

Original article

Machine learning for classifying chronic ankle instability based on ankle strength, range of motion, postural control and anatomical deformities in delivery service workers with a history of lateral ankle sprains

Ui-jae Hwang^{a,*}, Oh-yun Kwon^b, Jun-hee Kim^a, Gyeong-tae Gwak^a

^a Department of Physical Therapy, College of Health Science, Laboratory of KEMA AI Research (KAIR), Yonsei University, 234 Maeji-ri, Heungeop-Myeon, Wonju, Kangwon-Do, 220-710, South Korea

^b Department of Physical Therapy, College of Health Science, Laboratory of Kinetic Ergocise Based on Movement Analysis, Yonsei University, 234 Maeji-ri, Heungeop-Myeon, Wonju, Kangwon-Do, 220-710, South Korea

ARTICLE INFO

Keywords:

Ankle sprain
Ergonomics
Joint instability
Occupational health

ABSTRACT

Objective: Chronic ankle instability (CAI) frequently develops as a result of lateral ankle sprains (LAS) in delivery service workers (DSWs). Identifying risk factors for CAI is crucial for implementing targeted interventions. This study aimed to develop machine learning (ML) models for classifying CAI in DSWs with a history of LAS (DSWsLAS) and to identify key contributory factors.

Design: Exploratory, cross-sectional design.

Setting: and participants: A total of 121 DSWsLAS were screened for eligibility among 289 DSWs.

Methods: A total of 121 DSWsLAS were assessed for demographic characteristics, including ankle strength, range of motion, postural control, and anatomical deformities. Seven ML algorithms were trained and tested for classifying CAI. Principal component analysis (PCA) was used for feature extraction, and feature permutation importance (FPI) and Shapley additive explanations (SHAP) were employed to identify influential features.

Main outcome measures: Model performances were assessed using area under the curve (AUC). To interpret the classifications, we used FPI and SHAP values.

Results: PCA derived 7 principal components (PCs) accounting for 83.5% of the total variation in the data. The support vector machine (SVM) algorithm achieved the highest classifying performance (AUC = 0.817) among the ML models. FPI and SHAP revealed that PC1, PC2, PC5, and PC7 were the most influential features for classifying CAI in DSWsLAS.

Conclusions: The SVM algorithm, utilizing PCA-derived factors related to body mass index and ankle muscle strength demonstrated high classifying performance for diagnosis of CAI in DSWsLAS, emphasizing the importance of considering multiple contributory factors in the prevention and management of this condition.

1. Introduction

Chronic ankle instability (CAI) is a complex condition that often develops following a lateral ankle sprain. It is characterized by persistent symptoms such as pain, swelling, diminished self-reported function, and repetitive episodes or perceptions of the ankle giving way along with recurrent ankle sprains that continue for more than 12 months after the initial injury (Attenborough et al., 2014, Hertel and Corbett, 2019, Thompson et al., 2018). Delivery service workers (DSWs) are at high risk for ankle sprains due to the physically demanding nature of their work, which involves walking an average of 8 km per delivery shift while

carrying parcels of varying weights, sizes and shapes (Allen et al., 2018, Martinez-Sykora et al., 2020). Ankle sprains are the second most common work-related musculoskeletal disorder (González-Iñigo et al., 2017) and the ankle (23%) is the most frequently injured body part in DSWs (Bentley, 1998).

Ankle sprains lead to significant loss of work, with an average of 20 lost workdays per incident (González-Iñigo et al., 2017), and recurrent sprains resulting from CAI can further exacerbate productivity losses. Therefore, accurate classification of DSWs with and without CAI is crucial for effectively managing the condition and promoting timely return to work. Targeted interventions addressing the specific factors

* Corresponding author.

E-mail addresses: hwangu33@nate.com (U.-j. Hwang), kwonoy@yonsei.ac.kr (O.-y. Kwon), kemakjh@naver.com (J.-h. Kim), tytyte@naver.com (G.-t. Gwak).

<https://doi.org/10.1016/j.msksp.2024.103230>

Received 8 August 2024; Received in revised form 29 October 2024; Accepted 18 November 2024

Available online 19 November 2024

2468-7812/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

contributing to CAI are essential for facilitating recovery in affected workers. The functional insufficiencies that contribute to CAI following lateral ankle sprains arise from alterations in the motor component of sensorimotor control, including impaired postural control (Gribble et al., 2012, Linens et al., 2014, Wang et al., 2024), limited ankle range of motion (ROM) (Grindstaff et al., 2017, Terada et al., 2014), diminished ankle eversion strength (David et al., 2013, Thompson et al., 2018), anatomical deformities (Saki et al., 2021), and altered motor control of the proximal musculature surrounding the injured ankle (Son et al., 2019, Tavakoli et al., 2016).

Accurate classification of DSWs with CAI is crucial for implementing targeted interventions to mitigate the condition's impact. The recently updated CAI model (Hertel and Corbett, 2019) outlines a spectrum of pathomechanical, sensory-perceptual, and motor-behavioral impairments that contribute to the development of CAI. However, the precise contributions of these diverse factors to the development of CAI remain controversial among researchers and clinicians alike (Hertel and Corbett, 2019, Thompson et al., 2018), partly due to the complex interplay between mechanical and functional insufficiencies, which may occur independently or in combination and nonlinear, leading to heterogeneous presentations of CAI (Hertel and Corbett, 2019, Thompson et al., 2018). Traditional statistical methods may not adequately capture these complex, multifactorial relationships.

Machine learning (ML) techniques offer a promising approach for classifying CAI as they can model highly nonlinear relationships between a large number of variables and identify the most important contributory factors from a multidimensional perspective (Hwang and Gwak, 2023). ML algorithms learn from data to uncover intricate patterns and build classifying models that consider the combined influence of multiple factors (Hwang and Gwak, 2023). In the context of CAI, ML can potentially provide a more comprehensive understanding of how various pathomechanical, sensory-perceptual, and motor-behavioral impairments interact to influence the condition's development in

DSWs who had a history lateral ankle sprains (DSWsLAS).

However, to date, no studies have utilized ML to classify CAI in DSWsLAS or any other population. Therefore, the primary purpose of this study was to develop and compare ML models for classifying CAI in DSWsLAS, using measures of ankle strength, ROM, postural control and anatomical deformities. The secondary purpose was to identify the most important contributory factors from a multifactorial perspective to inform integrative preventative and rehabilitative strategies.

2. Materials and methods

2.1. Participants

DSW's data generated from musculoskeletal screening tests for preventing industrial accident were used by visiting a musculoskeletal health care center in delivery company from August 2021 to March 2022. Informed consent for the present study was waived by Institutional review board before the queries and analyses (1041849-202301-BM-016-01). A total of 121 DSWs with a history of at least one ankle sprain were screened for eligibility among 289 DSWs between August 2021 and March 2022 (Fig. 1.). CAI was defined as the occurrence of at least one ankle sprain resulting in lateral ankle pain, impaired physical activity, and at least one subsequent episode of 'giving way' or perceived ankle instability. Participants were confirmed to have recurrent bouts of instability by scoring at least 5 points on the Ankle Instability Instrument (Grindstaff et al., 2017). The Ankle Instability Instrument has been shown to be a reliable tool for identifying individuals with CAI (Docherty et al., 2006). Exclusion criteria included less than 6 months of work experience in parcel delivery service, a recent history (<6 months) of lower-extremity surgery, a diagnosis of ankle osteoarthritis, or a history of ankle surgery involving intra-articular fixation.

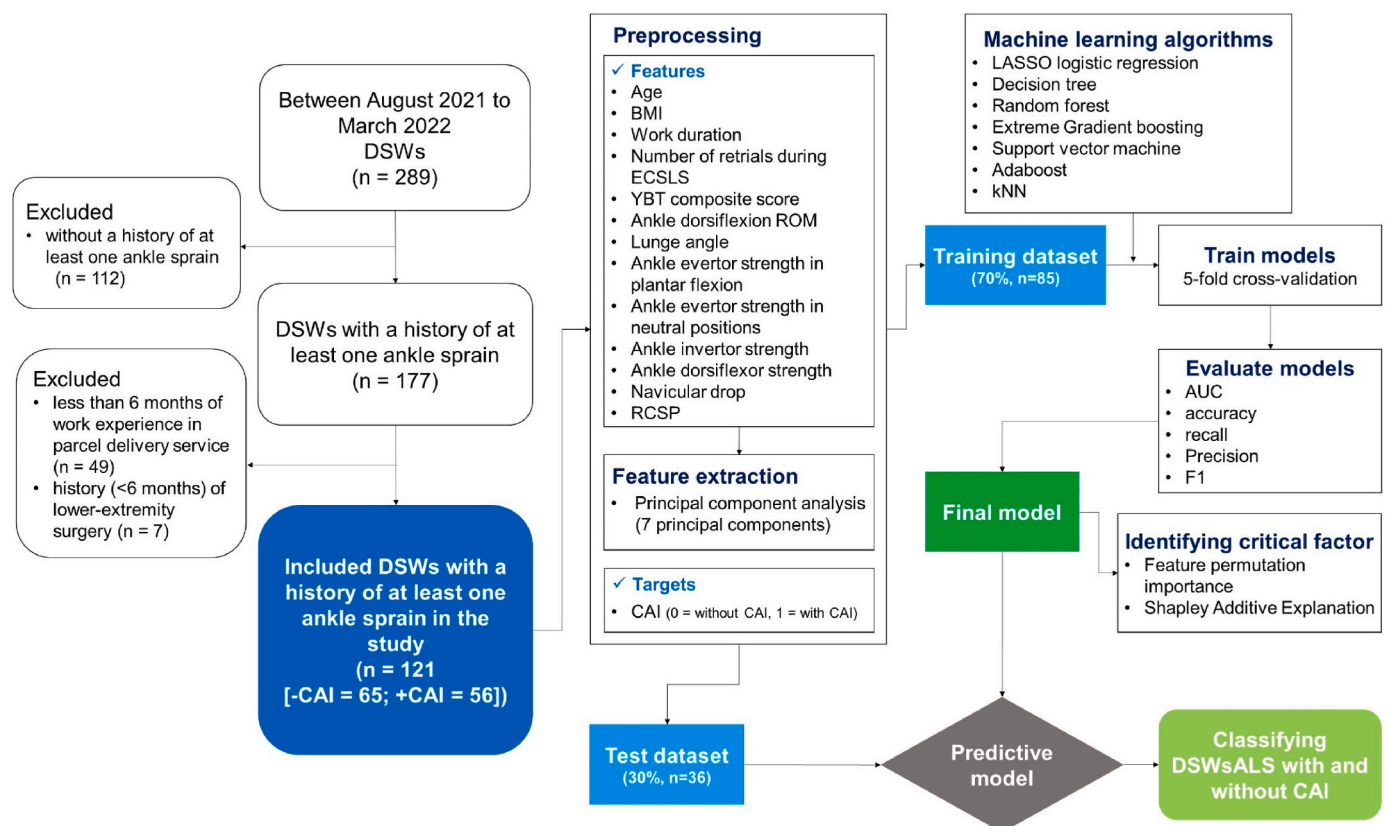


Fig. 1. Flowchart of participants selection and machine learning algorithm development.

2.2. Outcome measures

2.2.1. Postural control: Y-balance test and one leg standing with eye closed

To assess dynamic postural control, participants completed the Y-Balance Test (YBT), which involved reaching as far as possible with the non-stance limb in the anterior, posteromedial, and posterolateral directions (Linek et al., 2017). Participants performed six familiarization trials followed by two recorded trials, maintaining an upright posture with hands on their chest throughout the reaches (Linek et al., 2017). The YBT composite score was derived by summing the maximum reach distances in each direction and dividing by three times the limb length (Linek et al., 2017).

Static postural control was evaluated using the Eyes-Closed Single-Limb Stance (ECSLS) test, where participants stood on the test limb with hands on hips and eyes closed, aiming to maintain balance for 20 s (Doherty et al., 2014). If a loss of balance occurred, participants were instructed to quickly regain single-limb stance and continue the trial. During the ECSLS test, the number of times participants lost balance and had to retry the trial was counted (Doherty et al., 2014).

2.2.2. Ankle range of motions: ankle dorsiflexion range of motion and lunge angle

Ankle dorsiflexion ROM was assessed using an inclinometer while participants were seated with their knee extended and the distal portion of their leg unsupported off the edge of the examination table (Denegar et al., 2002). The inclinometer was positioned on the lateral aspect of the foot, and the examiner passively dorsiflexed the ankle until a firm end-feel was encountered, indicating the end of available ROM. To measure the weight-bearing lunge angle, an inclinometer was placed 15 cm distal to the tibial tuberosity (Hall and Docherty, 2017). Participants were instructed to lunge forward, aligning their second toe and heel on a straight line, while attempting to make contact between their knee and a vertical line on the wall. Throughout the lunge, participants were required to maintain heel contact with the ground. The lunge angle was recorded at the maximum anterior tibial advancement (Hall and Docherty, 2017).

2.2.3. Ankle joint muscles strength: evtor, invertor, and dorsiflexor

Ankle muscle strength was quantified using a strength sensor with the strap positioned at the distal metatarsals, allowing half of the foot to extend beyond the edge of the table. The opposite end of the sensor was secured to a floor-mounted absorber using a belt (Ahn et al., 2020). To assess evtor strength, participants were positioned in side-lying on the contralateral side with their hip and knee flexed to 90° (Ahn et al., 2020). Measurements were taken in both neutral (0°) and plantar flexed (50°) ankle positions. For invertor strength testing, participants were positioned in side-lying on the ipsilateral side with their hip and knee flexed to 90° and their ankle in a neutral position. Dorsiflexor strength was measured with participants seated on a table, hips and knees flexed to 90°, and the ankle in a neutral position.

2.2.4. Anatomical deformities: navicular drop and resting calcaneal stance position

Navicular drop was assessed with participants standing barefoot on the floor. The navicular tuberosity was identified and marked, and the subtalar neutral position was determined by maneuvering the foot into eversion and inversion until the talus was palpated to be congruent between the thumb and index finger (Zuil-Escobar et al., 2019). The vertical distance from the navicular tuberosity to the floor was measured in both subtalar neutral and relaxed standing postures. Resting calcaneal stance position (RCSP) was evaluated with participants lying prone on a plinth, with their feet extending just beyond the edge (Cho et al., 2019). A bisection line was drawn on the calcaneus using three reference points: the upper, middle, and lower aspects of the posterior calcaneus (Cho et al., 2019). RCSP was then measured in relaxed standing with feet shoulder-width apart, and the angle formed between the calcaneal

bisection line and the vertical was recorded, with calcaneal eversion and inversion represented by positive and negative values, respectively.

2.3. Experimental protocol

Participants initially completed a concise health history questionnaire and the Ankle Instability Instrument. The sequence of assessments was performed in the following order: anatomical deformities, postural control, ankle ROM, and ankle joint muscle strength. Prior to data collection, participants performed three practice trials for the ECSLS test to familiarize themselves with the testing procedure. For strength measurements, participants completed two submaximal practice contractions before the actual testing to ensure proper technique and familiarization with the testing position. The order of muscle strength tests was randomized, and a 3-min rest interval was provided between each strength measurement to minimize fatigue. All measurements were conducted by three experienced physical therapists (OYK, UJH and GTG) with more than 10 years of clinical experience in musculoskeletal assessment. Prior to data collection, the examiners underwent a standardization training program consisting of three days of practice sessions to ensure consistent measurement techniques. For DSWsLAS, regardless of the presence or absence of CAI, all evaluations were conducted on the affected limb. In cases where DSWsLAS exhibited bilateral CAI, the ankle with greater perceived instability was chosen for assessment.

2.4. Machine learning modeling

ML analysis was performed using the Orange data mining software (Orange 3.3.0, Ljubljana, Slovenia) and Python (Version 3.6.15; Python Software Foundation) with numpy (Version 1.19.5) for numerical computations and scikit-learn (Version 0.24.2) for the ML algorithms.

2.4.1. Pre-processing

Twelve features were included in the study. Exploratory data analysis was performed to identify missing values in the dataset. Upon further investigation, no missing data points were found that required imputation or removal of instances with unknown values. Given our relatively modest sample size ($n = 121$) and the importance of maintaining clinical relevance, we did not exclude any potential outliers. This decision was made to preserve all available clinical information and maintain the natural variability present in the patient population, as extreme values in clinical measures may represent meaningful physiological or functional variations rather than measurement errors.

Principal component analysis (PCA) was employed as a feature extraction technique to reduce the dimensionality of the dataset and identify underlying patterns among the variables (Gaonkar et al., 2017; Raihan-Al-Masud and Mondal, 2020). PCA was applied to examine the covariance and correlation matrices of the variables and to determine the number of principal components (PCs) needed to adequately represent the original set of twelve features. PCA is a dimensionality reduction technique that transforms potentially correlated variables into a smaller set of uncorrelated variables called PCs. This transformation maintains the maximum amount of information from the original variables while reducing redundancy. Each principal component represents a linear combination of the original variables, with the first component capturing the largest variance in the data, and subsequent components capturing progressively less variance. To ensure comparability and avoid scaling effects, the data were pre-processed by standardizing the variables prior to conducting PCA. The selection of PCs was based on the criterion that the cumulative proportion of explained variance should exceed 80% (Ge et al., 2023).

2.4.2. Machine learning algorithms

The complete dataset ($n = 121$) was split into a training set (70%, $n = 85$) for model development and a test set (30%, $n = 36$) for external

validation of classification performance (Fig. 1.). Seven ML algorithms, including LASSO logistic regression, decision tree, random forest, Extreme Gradient Boosting machine, support vector machine (SVM), AdaBoost, and k-nearest neighbors (kNN), were trained using 5-fold cross-validation (Fig. 1.). Hyperparameter tuning was performed using grid search on the training dataset. The optimal hyperparameters were determined after evaluating multiple combinations for each algorithm. For logistic regression, the optimal parameter was $C = 1$. The random forest model performed best with $\text{max_depth} = 20$ and $n_estimators = 200$. For gradient boosting, optimal parameters were $\text{learning_rate} = 1$, $\text{max_depth} = 5$, and $n_estimators = 50$. The SVM model achieved optimal performance with $C = 1$, $\text{gamma} = \text{'scale'}$, and $\text{kernel} = \text{'rbf'}$. The decision tree classifier performed best with $\text{criterion} = \text{'entropy'}$, $\text{max_depth} = \text{None}$, $\text{min_samples_leaf} = 2$, and $\text{min_samples_split} = 5$. For kNN, optimal parameters were $\text{algorithm} = \text{'auto'}$, $\text{metric} = \text{'euclidean'}$, $n_neighbors = 9$, and $\text{weights} = \text{'distance'}$. The AdaBoost classifier achieved best performance with $\text{algorithm} = \text{'SAMME'}$, a decision tree base estimator ($\text{max_depth} = 2$), $\text{learning_rate} = 1.0$, and $n_estimators = 50$.

2.4.3. Model validation

The primary metric used to evaluate model performance was the area under the receiver operating characteristic curve (AUC), which was calculated for both positive (CAI) and negative (non-CAI) classes and averaged to provide a single performance metric. This binary classification approach ensures balanced evaluation of the model's ability to correctly identify both CAI and non-CAI cases. Secondary performance measures included precision, sensitivity (recall), specificity and the F1 score. For calculating precision, recall, and F1 scores, we used a classification threshold of 0.5, which is the standard threshold that balances between sensitivity and specificity. The F1 score represents the harmonic mean of precision and recall, providing a single metric that considers both false positives and false negatives. This is particularly useful in our case where we need to balance the correct identification of both CAI and non-CAI cases. The classifying performance of the models was categorized based on the AUC value as follows: excellent (≥ 0.9), good (0.8–0.9), fair (0.7–0.8), and poor (< 0.7) (Zou et al., 2007). To prevent overfitting, several measures were implemented: (1) strict separation of training and test sets before any model development, (2) use of 5-fold cross-validation during training, (3) implementation of early stopping in gradient boosting, and (4) regularization parameters in SVM and logistic regression. The apparent higher performance of some models on the test set, particularly SVM, can be attributed to the random sampling variation in our relatively modest sample size and the effective generalization achieved through cross-validation and hyperparameter tuning process.

Feature permutation importance and Shapley additive explanations (SHAP) were used to identify the most influential features for the model's classifications. Feature permutation importance involves randomly shuffling the values of each feature in the training data and measuring the resulting increase in the model's classification error, quantified by the change in AUC (Hwang and Gwak, 2023). SHAP visualizes the impact and directionality of each classifying variable, generating summary plots that display features sorted by their relative importance and their SHAP values, indicating the magnitude and direction of each feature's impact on the model's output (Hwang et al., 2024b). These techniques provide a comprehensive assessment of each feature's relative importance and directional influence on the model's classifications.

3. Results

3.1. Patients characteristics

Table 1 shows the average values and variability of all the variables. Our study evaluated 121 DSWsLAS with and without CAI and were

Table 1

Mean (standard deviation) of baseline characteristics in DSWsLAS with and without CAI.

	without CAI (n = 65)	without CAI (n = 56)	p
Age (yr)	35.68 ± 8.10	37.60 ± 5.90	0.135
Weight (kg)	74.85 ± 12.10	78.01 ± 15.30	0.215
Height (cm)	169.57 ± 5.60	175.08 ± 6.00	0.067
Body mass index (kg/m ²)	24.30 ± 3.61	25.82 ± 4.16	0.035
Work duration (day)	388.92 ± 224.70	401.84 ± 159.90	0.714
Ankle Instability Instrument	2.80 ± 1.20	6.20 ± 1.30	<0.001
Ankle evertor strength in neutral (kgf)	25.55 ± 8.36	21.73 ± 6.94	0.007
Ankle evertor strength in plantar flexion (kgf)	18.02 ± 6.11	15.20 ± 4.12	0.003
Ankle invertor strength (kgf)	29.78 ± 8.17	26.98 ± 6.85	0.043
Ankle dorsiflexor strength (kgf)	40.42 ± 13.73	39.46 ± 11.00	0.67
Navicular drop (cm)	0.63 ± 0.31	0.68 ± 0.49	0.514
Resting calcaneal stance position (°)	0.66 ± 3.11	2.34 ± 3.50	0.007
Y-balance test composite score (%)	102.13 ± 97.59	83.03 ± 13.71	0.124
Lunge angle (°)	49.82 ± 9.98	38.02 ± 8.30	0.045
Ankle dorsiflexion ROM (°)	5.05 ± 3.30	1.98 ± 3.10	<0.001
Number of retrials during eyes-closed single-limb stance (trial)	2.17 ± 2.60	1.93 ± 1.70	0.548

analyzed using ML models. There were no significant differences between DSWsLAS with and without CAI in terms of age, work duration, ankle dorsiflexor strength, navicular drop, YBT composite score and number of retrials during ECSLS. The body mass index (BMI), Ankle Instability Instrument score and RCSP in DSWsLAS with CAI were statistically higher than that in DSWsLAS without CAI. The ankle evertor strength in neutral and plantar flexion position, ankle invertor strength, YBT composite score, lunge angle, ankle dorsiflexion ROM in DSWsLAS with CAI were statistically lower than that in DSWsLAS without CAI. All the DSWsLAS who participated in the present study were male.

3.2. Principal component analysis

PCA derived 7 PCs accounting for 83.5% of the total variation in the data (Table 2). These PCs can be summarised as representing the BMI, ankle evertor strength in neutral and plantar flexion, ankle invertor and dorsiflexor strength (PC1), ankle dorsiflexion ROM and RCSP (PC2), navicular drop and number of retrials during ECSLS (PC3), RCSP and number of retrials during ECSLS (PC4), age and work duration (PC5), lunge angle and YBT composite score (PC6), and navicular drop and YBT composite score (PC7) (Table 2 and Fig. 2). Principal component scores were calculated for each participant using the component loadings.

3.3. Classifying models of machine learning

The LASSO logistic regression, random forest, Extreme Gradient Boosting machine, SVM, and kNN algorithms demonstrated fair classifying performance, with AUC values exceeding 0.7 using the training dataset. SVM algorithm model for classifying DSWsLAS with and without CAI performed in the highest AUC using the test dataset (AUC, 0.817 [good]; F1, 0.749) (Table 3).

For feature permutation importance, PC1, PC5 and PC7 were the top three features for classifying DSWsLAS with and without CAI in the SVM model (Fig. 3.). For SHAP, low PC1, low PC2 and low PC5 were the top three features for classifying DSWsLAS with and without CAI in the SVM model (Fig. 3.).

4. Discussion

This study aimed to develop and compare ML models for classifying CAI in DSWsLAS using measures of ankle strength, ROM, postural

Table 2
Principal components loading of 12 features.

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Age (yr)	−0.165	0.346	−0.307	0.236	−0.691	0.221	0.027
Body mass index (kg/m2)	−0.517	−0.318	−0.338	−0.419	0.210	0.142	0.165
Work duration (day)	0.217	−0.338	0.455	−0.152	−0.498	0.324	0.287
Ankle invertor strength (kgf)	0.842	0.060	−0.068	0.133	−0.083	0.071	−0.013
Ankle evertor strength in neutral (kgf)	0.852	−0.027	−0.194	−0.130	−0.036	0.058	0.024
Ankle evertor strength in plantar flexion (kgf)	0.782	−0.040	−0.147	−0.200	0.020	0.004	0.072
Ankle dorsiflexor strength (kgf)	0.711	−0.206	0.113	−0.161	0.060	−0.107	−0.022
Navicular drop (cm)	−0.062	0.265	−0.500	−0.227	−0.139	−0.392	0.602
Resting calcaneal stance position (°)	−0.282	−0.570	−0.121	0.491	−0.179	−0.062	0.049
Y-balance test composite score (%)	0.137	0.088	0.226	0.516	0.426	0.276	0.602
Lunge angle (°)	−0.077	0.395	−0.290	−0.163	0.165	0.723	−0.093
Ankle dorsiflexion ROM (°)	0.048	0.793	0.305	0.136	0.005	−0.241	−0.041
Number of retrials during eyes-closed single-limb stance (trial)	−0.333	0.143	0.599	−0.479	−0.122	0.025	0.178

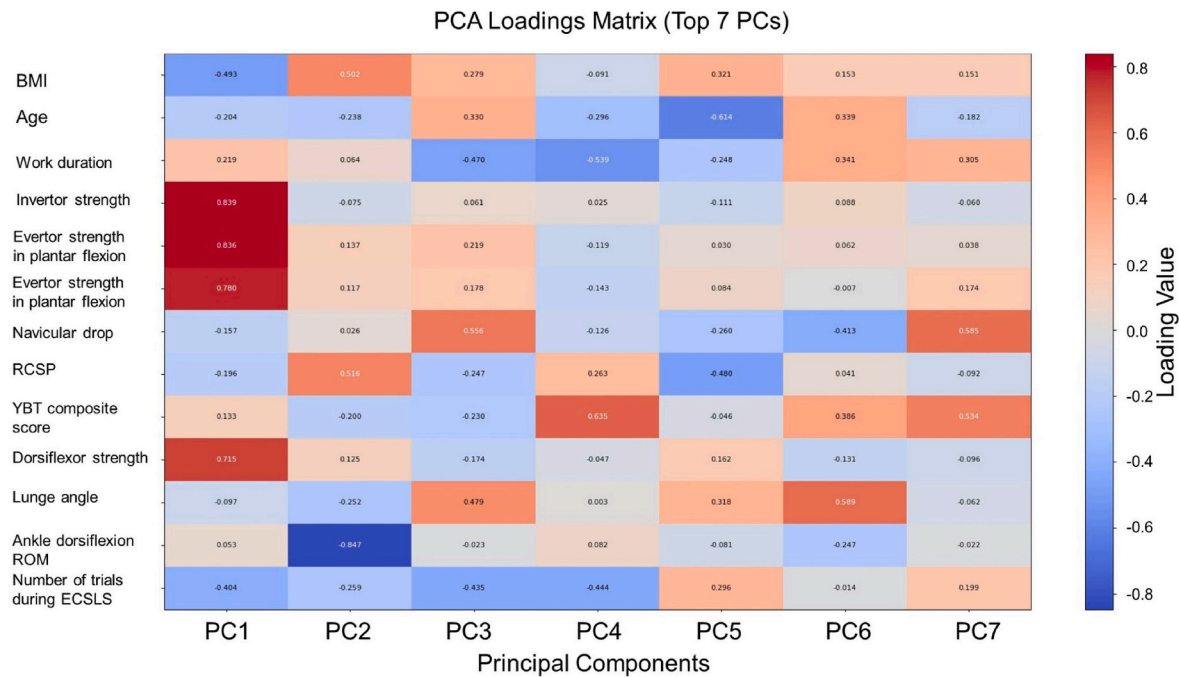


Fig. 2. Loading matrix in principal component analysis.

control, and anatomical deformities. Additionally, we sought to identify the most important contributory factors from a multifactorial perspective to inform integrative preventative and rehabilitative strategies. Our analysis provides actionable clinical insights despite the statistical complexity of PCs. The strong association of PC1 with CAI classification translates directly to clinical practice through its primary components: BMI and ankle evertor strength. SHAP analysis revealed that lower evertor strength (particularly in neutral and plantar flexed positions) and higher BMI were strongly associated with CAI presence in DSWsLAS. These findings suggest specific, measurable targets for clinical assessment and potential rehabilitation strategies. For instance, clinicians may prioritize evertor strengthening exercises and weight management strategies when treating DSWsLAS with ankle sprains.

Our findings demonstrate that the SVM algorithm achieved the highest classifying performance (AUC = 0.817, accuracy = 0.750, F1 = 0.749) among the seven ML models tested. Compared to previous studies using traditional statistical methods to classify CAI, our SVM model showed superior classifying performance. Combination of star excursion balance test reach distance and self-reported function could predict CAI with a sensitivity of 67% and specificity of 67% (de Noronha et al., 2013). Similarly, inability to perform a single-leg drop landing in the acute phase after ankle sprain was associated with CAI, with a sensitivity

of 58% and specificity of 80% (Doherty et al., 2016). Recent first-time ankle sprain (odds ratio (OR) = 8.23) and younger age (OR = 8.41) as independent predictors of recurrent ankle sprain, which is a key characteristic of CAI (Pourkazemi et al., 2018). In contrast, our SVM model achieved a higher specificity of 75% and sensitivity of 75%, highlighting the advantages of ML techniques for classifying CAI in DSWsLAS.

The high performance of the SVM model in our study compared to other algorithms can be attributed to several factors. First, SVM is well-suited to handle high-dimensional data, which is particularly relevant in our study, as we included a wide range of predictors spanning multiple domains (Cervantes et al., 2020). Second, SVM is effective in solving non-linear classification problems by mapping the input data into a higher-dimensional feature space, where a linear hyperplane can be used to separate the classes (Ghaddar and Naoum-Sawaya, 2018). This ability to capture complex, non-linear relationships between features and CAI status may have contributed to the SVM model's superior performance. Finally, SVM aims to find an optimal hyperplane that maximizes the margin between classes, making it robust to outliers and noise in the data (Ghaddar and Naoum-Sawaya, 2018). This property may have enhanced the generalizability of our SVM model to new, unseen data (Cervantes et al., 2020).

Feature permutation importance and SHAP revealed that the

Table 3

For classifying DSWsLAS with and without CAI, performance metrics of seven machine learning algorithms in the training and test dataset.

Performance metrics of seven machine learning algorithms in the training dataset						
Model	AUC	Acc	F1	Specificity	Sensitivity	Precision
AdaBoost	0.645	0.647	0.647	0.647	0.647	0.647
Extreme Gradient Boosting	0.747	0.682	0.683	0.682	0.682	0.683
kNN	0.722	0.706	0.705	0.692	0.706	0.72
LASSO	0.707	0.624	0.624	0.624	0.624	0.624
Logistic Regression						
Random Forest	0.721	0.682	0.68	0.682	0.682	0.682
Support Vector Machine	0.727	0.659	0.658	0.660	0.659	0.658
Decision Tree	0.647	0.600	0.600	0.600	0.600	0.600
Performance metrics of seven machine learning algorithms in the test dataset						
Model	AUC	Acc	F1	Specificity	Sensitivity	Precision
AdaBoost	0.525	0.528	0.527	0.530	0.528	0.526
Extreme Gradient Boosting	0.783	0.694	0.691	0.691	0.694	0.697
kNN	0.769	0.667	0.667	0.663	0.667	0.671
LASSO	0.731	0.639	0.635	0.639	0.639	0.639
Logistic Regression						
Random Forest	0.709	0.694	0.686	0.682	0.694	0.706
Support Vector Machine	0.817	0.750	0.749	0.750	0.750	0.750
Decision Tree	0.658	0.639	0.638	0.640	0.639	0.638

AUC, area under curve; Acc, Accuracy; kNN, k-nearest neighbors.

principal components (PC1, PC2, PC5, and PC7) related to BMI, ankle muscle strength, ankle dorsiflexion ROM, age, work duration, and anatomical deformities were the most influential features for classifying DSWsLAS with and without CAI. To provide clinically interpretable results, we examined the SHAP values of individual clinical measures within significant PCs. Within PC1, ankle evetor strength showed consistently negative SHAP values (indicating lower strength in CAI cases) while BMI showed positive values (indicating higher BMI in CAI cases). PC5's influence was primarily driven by age and work duration, with longer work duration associated with CAI classification. PC7's importance stemmed from the relationship between navicular drop and dynamic balance, where greater navicular drop combined with lower YBT scores was associated with CAI classification. These directional relationships provide specific guidance for clinical assessment. These findings highlight the multifactorial nature of CAI development and the importance of considering both modifiable and non-modifiable risk factors in the prevention and management of CAI. The identification of BMI and ankle muscle strength as key contributory factors aligns with previous research highlighting their role in the development of CAI (Fousekis et al., 2012, Hertel and Corbett, 2019, Thompson et al., 2018, Vuurberg et al., 2019). Our ML approach extends these findings by quantifying the relative importance of these factors in classifying CAI among DSWsLAS, emphasizing the need for targeted interventions to address BMI and ankle muscle strength deficits in this population. The identification of ankle dorsiflexion range of motion as an influential feature (PC2) aligns with previous research highlighting its role in the development of CAI. Individuals with CAI demonstrated reduced ankle dorsiflexion ROM compared to healthy controls, and this reduction was associated with impaired dynamic postural control(Terada et al., 2014). Similarly, ankle dorsiflexion range of motion influenced performance on the lateral step down test in individuals with CAI(Grindstaff et al.,

2017). Our findings extend this body of evidence by quantifying the relative importance of ankle dorsiflexion range of motion in classifying CAI among DSWsLAS, emphasizing the need for interventions targeting this modifiable risk factor. The importance of anatomical deformities, such as navicular drop and resting calcaneal stance position (PC7), in classifying CAI among DSWsLAS highlights the role of structural factors in the condition's development. Our cross-sectional analysis identified clinical features associated with CAI presence in DSWsLAS with prior ankle sprains. While causality cannot be established from this study design, these associations provide valuable insights for clinical assessment and rehabilitation planning. The identified features - particularly modifiable factors like ankle evetor strength and BMI - may serve as therapeutic targets, though longitudinal studies would be needed to establish their role in CAI development or prevention.

Notably, our study focused on a specific occupational group, DSWsLAS, which distinguishes it from previous research that primarily investigated CAI in athletes(Attenborough et al., 2014, Roos et al., 2017, Saki et al., 2021) (Roos et al., 2017; Attenborough et al., 2014; Saki et al., 2021).. The unique occupational demands faced by DSWs, such as prolonged walking and carrying loads of varying weights and sizes (Martinez-Sykora et al., 2020), may contribute to the development of CAI in this population(González-Iñigo et al., 2017). The physically demanding nature of delivery work may exacerbate the risk of ankle sprains and subsequent CAI, as well as influence the relative importance of specific risk factors identified in our study. The identification of work duration as an influential feature suggests that the cumulative exposure to occupational demands may play a more significant role in CAI development among DSWs compared to other populations (González-Iñigo et al., 2017). The repetitive stress and strain on the ankle joint associated with prolonged walking and load-carrying may increase the likelihood of recurrent sprains and CAI. Implementing workplace interventions to mitigate the risk of ankle sprains and promote effective rehabilitation is crucial for reducing the burden of CAI in this population.

While our previous study (Hwang et al., 2024a) investigated ankle kinematics to distinguish between healthy DSWs without LAS and those with CAI, the current study has a fundamentally different research objective and target population. Specifically, this study focuses on identifying factors that contribute to CAI development among DSWsLAS. Our findings provide novel insights into the pathomechanical cascade that may occur after an initial LAS, potentially leading to CAI development. The different target populations (healthy vs. CAI in the previous study; history of LAS with vs. without CAI in the current study) and different predictor variables (ankle kinematics vs. comprehensive clinical measures) make these studies complementary but distinct in their contributions to understanding CAI. While our previous work helps identify existing CAI in the general DSWs population, the current study provides crucial insights into the mechanisms and risk factors that may influence the progression from acute LAS to CAI, thereby informing both preventive strategies and early interventions for workers who have experienced an initial LAS.

This study has several limitations that should be acknowledged. First, while we used the validated Ankle Instability Instrument for CAI classification, our diagnostic approach did not include multiple self-reported questionnaires as recommended by the International Ankle Consortium (Gribble et al., 2013). Specifically, we did not utilize the Identification of Functional Ankle Instability questionnaire, which could have provided additional insights into the functional aspects of ankle instability. Future studies should consider incorporating multiple validated questionnaires to enhance the comprehensive assessment of CAI in this population. Second, the cross-sectional design precludes the establishment of causal relationships between the identified risk factors and CAI development. While our findings provide valuable insights into the associations between various impairments and CAI, prospective studies are needed to confirm the temporal and causal nature of these relationships. Third, our study population was limited to male DSWsLAS,

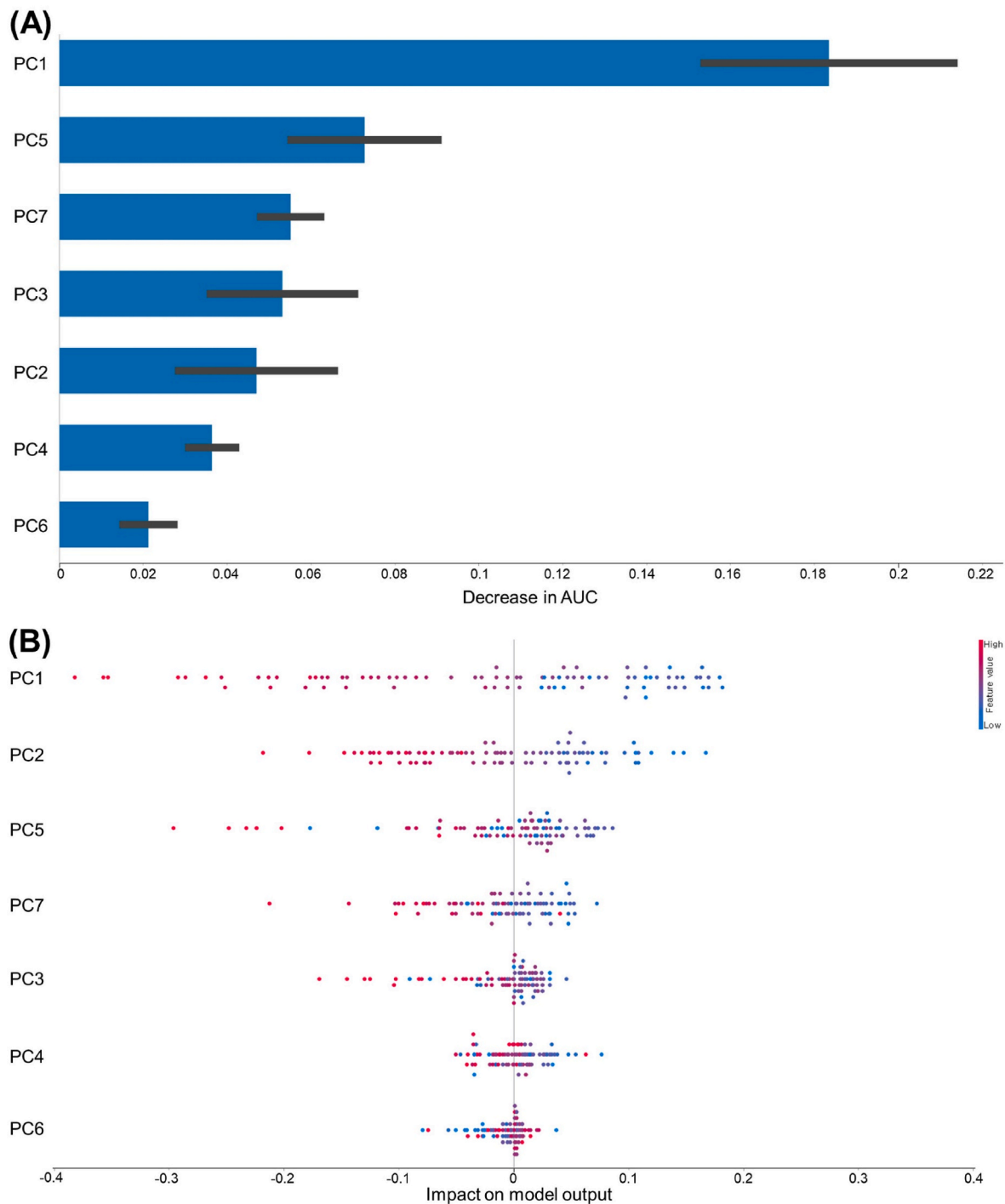


Fig. 3. A: Feature permutation importance of support vector machine model for classifying DSWsLAS with and without CAI; B: Shapley Additive Explanation analyses of support vector machine model for classifying DSWsLAS with and without CAI; C: Receiver operating characteristic curves of seven machine learning algorithms in the test dataset for for classifying DSWsLAS with and without CAI.

which may limit the generalizability of our results to other populations. Fourth, we did not analyze certain pathomechanical impairments, such as pathologic ligamentous laxity and sensory-perceptual impairments, which could potentially influence the development of CAI. Finally, while statistical tests exist for comparing different ML algorithms based on a single test dataset, a separate validation study using an independent sample would provide a more robust assessment of the generalizability and real-world performance of our models.

5. Conclusion

The present study demonstrates the utility of ML techniques, particularly the SVM algorithm, for classifying CAI in DSWsLAS and identifies key contributory factors from a multifactorial perspective. The high classifying performance of our model compared to previous studies highlights the potential of ML approaches to enhance CAI classification and management. BMI and ankle muscle strength emerged as the most influential features, underscoring the need for targeted, multifaceted

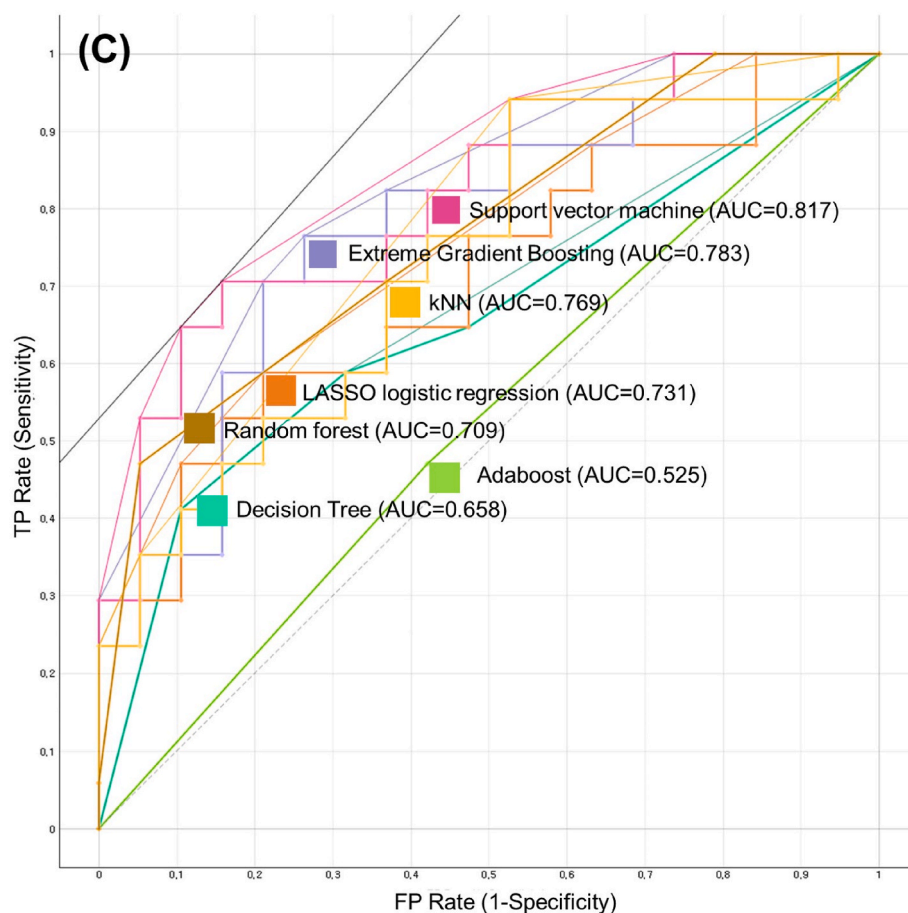


Fig. 3. (continued).

interventions to prevent and manage CAI in this high-risk occupational group. Healthcare professionals should consider these factors when designing comprehensive assessment and treatment strategies for DSWsLAS with CAI.

CRedit authorship contribution statement

Ui-jae Hwang: Writing – original draft, Methodology, Conceptualization. **Oh-yun Kwon:** Supervision, Project administration. **Jun-hee Kim:** Validation, Software, Data curation. **Gyeong-tae Gwak:** Formal analysis, Data curation.

Ethics approval statement

The present study conformed to the ethical guidelines of the 1975 Declarations of Helsinki. The requirement for informed consent was waived by Yonsei University Mirae Campus Institutional review board approval (1041849-202301-BM-016-01) because of the retrospective nature of the study. Ethical approval was obtained from Yonsei University Mirae Campus Institutional review board.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data statement

The datasets analyzed during the current study are available from the

corresponding author on reasonable request.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgments

We would like to thank all DSWs in our study for their active participation and cooperation.

References

- Ahn, S-h, Hwang, U.-j, Gwak, G-t, Yoo, H-i, Kwon, O-y, 2020. Comparison of the strength and electromyography of the evertor muscles with and without toe flexion in patients with chronic ankle instability. *Foot Ankle Int.* 41, 479–485.
- Allen, J., Pieczyk, M., Piotrowska, M., McLeod, F., Cherrett, T., Ghali, K., et al., 2018. Understanding the impact of e-commerce on last-mile light goods vehicle activity in urban areas: the case of London. *Transport. Res. Transport Environ.* 61, 325–338.
- Attenborough, A.S., Hiller, C.E., Smith, R.M., Stuelcken, M., Greene, A., Sinclair, P.J., 2014. Chronic ankle instability in sporting populations. *Sports Med.* 44, 1545–1556.
- Bentley, T.A., 1998. Slip, trip and fall accidents occurring during the delivery of mail. *Ergonomics* 41, 1859–1872.
- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., Lopez, A., 2020. A comprehensive survey on support vector machine classification: applications, challenges and trends. *Neurocomputing* 408, 189–215.
- Cho, Y., Park, J.-W., Nam, K., 2019. The relationship between foot posture index and resting calcaneal stance position in elementary school students. *Gait Posture* 74, 142–147.
- David, P., Halimi, M., Mora, I., Doutrelot, P.-L., Petitjean, M., 2013. Isokinetic testing of evertor and invertor muscles in patients with chronic ankle instability. *J. Appl. Biomech.* 29, 696–704.

- de Noronha, M., França, L.C., Haupenthal, A., Nunes, G., 2013. Intrinsic predictive factors for ankle sprain in active university students: a prospective study. *Scand. J. Med. Sci. Sports* 23, 541–547.
- Denegar, C.R., Hertel, J., Fonseca, J., 2002. The effect of lateral ankle sprain on dorsiflexion range of motion, posterior talar glide, and joint laxity. *J. Orthop. Sports Phys. Ther.* 32, 166–173.
- Docherty, C.L., Gansseder, B.M., Arnold, B.L., Hurwitz, S.R., 2006. Development and reliability of the ankle instability instrument. *J. Athl. Train.* 41, 154.
- Doherty, C., Bleakley, C., Hertel, J., Caulfield, B., Ryan, J., Delahunt, E., 2014. Balance failure in single limb stance due to ankle sprain injury: an analysis of center of pressure using the fractal dimension method. *Gait Posture* 40, 172–176.
- Doherty, C., Bleakley, C., Hertel, J., Caulfield, B., Ryan, J., Delahunt, E., 2016. Recovery from a first-time lateral ankle sprain and the predictors of chronic ankle instability: a prospective cohort analysis. *Am. J. Sports Med.* 44, 995–1003.
- Fousekis, K., Tsepis, E., Vagenas, G., 2012. Intrinsic risk factors of noncontact ankle sprains in soccer: a prospective study on 100 professional players. *Am. J. Sports Med.* 40, 1842–1850.
- Gaonkar, A.P., Kulkarni, R., Caytiles, R.D., Iyengar, N., 2017. Classification of lower back pain disorder using multiple machine learning techniques and identifying degree of importance of each parameter. *Int. J. Adv. Sci. Tech.* 105, 11–24.
- Ge, Y., Xu, B., Cai, H., Jing, W., Ouyang, Q., Yuan, Q., et al., 2023. Diagnostic role of plasma ORM2 in differentiating prostate cancer from benign prostatic hyperplasia. *J. Cancer Res. Clin. Oncol.* 149, 2301–2310.
- Ghaddar, B., Naoum-Sawaya, J., 2018. High dimensional data classification and feature selection using support vector machines. *Eur. J. Oper. Res.* 265, 993–1004.
- González-Inigo, S., Munuera-Martínez, P.V., Lafuente-Sotillos, G., Castillo-López, J.M., Ramos-Ortega, J., Domínguez-Maldonado, G., 2017. Ankle sprain as a work-related accident: status of proprioception after 2 weeks. *PeerJ* 5, e4163.
- Gribble, P.A., Delahunt, E., Bleakley, C., Caulfield, B., Docherty, C., Fourchet, F., et al., 2013. Selection criteria for patients with chronic ankle instability in controlled research: a position statement of the International Ankle Consortium. *J. Orthop. Sports Phys. Ther.* 43, 585–591.
- Gribble, P.A., Hertel, J., Plisky, P., 2012. Using the Star Excursion Balance Test to assess dynamic postural-control deficits and outcomes in lower extremity injury: a literature and systematic review. *J. Athl. Train.* 47, 339–357.
- Grindstaff, T.L., Dolan, N., Morton, S.K., 2017. Ankle dorsiflexion range of motion influences Lateral Step Down Test scores in individuals with chronic ankle instability. *Phys. Ther. Sport* 23, 75–81.
- Hall, E.A., Docherty, C.L., 2017. Validity of clinical outcome measures to evaluate ankle range of motion during the weight-bearing lunge test. *J. Sci. Med. Sport* 20, 618–621.
- Hertel, J., Corbett, R.O., 2019. An updated model of chronic ankle instability. *J. Athl. Train.* 54, 572–588.
- Hwang, U.-j., Gwak, G.-t., 2023. Machine learning vs. Logistic regression for classifying pressure pain hypersensitivity based on postural analysis data in food service workers with nonspecific neck/shoulder myofascial pain. *J. Musculoskeletal Sci. Tech.* 7, 71–79.
- Hwang, U.-j., Kwon, O.-y., Kim, J.-h., Gwak, G.-t., 2024a. Classification of chronic ankle instability using machine learning technique based on ankle kinematics during heel rise in delivery workers. *Digital Health* 10, 20552076241235116.
- Hwang, U.-j., Kwon, O.-y., Kim, J.-h., Yang, S., 2024b. Machine learning models for classifying non-specific neck pain using craniocervical posture and movement. *Musculoskeletal Sci. Practice* 71, 102945.
- Linek, P., Sikora, D., Wolny, T., Saulicz, E., 2017. Reliability and number of trials of Y Balance Test in adolescent athletes. *Musculoskeletal Sci. Practice* 31, 72–75.
- Linens, S.W., Ross, S.E., Arnold, B.L., Gayle, R., Pidcoe, P., 2014. Postural-stability tests that identify individuals with chronic ankle instability. *J. Athl. Train.* 49, 15–23.
- Martínez-Sykora, A., McLeod, F., Lamas-Fernandez, C., Bektaş, T., Cherrett, T., Allen, J., 2020. Optimised solutions to the last-mile delivery problem in London using a combination of walking and driving. *Ann. Oper. Res.* 295, 645–693.
- Pourkazemi, F., Hiller, C.E., Raymond, J., Black, D., Nightingale, E.J., Refshauge, K.M., 2018. Predictors of recurrent sprains after an index lateral ankle sprain: a longitudinal study. *Physiotherapy* 104, 430–437.
- Raihan-Al-Masud, M., Mondal, M.R.H., 2020. Data-driven diagnosis of spinal abnormalities using feature selection and machine learning algorithms. *PLoS One* 15, e0228422.
- Roos, K.G., Kerr, Z.Y., Mauntel, T.C., Djoko, A., Dompier, T.P., Wikstrom, E.A., 2017. The epidemiology of lateral ligament complex ankle sprains in National Collegiate Athletic Association sports. *Am. J. Sports Med.* 45, 201–209.
- Saki, F., Yalfani, A., Fousekis, K., Sodejani, S.H., Ramezani, F., 2021. Anatomical risk factors of lateral ankle sprain in adolescent athletes: a prospective cohort study. *Phys. Ther. Sport* 48, 26–34.
- Son, S.J., Kim, H., Seeley, M.K., Hopkins, J.T., 2019. Altered walking neuromechanics in patients with chronic ankle instability. *J. Athl. Train.* 54, 684–697.
- Tavakoli, S., Forghany, S., Nester, C., 2016. The effect of dual tasking on foot kinematics in people with functional ankle instability. *Gait Posture* 49, 364–370.
- Terada, M., Harkey, M.S., Wells, A.M., Pietrosimone, B.G., Gribble, P.A., 2014. The influence of ankle dorsiflexion and self-reported patient outcomes on dynamic postural control in participants with chronic ankle instability. *Gait Posture* 40, 193–197.
- Thompson, C., Schabrun, S., Romero, R., Bialocerkowski, A., van Dieen, J., Marshall, P., 2018. Factors contributing to chronic ankle instability: a systematic review and meta-analysis of systematic reviews. *Sports Med.* 48, 189–205.
- Vuurberg, G., Alting, N., Rajai, M., Blankevoort, L., Kerkhoffs, G.M., 2019. Weight, BMI and stability are risk factors associated with lateral ankle sprains and chronic ankle instability: a meta-analysis. *J. SAKOS* 4, 313–327.
- Wang, X., Wang, Z., Adams, R., Ganderton, C., Lyu, J., Han, J., 2024. Ankle inversion proprioception measured during stair descent can identify chronic ankle instability. *Musculoskeletal Sci. Practice* 72, 102958.
- Zou, K.H., O'Malley, A.J., Mauri, L., 2007. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* 115, 654–657.
- Zuil-Escobar, J.C., Martínez-Cepa, C.B., Martín-Urrialde, J.A., Gómez-Conesa, A., 2019. Evaluating the medial longitudinal arch of the foot: correlations, reliability, and accuracy in people with a low arch. *Phys. Ther.* 99, 364–372.