

Core stability status classification based on mediolateral head motion during rhythmic movements and functional movement tests

DIGITAL HEALTH
Volume 9: 1–13
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076231186217
journals.sagepub.com/home/dhj



Siwoo Jeong¹, Si-Hyun Kim² and Kyue-Nam Park¹ 

Abstract

Objective: Core stability assessment is paramount for the prevention of low back pain, with core stability being considered as the most critical factor in such pain. The objective of this study was to develop a simple model for the automated assessment of core stability status.

Methods: To assess core stability—defined as the ability to control trunk position relative to the pelvic position—we used an inertial measurement unit sensor embedded within a wireless earbud to estimate the mediolateral head angle during rhythmic movements (RMs) such as cycling, walking, and running. The activities of muscles around the trunk were analyzed by an experienced, highly trained individual. Functional movement tests (FMTs) were performed, including single-leg squat, lunge, and side lunge. Data was collected from 77 participants, who were then classified into good and poor core stability groups based on their Sahrman core stability test scores.

Results: From the head angle data, we extrapolated the symmetry index (SI) and amplitude of mediolateral head motion (Amp). Support vector machine and neural network models were trained and validated using these features. In both models, the accuracy was similar across three feature sets for RMs, FMTs, and full, and support vector machine accuracy (~87%) is greater than neural network (~75%).

Conclusion: The use of this model, trained with head motion-related features obtained during RMs or FMTs, can help to accurately classify core stability status during activities.

Keywords

Core stability, mediolateral head motion, cycling, gait, functional movement tests, machine learning

Submission date: 15 December 2022; Acceptance date: 8 June 2023

Introduction

Core stability is regarded as the ability to control the position of the trunk relative to the pelvic position,^{1–3} which helps to maintain balance and a stable posture. Poor core stability alters trunk motion patterns and can result in musculoskeletal injury or pain.^{4–6} Therefore, core stability assessment is important for the prevention of musculoskeletal injury and pain.⁷ The incorporation of core stability into rehabilitation and training programs may allow some individuals to improve their core stability before injury onset. However, the requirement for specialized equipment to measure muscle contractions (e.g.

electromyography or ultrasound machines)^{8–13} and the need for highly experienced operators to conduct clinical tests hinder assessments of core stability.^{14–17}

¹Department of Physical Therapy, Jeonju University, Jeonju, Korea

²Department of Physical Therapy, Sangji University, Wonju, Korea

Corresponding author:

Kyue-Nam Park, Department of Physical Therapy, 303, Cheonjam-ro, Wansan-gu, Jeonju-si, Jeollabuk-do, Republic of Korea.
Email: knpark@jj.ac.kr



Research over the past decade has utilized inertial measurement units (IMUs) for the assessment of various tasks such as functional movement tests (FMTs), walking, running, and single-leg squats (SLS).^{18–23} This emerging trend in literature is partly driven by the advantages of IMUs, including their portability, low cost, and ability to provide real-time feedback.^{24,25} Even ear-worn sensors or head-worn sensors have been employed in recent studies for gait analysis and balance assessment, further illustrating the versatility of this technology.^{26,27} Although these studies were not primarily focused on the assessment of core stability, they have gradually expanded the scope of IMU applications in the field.

Trunk symmetry and range of motion during distal extremity movements can serve as indices of core stability. To achieve core stability, core muscle activities should involuntarily occur prior to the movement of distal extremities.^{28–31} For example, rapid arm raising involves core muscle contractions that begin prior to movement.^{30,32} These preprogrammed, feedforward core muscle contractions provide the postural stability necessary for distal extremity movements.^{29,33,34} Insufficient core stability leads to an abnormal range of reactive trunk motion. Low back pain patients exhibit abnormal core muscle activation; similarly, cyclists with low back pain showed a loss of co-contraction of core muscles during trunk motion.³⁵ Individuals with low back pain also exhibit asymmetrical trunk motion during sitting and standing.³⁶ An unstable core can lead to low back pain and kinematic asymmetry during movements of the distal segment of the trunk.

A recently developed automated assessment system acquires information that aids clinical decision-making through joint kinematics (measured by an IMU) and sensors worn on the body during FMTs.^{37–40} This information can be used to develop machine learning models. In

most studies using this system, greater than or equal to three sensors were placed on several joints to obtain kinematic data.^{39,40} However, this method is not user-friendly; the sensors must be attached to bare skin or tight clothing to minimize the effect of cloth motion. To overcome this limitation, we developed a system that measures the symmetry index (SI) and amplitude (Amp) of mediolateral head motion using an IMU sensor equipped with an easy-to-use wireless earbud (Figure 1). According to Morrison et al. (2015), head angle serves as a proxy for trunk motion during rhythmic movements (RMs); mediolateral head motion is associated with mediolateral trunk motion during gait. Our system for measurement of mediolateral head motion could replace existing systems that measure trunk motion, although the wireless earbud has only one IMU sensor.

This study evaluates a novel approach to core stability assessment by developing an automated system that leverages a combination of modern technology and machine learning algorithms. The novelty of this study lies in the use of a wireless earbud equipped with an IMU sensor, designed to measure the SI and Amp derived from the head angle during motor tasks. This head angle serves as a proxy for trunk motion, thus providing indirect measures of core stability.

We hypothesized that this proposed system could effectively assess core stability and categorize individuals into good and poor core stability groups based on these measurements. Participants performed motor tasks that trigger reactive trunk movement, such as cycling, walking, and running. Mediolateral trunk motion, which arises in response to the reaction force during these tasks, is dependent on core stability. This was also demonstrated in FMTs, such as SLS, lunge, and side lunge, where controlled trunk motion is essential to maintain balance and prevent falling.

Participants initially underwent the Sahrman core stability test (SCST), administered by a highly trained, experienced operator, which objectively evaluates core stability. Based on their SCST scores, participants were then categorized into good and poor core stability groups. Subsequently, the SI and Amp were extracted from the head angle measurements obtained during the motor tasks, resulting in three sets of labeled data: a full feature set, a RM feature set, and an FMT feature set. Each feature set was utilized to train and validate machine learning models. To our knowledge, this study is among the first to apply machine learning models for such a purpose, marking a significant step forward in the field of core stability assessment.

Methods

This study is a cross-sectional investigation designed to evaluate the different machine learning classifiers in predicting core stability status. The study involved a total of 77 participants, who were assessed on their core stability

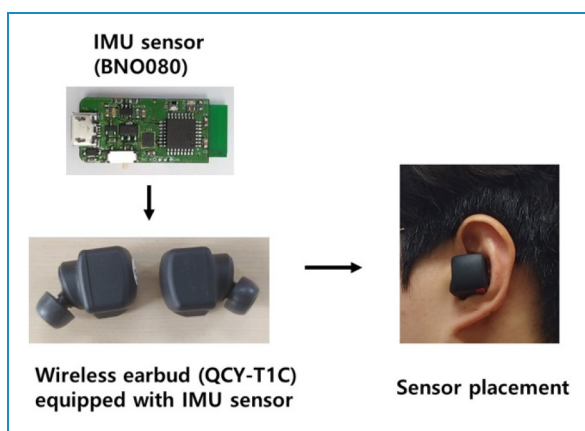


Figure 1. IMU sensor worn during motor tasks. Each participant wore a wireless earbud equipped with an IMU sensor in the left ear to measure head motion during the performance of motor tasks. IMU: inertial measurement unit.

using the SCST. Participants then performed a series of FMTs and RMs, during which mediolateral head angle data were recorded using an IMU. Three different feature sets were derived from these data, and support vector machine (SVM) and neural network (NN) were trained on

these features to predict core stability status. The performance of these machine learning models was assessed through various metrics including accuracy, sensitivity, specificity, and G-index (Figure 2).

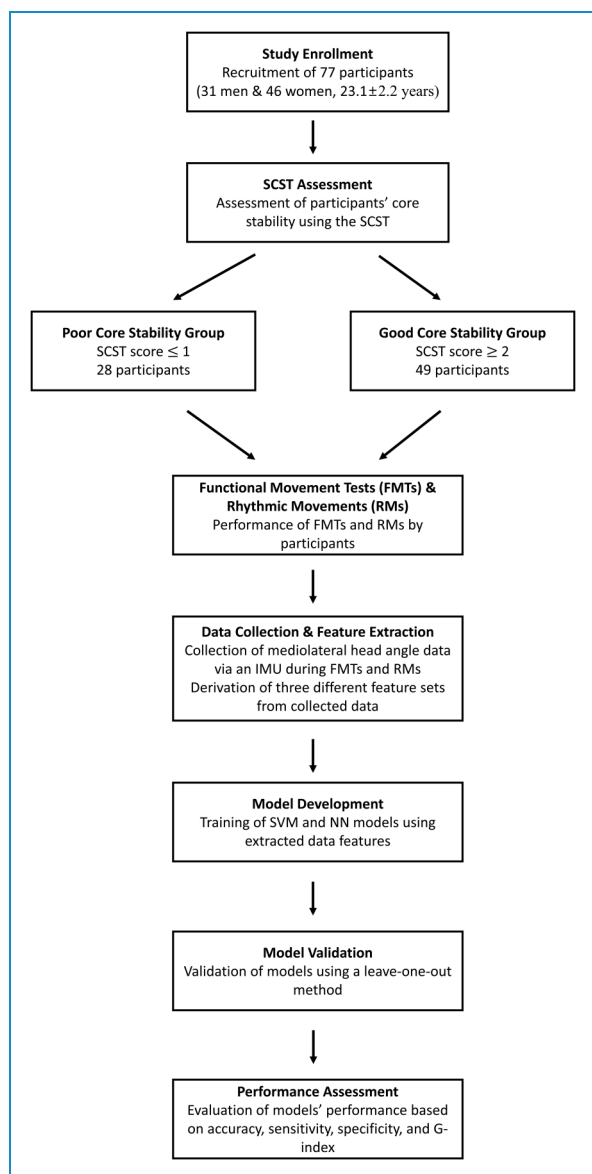


Figure 2. Flow diagram illustrating the methodological process of the study. The process starts with the recruitment of participants, followed by the assessment of their core stability using the SCST. The participants then performed FMTs and RMs, while the mediolateral head angle data were recorded using an IMU. These data are used to derive three different feature sets, which are utilized in the training of SVM and NN models. These models are validated using a leave-one-out method, and their performance is assessed in terms of accuracy, sensitivity, specificity, and G-index. FMTs: functional movement tests; IMU: inertial measurement unit; NN: neural network; RMs: rhythmic movements; SCST: Sahrman core stability test; SVM: support vector machine.

Participants

We enrolled 77 participants in this study (31 men and 46 women; mean age, 23.1 ± 2.2 years; mean body weight, 63.0 ± 13.3 kg). The exclusion criteria were as follows: overweight; current pregnancy; psychological, neurological, vestibular, or cardiopulmonary disorder; and musculoskeletal pathologies. All participants provided demographic information and visual analog scale pain scores. Anthropometric data were also obtained, including height and weight. All participants provided written informed consent, and the study protocol was approved by the Institutional Review Board of Jeonju University.

Sahrman core stability test

Each participant completed the SCST as an objective measure of core stability (Figure 3). The SCST comprises five tasks of increasing difficulty. The inflatable pad of a pressure biofeedback unit (Stabilizer; Chattanooga Group, Inc., Hixson, TN, USA) was placed in the natural lordotic curve when the participant was in the supine position. After the pad had been inflated to 40 mmHg, the participant began the SCST. When the participant had completed one of the SCST tasks with a deviation of ≤ 10 mmHg, the next task began. Task performance was scored using a 5-point scale. Participants with scores < 2 were assigned to the poor core stability group, whereas participants with scores of 2–5 were assigned to the good core stability group.

Rhythmic movements

Cycling

As a warm-up activity, each participant cycled for 5 min at their preferred speed, followed by a 5-min rest period; the participant was then asked to cycle as fast as possible. The fastest measured speed was used to calculate the target speed (70% of the fastest speed). The participant then cycled for 1 min at the target speed (Figure 4A, left panel).

Walking and running

The walking and running speeds were set as 4.5 and 9 km/h, respectively, and each participant performed a 30-s practice session to ensure familiarity with these speeds. After the practice session, the participant was asked to walk and

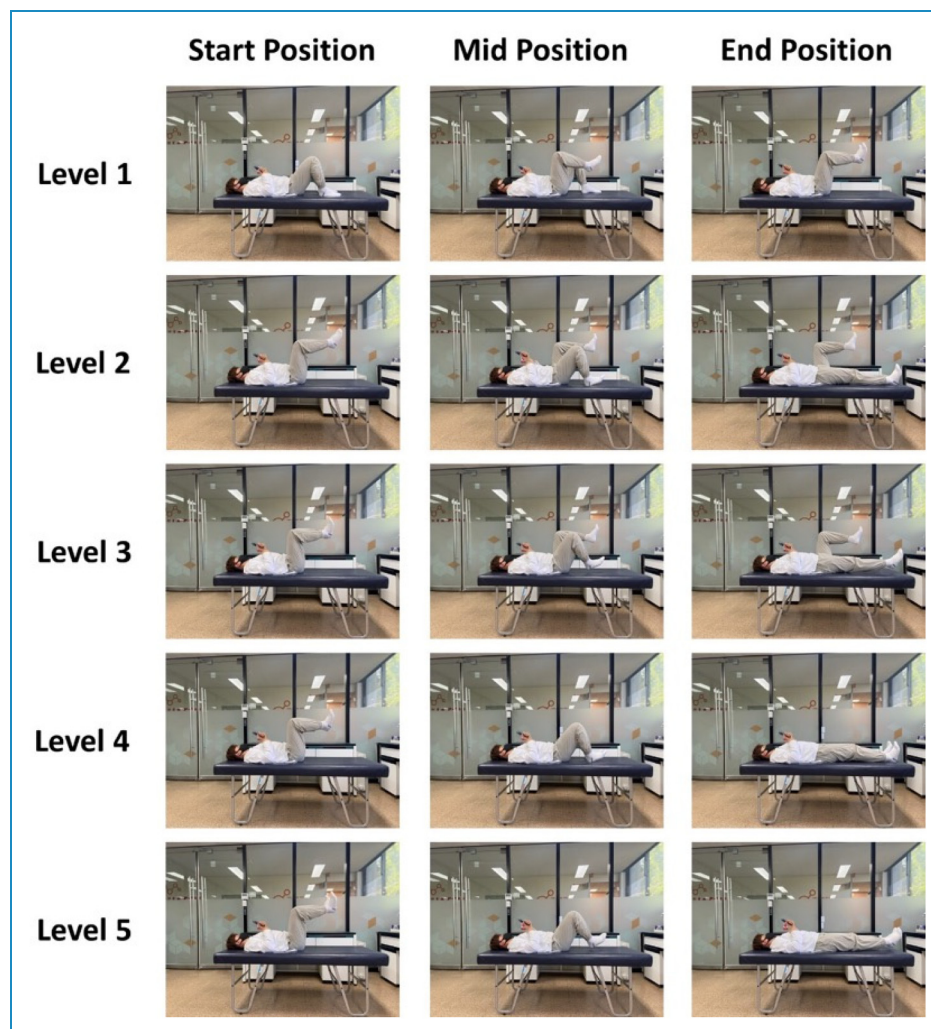


Figure 3. The five levels of the Sahrman core stability test.

run for 1 min on a treadmill (Xiaomi, Inc., Beijing, China) at the predefined speed (Figure 4A, middle and right panels).

Functional movement tests

Single-leg squat

Each participant performed five repetitions of the SLS, with both the left and right legs. The participant began in a standing position with the toes of the weight-bearing leg pointing forward, hands on the waist, and nonweight-bearing leg flexed at the knee (at 90°). The participant was then instructed to lift the heel while performing the SLS, without losing their balance (Figure 4B, right panel). In the event that a participant did not follow the instructions or the nonweight-bearing leg contacted the ground, the trial was considered unsuccessful and the participant was asked to perform it again.

Side lunge

Each participant performed five side lunges on both the right and left sides. Both start and end positions for the side lunge comprised standing upright with both feet together and hands on the waist. From the starting position, the participant lunged to the side and then pushed back to the starting position. The stride length was the length that allowed both feet to point forward with the knee of the front leg flexed at $\sim 90^\circ$ while remaining over the foot, with the back leg oriented in a straight position (Figure 4B, middle panel). If the participant lost their balance, they were asked to perform the side lunge again.

Lunge

Each participant performed five lunges with both the right and left legs. Both start and end positions comprised standing upright with the toes pointing forward and hands on the



Figure 4. Motor tasks. Each participant performed RMs (A) and FMTs (B) while wearing the wireless earbud. (A) RMs included cycling, running, and walking. Each participant performed each motor task for 1 min. (B) FMTs included lunge, side lunge, and single-leg squat exercises. All tasks were performed five times, with both the right and left leg, at the preferred speed. FMTs: functional movement tests; RMs: rhythmic movements.

waist. The stride length during the lunge was the length that allowed 90° knee flexion for both legs (Figure 4B, left). All participants were able to complete the lunges without loss of balance.

Data collection

A high-resolution IMU (BNO080; Ceva Technologies, Rockville, MD, USA) was used for head angle measurements. The IMU sensor was embedded into a left-sided wireless earbud (QCY-T1C; Dongguan Hele Electronics, Dongguan, China) (see Figure 1). The acceleration data were transferred to a self-developed mobile app via Bluetooth at 100 Hz. This app computed the mediolateral head angle in real time, using methods established in previous studies.^{41,42} Each sample carried signed 16-bit acceleration outputs for all three axes x , y , and z . These acceleration outputs were scaled to equate to a proportional value of Earth's gravity (G_{px} , G_{py} , and G_{pz}). The full-scale

IMU sensor value was equivalent to the maximum signed 16-bit integer value at 8 G.

The IMU's orientation can be delineated using its roll (ϕ), pitch (θ), and yaw (ψ). The corresponding matrices for roll, pitch, and yaw are as follows (equations 1–3):

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix} \quad (1)$$

$$R_y(\theta) = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (2)$$

$$R_z(\psi) = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

When the IMU sensor undergoes rotation, the vector of Earth's gravitational field (g) also rotates, influenced by the pitch, roll, and yaw matrices (equation 4). When an arbitrary accelerometer reading is G_p , it can be expressed as follows (equation 5):

$$R_y(\phi)R_x(\theta)R_z(\psi) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \theta \cos \phi \\ \sin \phi \\ \cos \theta \cos \phi \end{pmatrix} \quad (4)$$

$$\frac{1}{\sqrt{G_{px}^2 + G_{py}^2 + G_{pz}^2}} \begin{pmatrix} G_{px} \\ G_{py} \\ G_{pz} \end{pmatrix} = \begin{pmatrix} -\sin \theta \cos \phi \\ \sin \phi \\ \cos \theta \cos \phi \end{pmatrix} \quad (5)$$

The roll and pitch angles can then be determined from the following (equations 6–7):

$$\phi = \tan^{-1} \left(\frac{G_{py}}{\sqrt{G_{px}^2 + G_{pz}^2}} \right) \quad (6)$$

$$\theta = \tan^{-1} \left(\frac{-G_{px}}{G_{pz}} \right) \quad (7)$$

To filter out linear acceleration brought on by the participant's movements, the accelerometer output underwent a low-pass filter before angle calculation. This step is crucial to avoid the inaccuracy in angle estimation due to additional acceleration from the participant's movement. The accelerometer output should ideally be near 1 G, signifying the existence of only gravitational acceleration. After the low-pass filtering process, the accelerometer output magnitude converged to 1 G, within a 5% difference.

For calibration of offsets, an initial 1-s measurement was taken. The average offsets recorded in the sagittal and frontal angles during this time were utilized for offset removal. During this 1-s calibration period, the subject was instructed to stay still, while 100 data samples were gathered and evaluated for calibration accuracy. Calibration was considered

successful only if the difference between the maximum and minimum values of both ϕ and θ was less than 0.5° . If this condition was not met, the testing process restarted from the calibration stage. Each participant wore the wireless earbud while performing the RMs and clinical tests (Figure 5). Frontal plane head angle data were collected at 100 Hz.

Feature extraction

Two features were extracted from mediolateral head angle data recorded during the FMTs and RMs: the SI and Amp. During the FMTs, participants performed SLS, lunge, and side lunge exercises (five repetitions of each) with both the right and left legs. When performing the SLS and lunge, the head was tilted to the same side as the weight-bearing leg side; it was tilted to the opposite side during the side lunge. The maximum head angles on the right and left sides were determined for each trial. The mean Amp was calculated by summing the maximum right and left head angles during cycling (five trials) (Figure 6):

$$\begin{aligned} \text{amp}_{\max} &= \sum (\text{left}_{\max} + \text{right}_{\max}), \\ \text{amp}_{\min} &= \sum (\text{left}_{\min} + \text{right}_{\min}), \\ \text{amp}_{\text{mean}} &= \sum (\text{left}_{\text{mean}} + \text{right}_{\text{mean}}), \end{aligned}$$

where left_{\max} , left_{\min} , and $\text{left}_{\text{mean}}$ represent the maximum, minimum, and mean values of five maximum left head angles, respectively; right_{\max} , right_{\min} , and $\text{right}_{\text{mean}}$ indicate the maximum, minimum, and mean values of five maximum right head angles, respectively.

The SI during the FMTs reflected the difference between the left and the right maximum head angles (Figure 6):

$$\begin{aligned} \text{SI}_{\max} &= \sum |\text{left}_{\max} - \text{right}_{\max}|, \\ \text{SI}_{\min} &= \sum |\text{left}_{\min} - \text{right}_{\min}|, \\ \text{SI}_{\text{mean}} &= \sum |\text{left}_{\text{mean}} - \text{right}_{\text{mean}}|. \end{aligned}$$

During RMs, the maximum right and left head angles were extracted over a 1-min period of cycling. Maximum, minimum, and mean Amp values were obtained from pairs of maximum right and left head angles (Figure 7):

$$\begin{aligned} \text{Amp}_{\max} &= \text{Max}(\text{left}_i + \text{right}_i), \\ \text{Amp}_{\min} &= \text{Min}(\text{left}_i + \text{right}_i), \\ \text{Amp}_{\text{mean}} &= \frac{\sum_i (\text{left}_i + \text{right}_i)}{n}, \end{aligned}$$

where left_i and right_i represent the maximum left and right values for the i th cycling trials, respectively, and n is the total number of trials.

The SI during RMs was also calculated from the maximum right and left head angles obtained over a 1-min period for each cycling trial (Figure 7):

$$\begin{aligned} \text{SI}_{\max} &= \text{Max}(|\text{left}_i - \text{right}_i|), \\ \text{SI}_{\min} &= \text{Min}(|\text{left}_i - \text{right}_i|), \\ \text{SI}_{\text{mean}} &= \frac{\sum_i (|\text{left}_i - \text{right}_i|)}{n}. \end{aligned}$$

An SI closer to 0° indicates a more symmetrical mediolateral head angle.

Feature set

For both the FMTs (SLS, lunge, and side lunge) and RMs (cycling, running, and walking), the maximum, minimum, and mean Amp and SI values were used to derive three feature sets (Table 1):

- A full feature set including all Amp_{\max} , Amp_{\min} , Amp_{mean} , SI_{\max} , SI_{\min} , and SI_{mean} values from the FMTs and RMs, as well as the maximum cycling speed (37 features).
- A feature set including the Amp_{\max} , Amp_{\min} , Amp_{mean} , SI_{\max} , SI_{\min} , and SI_{mean} values only from the FMTs (18 features).
- A feature set including the maximum cycling speed and Amp_{\max} , Amp_{\min} , Amp_{mean} , SI_{\max} , SI_{\min} , and SI_{mean} values only from the RMs (19 features).

Two-class classifiers: training and validation

Two-class classifiers were trained to determine core stability status. Using MATLAB R2022a software (MathWorks, Natick, MA, USA), six models were generated: SVM, decision tree, discriminant analysis, logistic regression, naïve Bayes, k-nearest neighbor, and NN models. The SVM and NN models exhibited the best classification performance, and we only report the results for these classifiers below.

Before model training, oversampling was conducted to balance the number of samples in the poor and good core stability groups, as the good core stability group initially had more samples. Specifically, we implemented random oversampling, a technique that works by duplicating examples from the minority class in the training data set. In this method, samples from the minority class (in our case, the poor core stability group) are selected at random and added to the training data set, increasing its size and thus balancing the class distribution.

Achieving balancing the class distribution can enhance the performance of the machine learning model,⁴³ particularly in our case where the data imbalance was significant. It's important to note that this method may increase the likelihood of overfitting since it duplicates the minority class events. However, to mitigate this potential issue, we repeated the procedure 10 times, which helped ensure a

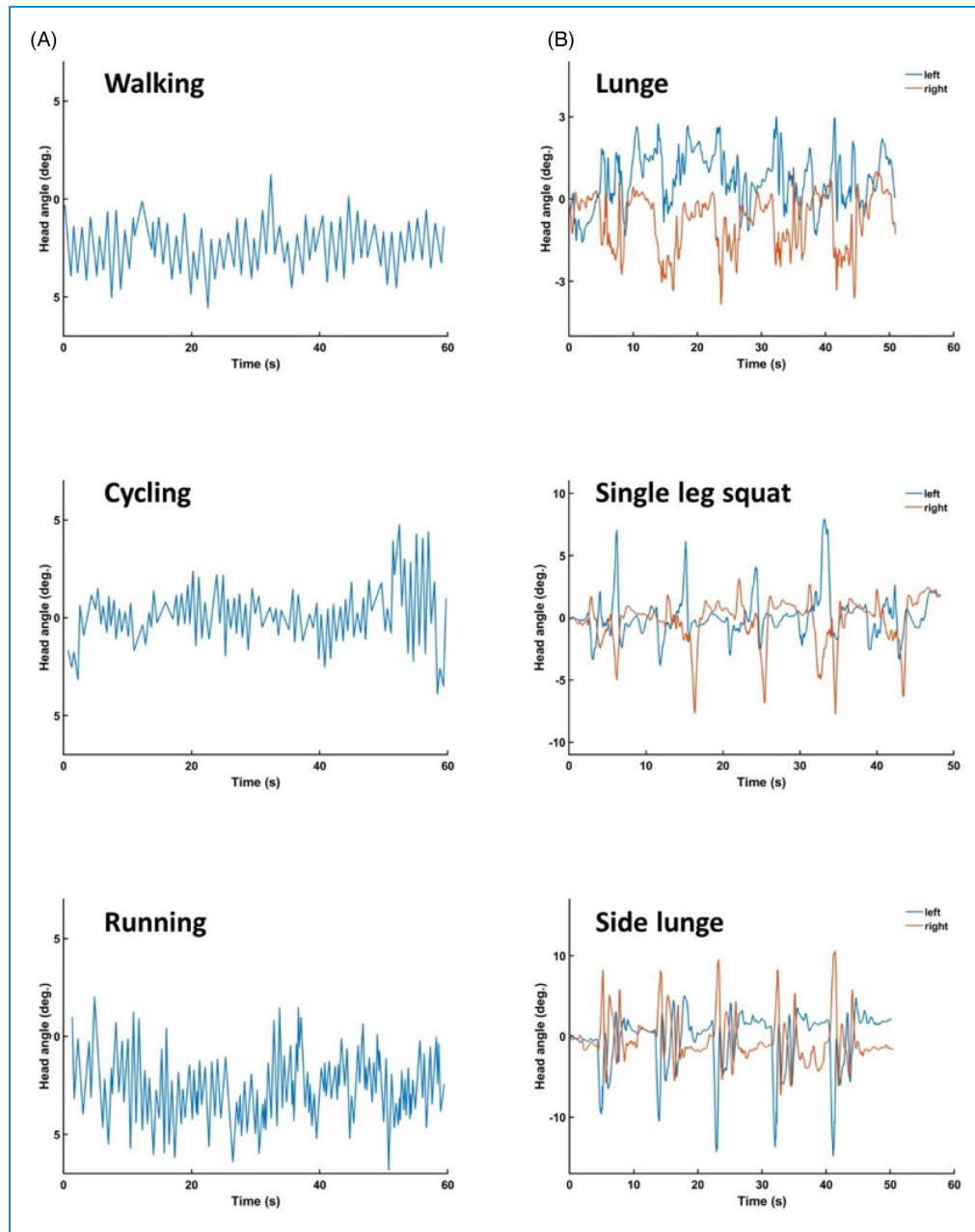


Figure 5. Example mediolateral head angle data obtained during RMs (A) and FMTs (B). (A) Each participant performed RMs, including walking, cycling, and running, for 1 min at a constant speed. (B) Each participant also completed lunge, single-leg squat, and side lunge FMTs (five repetitions of each exercise). FMTs: functional movement tests; RMs: rhythmic movements.

more diverse set of minority class examples in the training data set, thus reducing the probability of overfitting.

SVM models search the optimal hyperplane at the point where the margin between two classes of data is maximal. According to Ganyun et al. (2005), SVM models are appropriate for small data sets.⁴⁴ Our SVM model was implemented in MATLAB with the “fitsvm” function. We employed a Gaussian (also known as radial basis function (RBF)) kernel function due to its flexibility in handling nonlinear

classification problems. The RBF kernel introduces non-linearity without needing to increase the computational cost excessively. The kernel scale parameter was set to 1.1, controlling the sensitivity of the Gaussian kernel and indirectly influencing the decision boundary.

The NN model was trained using MATLAB’s “fitnet” function. The model consisted of an input layer, three hidden layers, and an output layer. The input layer was designed to accommodate the dimensionality of our feature

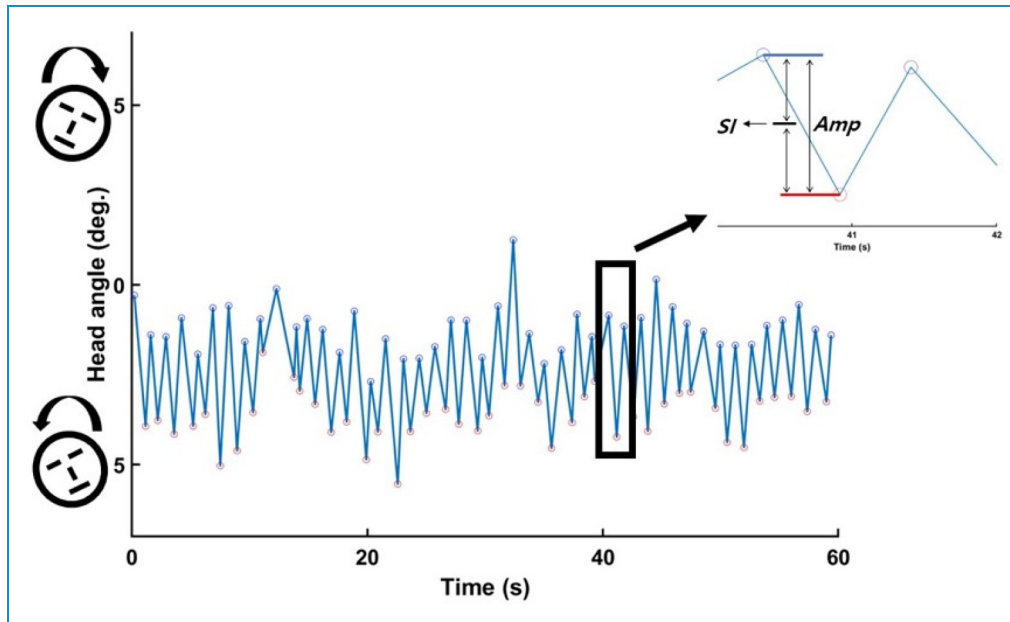


Figure 6. Mediolateral head angle during walking. Each participant walked for 1 min at 4.5 km/h. Angles above and below 0° represent left and right head tilt directions, respectively. The left (upper circles) and right (lower circles) peak angles calculated during each cycling exercise were used to calculate the SI and Amp of the head angle; an SI value closer to 0° indicated more symmetric head movement in the frontal plane. The mean SI and Amp for the five cycling repetitions were calculated, and the maximum and minimum SI and Amp were extracted. Amp: amplitude; SI: symmetry index.

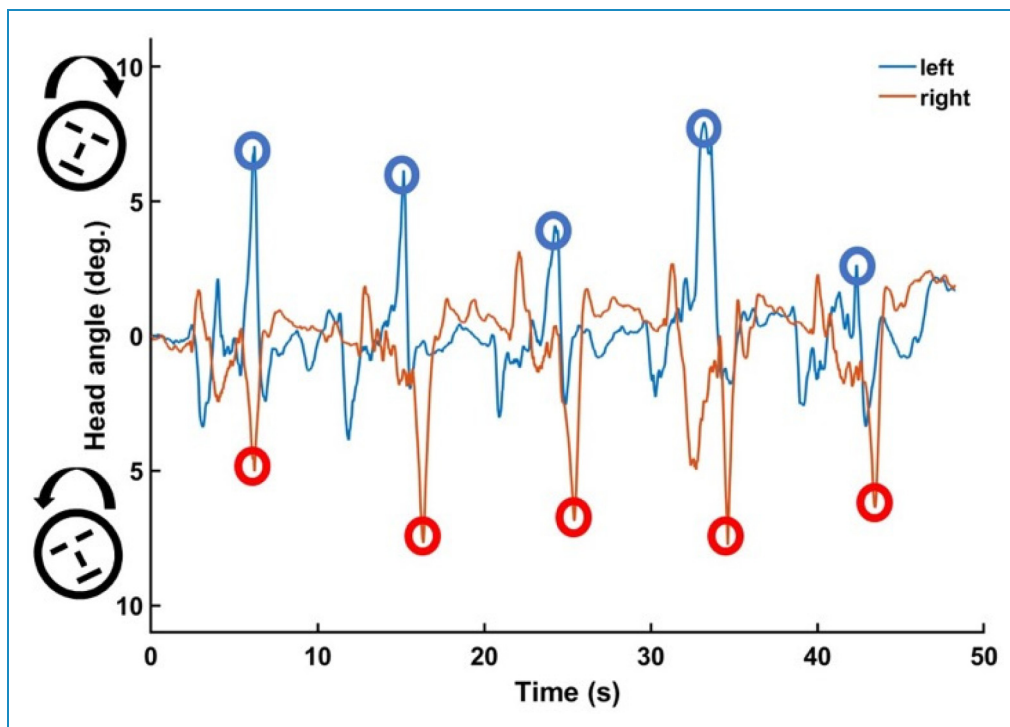


Figure 7. Mediolateral head angle during SLS. Each participant performed five repetitions of the SLS exercise at their preferred speed. Angles above and below 0° represent left and right head tilt directions, respectively. The left (upper circles) and right (lower circles) peak angles for each SLS repetition were used to calculate the SI and Amp of the head angle in the frontal plane. A SI value of 0° indicates perfectly symmetrical mediolateral head motion. The mean SI and Amp for the five SLS repetitions were calculated, and the maximum and minimum SI and Amp were extracted. Amp: amplitude; SLS: single-leg squat; SI: symmetry index.

Table 1. Feature sets used for model training.

Feature set	Data source	Description
FMTs feature set	IMU sensor	Comprises maximum, minimum, and mean symmetry index and amplitude data obtained during mediolateral head motion when performing FMTs
RMs feature set	IMU sensor	Comprises maximum, minimum, and mean symmetry index and amplitude data obtained during mediolateral head motion while cycling, running, and walking

FMTs: functional movement tests; IMU: inertial measurement unit; RMs: rhythmic movements.

set. Each hidden layer contained 10 neurons, with the Rectified Linear Unit (ReLU) activation function applied to their outputs. This nonlinear activation function introduces nonlinearity into our model, helping it learn from the complex relationships in our data. The output layer of the NN was designed to classify each sample into one of the two classes: “good” and “poor.” Although this is a binary classification task, MATLAB’s “fitnet” function defaults to the softmax activation function for the output layer, regardless of the number of classes. This is because softmax is a more general function that is used for multiclass classification problems, and it simplifies to the sigmoid function in the case of binary classification. To optimize the weights of the NN, the “fitnet” function uses an iterative training process with a maximum iteration limit of 1000. This function automatically manages the batch size and the number of epochs for training, based on the size of the training data.

Leave-one-participant-out cross-validation was used to evaluate the accuracies of all models. This validation method is appropriate for small sample sizes,³⁹ as in our study. The model was trained using the full data set after exclusion of one participant’s data; the trained model then labeled the data of the excluded participant based on the input.

Performance evaluation

Model performance was evaluated by comparison of predicted and actual labels. Accuracy, sensitivity, specificity, and precision values, as well as the *F1*-score and G-index, were calculated as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%,$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\%,$$

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100\%,$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\%,$$

$$F1 - \text{score} = \frac{\text{Sensitivity} \times \text{Precision}}{\text{Sensitivity} + \text{Precision}} \times 2,$$

$$G\text{-index} = \sqrt{(1 - \text{Sensitivity})^2 + (1 - \text{Precision})^2},$$

where TP is true-positive (number of participants correctly included in good core stability group), TN true-negative (number of participants correctly included in poor core stability group), FP is false-positive (number of participants incorrectly included in good core stability group), and FN is false-negative (number of participants incorrectly included in poor core stability group).

The SVM and NN binary classifier models were run 10 times, and the results are reported as means and standard deviations.

Results

Demographic data

There were no significant differences in age, height, weight, body mass index, or visual analog scale pain scores between the good core stability and poor core stability groups (Table 2).

Classification accuracy

The results of leave-one-participant-out cross-validation are reported in Table 3. Compared with the NN model (73.7 ± 3.6%), the SVM model showed higher classification accuracy (86.6 ± 1.7%) in the RM feature set (Table 3). The SVM model also showed better specificity, precision, *F1*-score, and G-index values for all feature sets; however, the NN model had higher sensitivity (Table 3). The SI and Amp of mediolateral head motion during the SLS were extracted from the FMT data set; the SI and Amp of mediolateral head motion during cycling, walking, and running were extracted from the RM data set. The SVM and NN models both exhibited better performance when trained on the RM feature set, compared with the FMT and full feature sets.

Table 2. Demographics of good core stability and poor core stability groups.

	Good core stability	Poor core stability	P-value
No. of participants	49	28	-
Sex (male/female)	17/32	14/14	-
Age (years)	23.0 ± 1.8	23.3 ± 2.8	0.66
Height (cm)	166.4 ± 9.1	167.1 ± 8.8	0.74
Weight (kg)	62.0 ± 13.9	64.7 ± 12.4	0.39
BMI (kg/m ²)	22.4 ± 3.1	23.0 ± 3.1	0.37
VAS score	3.1 ± 3.2	4.4 ± 2.6	0.074

BMI: body mass index, VAS: visual analog scale.

Discussion

This study was conducted to develop machine learning models that could classify young people without musculoskeletal disorders according to core stability status, as determined using the SCST. Among the 77 participants, 49 and 28 were classified into good core stability and poor core stability groups, respectively. There were no significant demographic differences between the groups, which minimized the risk of confounding bias. The SI and Amp of mediolateral head motion extracted during the FMTs and RMs were used to train SVM and NN models. Model classification accuracy was similar across all three data sets (full, FMT, and RM), and the SVM model showed superior performance compared with the NN model.

In this study, we achieved high classification accuracy (~85%) with the SVM model, even without using variables such as jerking movements, acceleration, or velocity that previous studies have used.^{37–39,45} The novelty in our approach was the use of mediolateral head motion during RMs, as a proxy for trunk motion and thus, core stability. It implies that head motion during RMs, as measured by our system, is reflective of the preprogrammed, feedforward contractions of the core muscles that provide postural stability necessary for distal extremity movements.^{29,34} These contractions have been previously found to be closely related to core stability, particularly during the movement of distal extremities.

We examined the classification performance of SVM and NN models using different feature sets to identify individuals with good or poor core stability. Interestingly, we found that all feature sets produced similar accuracy levels in both models. These results suggest that the chosen features may not have a significant impact on the

classification performance of the models and that a simpler feature set may be sufficient for accurate classification of core stability status. However, it is important to consider that the nature of the tasks during which the features are measured may influence the involuntary reactive trunk motion and, thus, the accuracy of core stability assessment. For example, the use of different FMTs or exercises could result in different levels of trunk motion, affecting the classification accuracy. Further investigation is needed to determine the most appropriate feature set and task for accurate classification of core stability status.

The SVM model achieved a G-index close to 0.25, which is typically considered indicative of a well-optimized model.⁴⁶ It also exhibited very high specificity and precision but relatively lower sensitivity, resulting in a higher number of false negatives. This means that a significant number of individuals with good core stability were misclassified as having poor core stability. However, the high specificity and precision suggest that those identified as having poor core stability are very likely to actually have poor core stability. Therefore, the SVM model could be useful in identifying individuals with poor core stability but may need to be combined with other measures to accurately identify those with good core stability. In contrast, the NN model showed a more balanced number of false positives and false negatives, indicating that it may be better at accurately identifying individuals with good core stability. However, the lower specificity and precision suggest that individuals identified as having poor core stability may not necessarily have poor core stability. Therefore, the NN model could be useful in identifying individuals with good core stability. Overall, these findings suggest that both models have their strengths and weaknesses and that combining them could result in a more accurate classification of core stability status.

While studies specifically aimed at classifying core stability status are relatively few, there have been efforts to develop machine learning models based on signals from IMUs or other body-worn sensors for clinical objectives. For instance, one model, trained using four reliable accelerometer-derived features, classified participants according to their susceptibility to falls with an accuracy of 74.4%.³⁸ Another model, trained with data acquired using a pressure platform and inertial sensor placed on the lower back, achieved ~73% when classifying patients according to fall history.³⁷ Another sensor-based machine learning model categorized nonspecific low back pain patients into high- and low-risk groups with 75% accuracy.³⁹

Our study, to our knowledge, is among the first to use machine learning to classify individuals based on core stability status, achieving a classification accuracy of ~85% using mediolateral head motion during RMs as the primary feature. This highlights the potential of our novel approach in leveraging head motion as a representative of core stability, despite the lack of inclusion of certain

Table 3. Accuracy, sensitivity, specificity, precision, F1-score, and G-index values of neural network and support vector machine models used to classify participants according to core stability status.

	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-index
SVM						
Full	86.5 ± 1.6%	73.1 ± 3.2%	100.0 ± 0.0%	100.0 ± 0.0%	84.4 ± 2.1	0.27 ± 0.03
FMTs	86.9 ± 1.7%	73.9 ± 3.4%	100.0 ± 0.0%	100.0 ± 0.0%	84.9 ± 2.3	0.26 ± 0.03
RMs	86.6 ± 1.7%	75.3 ± 3.4%	100.0 ± 0.0%	100.0 ± 0.0%	84.5 ± 2.2	0.27 ± 0.03
NN						
Full	71.9 ± 3.9%	81.8 ± 2.4%	62.0 ± 8.1%	68.6 ± 4.8%	78.0 ± 1.9	0.36 ± 0.04
FMTs	74.5 ± 2.5%	83.9 ± 4.7%	65.1 ± 6.3%	70.8 ± 3.2%	79.6 ± 2.8	0.34 ± 0.03
RMs	73.7 ± 3.6%	83.5 ± 4.9%	63.9 ± 5.3%	69.9 ± 3.3%	79.1 ± 3.5	0.35 ± 0.04

FMTs: functional movement tests; RMs: rhythmic movements; NN: neural network; SVM: support vector machine.

variables used in previous research. However, in the absence of directly comparable studies, further research is required to validate our approach.

It is worth noting that our approach offers advantages in terms of simplicity and user-friendliness, with our wireless earbud IMU sensor being less intrusive and easier to use than multiple sensors placed on the body. Previous models generally used greater than or equal to two IMU or body-worn sensors placed on the lower extremities or lower lumbar spine.^{39,45} To obtain accurate body kinematic data, the effect of clothes should be minimized; thus, participants are typically instructed to attach the sensors to bare skin or to wear tight clothing. In contrast, when measuring head motion via the IMU sensor embedded in our wireless earbud, there is no need to consider the effect of clothing. Moreover, the involvement of a highly trained operator for head motion measurement is unnecessary when using our earbud.

Automated assessment of core stability can help prevent low back pain. In previous studies, low back pain patients tended to have poor core stability.^{4,28,47,48} The assessment of core stability during daily activities (e.g. walking, running, and cycling) may facilitate prevention of low back pain in at-risk individuals.

In this study, we made the decision not to perform feature scaling prior to our cross-validation process. This decision was based on our objective to avoid potential data leakage, which can lead to overoptimistic estimates of model performance. Additionally, given the nature of our data and the robustness of the machine learning models used, we found that our models performed well even without feature scaling. This suggests that our models were not overly sensitive to the range of feature values, providing further assurance of their robustness.

However, we acknowledge that the decision to scale features or not may depend on the specific characteristics of the data set and the machine learning model used. Future research should take these factors into consideration when designing their methodology.

Despite the promising results, our study has several limitations. One of the main limitations relates to the sample size. As noted above, the performance of our NN model was inferior to that of the SVM model, which we attribute, at least in part, to our sample size. According to Alwosheel et al. (2018), the minimum sample size necessary for artificial NN models is 50-fold greater than the number of weights⁴⁹; however, in the present study, the number of samples was considerably below this threshold. The small sample size may have limited the accuracy of our NN model. Another potential limitation is our reliance on the SCST to determine core stability status. The SCST is often used to evaluate the ability to control the lumbopelvic segment^{50,51}; however, core stability is also determined by muscular capacity in the lumbopelvic-hip complex.^{2,52–54} We acknowledge that the SCST cannot evaluate all factors that influence core stability, particularly the endurance and strength of the lumbopelvic-hip complex. In future work, we aim to address these limitations by increasing the number of samples and developing a new and more comprehensive test for assessing core stability.

Conclusion

The present study investigated the capacity of an IMU sensor embedded within a wireless earbud to determine core stability status. Our SVM machine model classified participants into good core stability and poor core stability groups with an accuracy of >85%. The SVM model was

trained and validated using features extracted from medio-lateral head angle during RMs such as cycling, running, and walking, where this angle partly reflects reactive trunk motion. The present findings may facilitate the development of automated core stability assessment systems that allow users to easily determine their core stability status during activities such as cycling, running, and walking. Such systems could help to prevent core stability-related low back pain in at-risk individuals through the incorporation of core stability exercises into regular exercise programs prior to the onset of low back pain.

Acknowledgements: We thank, above all, our study participants for their time, dedication, and enthusiasm.

Contributorship: SJ, S-HK and K-NP: authoring of first draft of the manuscript, study design, manuscript editing, supervision, and project administration. SJ and K-NP: conceptualization, methodology, and writing. All authors read and approved the final manuscript.

Declaration of conflicting interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Research Foundation of Korea (NRF), (grant number No. 2020R1A2C2 102729).

Ethical approval: The ethics committee of the Jeonju University approved this study (number: jjIRB-220728-HR-2022-0302).

Guarantor: Kyue-Nam, Park.

ORCID iD: Kyue-Nam Park  <https://orcid.org/0000-0003-3521-3121>

References

1. Akuthota V, Ferreiro A, Moore T, et al. Core stability exercise principles. *Curr Sports Med Rep* 2008; 7: 39–44.
2. Kibler WB, Press J and Sciascia A. The role of core stability in athletic function. *Sports Med* 2006; 36: 189–198.
3. Leetun DT, Ireland ML, Willson JD, et al. Core stability measures as risk factors for lower extremity injury in athletes. *Med Sci Sports Exerc* 2004; 36: 926–934.
4. Aluko A, DeSouza L and Peacock J. The effect of core stability exercises on variations in acceleration of trunk movement, pain, and disability during an episode of acute nonspecific low back pain: a pilot clinical trial. *J Manipulative Physiol Ther* 2013; 36: 497–504.
5. Willson JD, Dougherty CP, Ireland ML, et al. Core stability and its relationship to lower extremity function and injury. *J Am Acad Orthop Surg* 2005; 13: 316–325.
6. Wirth K, Hartmann H, Mickel C, et al. Core stability in athletes: a critical analysis of current guidelines. *Sports Med* 2017; 47: 401–414.
7. Watanabe Y, Kato K, Otoshi K, et al. Associations between core stability and low back pain in high school baseball players: a cross-sectional study. *J Orthop Sci Epub* ahead of print 22 June 2021. DOI: 10.1016/J.JOS.2021.05.010
8. Hodges PW and Richardson CA. Inefficient muscular stabilization of the lumbar spine associated with low back pain. A motor control evaluation of transversus abdominis. *Spine (Phila Pa 1976)* 1996; 21: 2640–2650.
9. Shamsi M, Sarrafzadeh J, Jamshidi A, et al. The effect of core stability and general exercise on abdominal muscle thickness in non-specific chronic low back pain using ultrasound imaging. *Physiother Theory Pract* 2016; 32: 277–283.
10. Cosio-Lima LM, Reynolds KL, Winter C, et al. Effects of physioball and conventional floor exercises on early phase adaptations in back and abdominal core stability and balance in women. *The Journal of Strength & Conditioning Research* 2003; 17: 721–725.
11. Zheng Y, Ke S, Lin C, et al. Effect of core stability training monitored by rehabilitative ultrasound image and surface electromyogram in local core muscles of healthy people. *Pain Res Manag* 2019; 2019: 9130959.
12. Czaprowski D, Afeltowicz A, Gębicka A, et al. Abdominal muscle EMG-activity during bridge exercises on stable and unstable surfaces. *Phys Ther Sport* 2014; 15: 162–168.
13. Hibbs AE, Thompson KG, French DN, et al. Peak and average rectified EMG measures: which method of data reduction should be used for assessing core training exercises? *J Electromyogr Kinesiol* 2011; 21: 102–111.
14. Ahmed S, Saraswat A and Esht V. Correlation of core stability with balance, agility and upper limb power in badminton players: a cross-sectional study. *Sport Sci Health* 2022; 18: 165–169.
15. Bliss LS and Teeple P. Core stability: the centerpiece of any training program. *Curr Sports Med Rep* 2005; 4: 179–183.
16. Solana-Tramunt M, Ortigón A, Morales J, et al. Diagnostic accuracy of lumbopelvic motor control tests using pressure biofeedback unit in professional swimmers: a cross-sectional study. *J Orthop* 2019; 16: 590–595.
17. Haskins IN, Prabhu AS, Jensen KK, et al. Effect of transversus abdominis release on core stability: short-term results from a single institution. *Surgery* 2019; 165: 412–416.
18. Picerno P. 25 Years of lower limb joint kinematics by using inertial and magnetic sensors: a review of methodological approaches. *Gait Posture* 2017; 51: 239–246.
19. Favre J, Aissaoui R, Jolles BM, et al. Functional calibration procedure for 3D knee joint angle description using inertial sensors. *J Biomech* 2009; 42: 2330–2335.
20. Tao W, Liu T, Zheng R, et al. Gait analysis using wearable sensors. *Sensors* 2012; 12: 2255–2283.
21. Mancini M and Horak FB. The relevance of clinical balance assessment tools to differentiate balance deficits. *Eur J Phys Rehabil Med* 2010; 46: 239.
22. Papi E, Osei-Kuffour D, Chen Y-MA, et al. Use of wearable technology for performance assessment: a validation study. *Med Eng Phys* 2015; 37: 698–704.

23. Whelan DF, O'Reilly MA, Ward TE, et al. Technology in rehabilitation: evaluating the single leg squat exercise with wearable inertial measurement units. *Methods Inf Med* 2017; 56: 88–94.
24. Jensen U, Schmidt M, Hennig M, et al. An IMU-based mobile system for golf putt analysis. *Sports Eng* 2015; 18: 123–133.
25. Schmidt M, Rheinländer C, Nolte KF, et al. IMU-based determination of stance duration during sprinting. *Procedia Eng* 2016; 147: 747–752.
26. Hwang T-H and Effenberg AO. Head trajectory diagrams for gait symmetry analysis using a single head-worn IMU. *Sensors* 2021; 21: 6621.
27. Diao Y, Ma Y, Xu D, et al. A novel gait parameter estimation method for healthy adults and postoperative patients with an ear-worn sensor. *Physiol Meas* 2020; 41: 05NT01.
28. Hodges PW and Richardson CA. Delayed postural contraction of transversus abdominis in low back pain associated with movement of the lower limb. *J Spinal Disord* 1998; 11: 46–56.
29. Hodges PW and Richardson CA. Relationship between limb movement speed and associated contraction of the trunk muscles. *Ergonomics* Epub ahead of print 1997; 40. DOI: 10.1080/001401397187469.
30. Hodges P, Cresswell A and Thorstensson A. Preparatory trunk motion accompanies rapid upper limb movement. *Exp Brain Res* 1999; 124: 69–79.
31. Hodges PW and Richardson CA. Contraction of the abdominal muscles associated with movement of the lower limb. *Phys Ther* 1997; 77: 132–134.
32. Chiou SY, Hurry M, Reed T, et al. Cortical contributions to anticipatory postural adjustments in the trunk. *Journal of Physiology* Epub ahead of print 2018; 596. DOI: 10.1113/JP275312.
33. Kolář P, Šulc J, Kynčl M, et al. Postural function of the diaphragm in persons with and without chronic low back pain. *Journal of Orthopaedic and Sports Physical Therapy* Epub ahead of print 2012; 42. DOI: 10.2519/jospt.2012.3830
34. Cordo PJ and Nashner LM. Properties of postural adjustments associated with rapid arm movements. *J Neurophysiol* Epub ahead of print 1982; 47. DOI: 10.1152/jn.1982.47.2.287
35. Burnett AF, Cornelius MW, Dankaerts W, et al. Spinal kinematics and trunk muscle activity in cyclists: a comparison between healthy controls and non-specific chronic low back pain subjects - A pilot investigation. *Man Ther* Epub ahead of print 2004; 9. DOI: 10.1016/j.math.2004.06.002
36. Al-Eisa E, Egan D, Deluzio K, et al. Effects of pelvic asymmetry and low back pain on trunk kinematics during sitting: a comparison with standing. *Spine (Phila Pa 1976)* 2006; 31: E135–E143.
37. Greene BR, McGrath D, Walsh L, et al. Quantitative falls risk estimation through multi-sensor assessment of standing balance. *Physiol Meas* 2012; 33: 2049.
38. Doheny EP, Walsh C, Foran T, et al. Falls classification using tri-axial accelerometers during the five-times-sit-to-stand test. *Gait Posture* 2013; 38: 1021–1025.
39. Abdollahi M, Ashouri S, Abedi M, et al. Using a motion sensor to categorize nonspecific low back pain patients: a machine learning approach. *Sensors* 2020; 20: 3600.
40. Kianifar R, Lee A, Raina S, et al. Automated assessment of dynamic knee valgus and risk of knee injury during the single leg squat. *IEEE J Transl Eng Health Med* 2017; 5: 1–13.
41. Jung S and Park K-N. Is mediolateral head motion during the seated hip flexion test and indoor cycling test associated with core stability? *Journal of Musculoskeletal Science and Technology* 2022; 6: 64–70.
42. Kim AR, Park JH, Kim SH, et al. The validity of wireless earbud-type wearable sensors for head angle estimation and the relationships of head with trunk, pelvis, hip, and knee during workouts. *Sensors* Epub ahead of print 2022; 22. DOI: 10.3390/s22020597
43. Batista GE, Prati RC and Monard MC. A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter* 2004; 6: 20–29.
44. Ganyun LV, Haozhong C, Haibao Z, et al. Fault diagnosis of power transformer based on multi-layer SVM classifier. *Electr Power Syst Res* 2005; 74: 1–7.
45. Taborri J, Palermo E and Rossi S. Automatic detection of faults in race walking: a comparative analysis of machine-learning algorithms fed with inertial sensor data. *Sensors* 2019; 19: 1461.
46. Taborri J, Scalona E, Palermo E, et al. Validation of inter-subject training for hidden Markov models applied to gait phase detection in children with cerebral palsy. *Sensors* 2015; 15: 24514–24529.
47. Vasseljen O, Unsgaard-Tøndel M, Westad C, et al. Effect of core stability exercises on feed-forward activation of deep abdominal muscles in chronic low back pain: a randomized controlled trial.
48. Panjabi MM. Clinical spinal instability and low back pain. *J Electromyogr Kinesiol* 2003; 13: 371–379.
49. Alwosheel A, van Cranenburgh S and Chorus CG. Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling* 2018; 28: 167–182.
50. Aggarwal A, Kumar S, Madan R, et al. Relationship among different tests of evaluating low back core stability. *J Musculoskelet Res* 2011; 14: 1250004.
51. Chan EWM, Hamid MSA, Nadzalan AM, et al. Abdominal muscle activation: an EMG study of the Sahrman five-level core stability test. *Hong Kong Physiother J* 2020; 40: 89–97.
52. Waldhelm A and Li L. Endurance tests are the most reliable core stability related measurements. *J Sport Health Sci* 2012; 1: 121–128.
53. Behm DG, Leonard AM, Young WB, et al. Trunk muscle electromyographic activity with unstable and unilateral exercises. *J Strength Cond Res* 2005; 19: 193–201.
54. Hsu S-L, Oda H, Shirahata S, et al. Effects of core strength training on core stability. *J Phys Ther Sci* 2018; 30: 1014–1018.