctoral dissertation, Université catholique de Louvain).

[43] Ahmad, N., Ghazilla, R. A. R., Khairi, N. M., & Kasi, V. (2013). Reviews on various inertial measurement unit (IMU) sensor applications. International Journal of Signal Processing Systems, 1(2), 256-262.

[44] Wang, Y., Chen, M., Wang, X., Chan, R. H., & Li, W. J. (2018). IoT for next-generation racket sports training. IEEE Internet of Things Journal, 5(6), 4558-4566.

[45] Seel, T., Raisch, J., & Schauer, T. (2014). IMU-based joint angle measurement for gait analysis. Sensors, 14(4), 6891-6909.

[46] Faisal, A. I., Majumder, S., Scott, R., Mondal, T., Cowan, D., & Deen, M. J. (2020). A simple, low-cost multi-sensor-based smart wearable knee monitoring system. IEEE Sensors Journal, 21(6), 8253-8266.

[47] Cuzzolin, F., Sapienza, M., Esser, P., Saha, S., Franssen, M. M., Collett, J., & Dawes, H. (2017). Metric learning for Parkinsonian identification from IMU gait measurements. Gait & posture, 54, 127-132.

[48] Pham, M. H., Warmerdam, E., Elshehabi, M., Schlenstedt, C., Bergeest, L. M., Heller, M., ... & Maetzler, W. (2018). Validation of a lower back "wearable"-based sit-to-stand and stand-to-sit algorithm for patients with Parkinson's disease and older adults in a home-like environment. Frontiers in neurology, 9, 652.

[49] Witchel, H. J., Oberndorfer, C., Needham, R., Healy, A., Westling, C. E., Guppy, J. H., ... & Klucken, J. (2018). Thigh-derived inertial sensor metrics to assess the sit-to-stand and stand-to-sit transitions in the timed up and go (TUG) task for quantifying mobility impairment in multiple sclerosis. Frontiers in neurology, 9, 684.

[50] Ko, J. B., Kim, K. B., Shin, Y. S., Han, H., Han, S. K., Jung, D. Y., & Hong, J. S. (2021). Predicting Sarcopenia of Female Elderly from Physical Activity Performance Measurement Using Machine Learning Classifiers. Clinical Interventions in Aging, 16, 1723.

[51] Kim, J. K., Bae, M. N., Lee, K., Kim, J. C., & Hong, S. G. (2022). Explainable Artificial Intelligence and Wearable Sensor-Based Gait Analysis to Identify Patients with Osteopenia and Sarcopenia in Daily Life. Biosensors, 12(3), 167.