ISSN: 2476-6909; Modares. Mech. Eng. *2026;26(12):949-958.*

Modares Mechanical Engineering

DOI: 10.48311/mme.2026.119233.82975
journal homepage: mme.modares.ac.ir

Assessing Postural Instability in Individuals with Low Back Pain Through Measurement of Center of Pressure and Sport Intensity: A Machine Learning Approach

Pourya Dorostkar Rouhani, Mahmood-Reza Azghani*, Amin Partovifard, Mahrad Pouryosef Miandoab

Department of Biomedical Engineering, Sahand University of Technology, Tabriz, Iran

ARTICLE INFO

Article Type
Original Research

Article History
Received: February 24, 2026
Revised: June 21, 2026
Accepted: June 23, 2026
ePublished: September 17, 2026

ABSTRACT

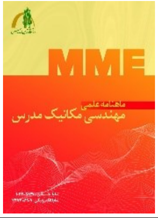
Falling is a major public health concern, especially for those dealing with low back pain, as it can lead to severe physical and psychological setbacks. While traditional clinical tests often miss the subtle nuances of balance impairment, quantitative Center of Pressure metrics offer a much clearer picture of stability. By combining these COP features with the right exercise intensity and advanced machine learning, we can create a more powerful tool for predicting and analyzing balance issues in patients with back pain.

In this study, we collected Center of Pressure data and extracted the key features defining postural stability. Following rigorous preprocessing and normalization, we evaluated several machine learning models specifically Logistic Regression, Random Forest, and Support Vector Machines to classify individuals with stable versus impaired balance. The results demonstrated that these machine learning architectures performed exceptionally well in distinguishing between the two groups. Among the features extracted from the Center of Pressure data, velocity metrics and displacement amplitude emerged as the most potent predictors of poor balance. Through stratified cross-validation, our models achieved high accuracy and stability, providing an effective and reliable framework for identifying balance deficits

Keywords: Center of Pressure, Physical Exertion, Low Back Pain, Neural Networks, Computer, Postural Balance

How to cite this article
Dorostkar Rouhani P, Azghani M.R, Partovifard A, Pouryosef Miandoab M, Assessing Postural Instability in Individuals with Low Back Pain Through Measurement of Center of Pressure and Sport Intensity: A Machine Learning Approach. Modares Mechanical Engineering; *2026;26(12):949-958.*

*Corresponding author's email: azghani@sut.ac.ir
*Corresponding ORCID ID: 0000-0001-9380-2675



شاپا: ۶۹۰۹-۲۴۷۶؛ مهندسی مکانیک مدرس. ۹۵۸-۹۴۹:(۱۲)۲۶؛۱۴۰۵

مهندسی مکانیک مدرس

DOI: 10.48311/mme.2026.119233.82975

صفحه اصلی مجله: mme.modares.ac.ir

ارزیابی ناپایداری ایستایی در افراد کمردردی از طریق اندازه‌گیری مرکز فشار و شدت تمرین‌های ورزشی: یک رویکرد یادگیری ماشین

پوریا درستکارروحانی، محمودرضا آذغانی*، امین پرتوی فرد، مهراد پوریوسف میاندوآب

گروه مهندسی پزشکی، دانشگاه صنعتی سهند، تبریز، ایران

اطلاعات مقاله

نوع مقاله

مقاله پژوهشی

تاریخچه مقاله

دریافت: ۱۴۰۴/۱۲/۰۵

بازنگری: ۱۴۰۵/۰۳/۳۱

پذیرش: ۱۴۰۵/۰۴/۰۲

ارائه آنلاین: ۱۴۰۵/۰۶/۲۶

چکیده

سقوط یکی از نگرانی‌های اصلی در حوزه سلامت عمومی است، به‌ویژه برای افرادی که با کمردرد دست‌وپنجه نرم می‌کنند؛ چرا که این عارضه می‌تواند منجر به پیامدهای ناگوار جسمی شود. در حالی که تست‌های بالینی سنتی اغلب در تشخیص اختلال تعادل ناتوان هستند، شاخص‌های کمی مرکز فشار تصویر بسیار دقیق‌تری از وضعیت پایداری ارائه می‌دهند. ما می‌توانیم با ترکیب این ویژگی‌های مرکز فشار با شدت تمرینات ورزشی و بهره‌گیری از یادگیری ماشین پیشرفته، ابزاری قدرتمند برای پیش‌بینی و تحلیل مشکلات تعادلی در بیماران مبتلا به کمردرد ایجاد کنیم.

در این پژوهش، داده‌های مرکز فشار جمع‌آوری و ویژگی‌های کلیدی معرف پایداری قامت استخراج گردید. پس از انجام مراحل دقیق پیش‌پردازش و نرمال‌سازی، چندین مدل یادگیری ماشین از جمله رگرسیون لجستیک، جنگل تصادفی و ماشین بردار پشتیبان را برای طبقه‌بندی افراد در دو گروه تعادل پایدار و تعادل ضعیف مورد ارزیابی قرار دادیم.

نتایج نشان داد که این ساختارهای یادگیری ماشین در تفکیک دو گروه عملکرد فوق‌العاده‌ای دارند. در میان ویژگی‌های استخراج‌شده از داده‌های مرکز فشار، شاخص‌های سرعت و دامنه جابه‌جایی به عنوان قوی‌ترین پیش‌بین‌های تعادل ضعیف شناخته شدند. مدل‌های ما از طریق اعتبارسنجی متقاطع لایه‌ای، به دقت و پایداری بالایی دست یافتند که چارچوبی موثر و قابل‌اعتماد را برای شناسایی نقص‌های تعادلی فراهم می‌کند.

کلیدواژه‌ها: مرکز فشار، اجهاد جسمانی، کمردرد، شبکه‌های عصبی، رایانه‌ای، تعادل ایستایی

نحوه ارجاع به این مقاله

درستکارروحانی پوریا، آذغانی محمودرضا، پرتوی فرد امین، پوریوسف میاندوآب مهراد، ارزیابی ناپایداری ایستایی در افراد کمردردی از طریق اندازه‌گیری مرکز فشار و شدت تمرین‌های ورزشی: یک رویکرد یادگیری ماشین، مهندسی مکانیک مدرس. ۹۵۸-۹۴۹:(۱۲)۲۶؛۱۴۰۵

*پست الکترونیکی نویسنده عهده‌دار مکاتبات: azghani@sut.ac.ir

*شناسه ارکید نویسنده عهده‌دار مکاتبات: ۲۶۷۵-۹۳۸۰-۰۰۰۱-۰۰۰۰

1- Introduction

Few public health challenges are as significant as the risk of falling, particularly for individuals living with chronic low back pain. While falls among older populations remain a concern for those with LBP, these incidents often serve as a catalyst for serious injury, a loss of confidence, and a permanent limitation of mobility both in daily life and athletic performance. This vulnerability is rooted in the deterioration of postural control, as LBP (Low back pain) fundamentally alters the body's sensory and motor systems, leading to neuromuscular shifts that degrade stability. In response to these risks, a growing body of evidence highlights the power of regular physical activity and targeted exercise as a primary defense. By building muscular strength and sharpening postural control [1], these programs can significantly lower the risk of falls [2].

Crucially, it isn't just about moving the type and intensity of that activity are what truly dictate the quality of balance. Different intensity levels don't just affect the body they leave distinct imprints on balance indices and overall postural stability [3–5]. High intensity training particularly in dynamic sports like soccer or gymnastics, has been shown to refine neuromuscular efficiency reduce COP (Center of pressure) path length, and tighten mediolateral sway range essentially teaching the CNS (Central Nervous System) to use more precise, energy efficient stabilization strategies [6]. In contrast, chronic participation in predominantly static, low intensity activities does not produce the same postural adaptations, the sensorimotor demands simply aren't high enough to drive significant recalibration of the postural control system. What makes this especially relevant for our population is that sport type shapes which COP parameters are affected and in which direction. Athletes from high demand dynamic sports tend to show lower COP displacement amplitude but higher mean velocity and frequency reflecting tighter, faster corrections rather than wide, slow sways [7]. This means that when LBP disrupts the sensorimotor loop the signature it leaves on COP data will look different depending on a person's athletic history. A patient who spent years in high-intensity sport has a different neuromuscular baseline than a sedentary individual and that difference matters when trying to classify impairment. Understanding how an individual's specific history of sport intensity interacts with their LBP to shape COP behavior is therefore critical, yet largely underexplored. While we often rely on the Berg Balance Scale and the TUG test for their accessibility, they offer a somewhat incomplete picture. These tests miss the nuances of a patient's stability and aren't sensitive enough to detect minor balance shifts meaning many at risk individuals go unnoticed because they can pass these basic clinical hurdles [4]. Force platforms, by contrast, enable precise quantification of objective parameters such as COP displacement velocity, and sway amplitude, offering far higher accuracy and sensitivity [8].

Recent studies have shown that portable platforms demonstrate good to excellent reliability in measuring COP velocity, although amplitude based indices may show lower stability under certain conditions. COP remains one of the most fundamental indicators in the analysis of postural control, and numerous parameters can be derived from it. Beyond LBP related factors, both the intensity and type of physical activity influence COP characteristics underscoring the importance of examining the relationship between exercise history and postural control [5]. Common COP derived metrics displacement velocity, movement amplitude dominant frequency, and sway area each reflect distinct motor strategies involved in maintaining balance [9]. While previous research has shown that these indices differentiate between groups, most of this work has been limited to classical statistical analyses like t-tests and ANOVA, with fewer studies moving toward machine learning and individualized predictive modeling. Thus despite the substantial value of COP indices, their optimal utilization requires more advanced data analytic techniques [9,10].

Recent advances in data analytics and machine learning have demonstrated that integrating COP features with ML models can improve fall risk prediction [11], using force platform time series features combined with deep neural networks, identified fall risk levels across age groups with high accuracy. Similarly [12] showed that COP paths recorded via inertial sensors could be predicted with low error suggesting a portable alternative to traditional force plates. However, a critical gap remains: no existing model has specifically targeted balance impairment in individuals with LBP by combining comprehensive COP metrics with their documented history of sport intensity. Therefore, the aim of this study was to classify postural instability in individuals with low back pain using COP derived features and exercise intensity data via logistic regression random forest, and support vector machine models, and to identify the strongest predictors of impaired balance.

2- Methods

2-1- Data Collection

This study was conducted under a larger project, named Extraction of Anthropometric Data of the Iranian Foot, with ethical clearance granted by the institutional review board IR.TBZMED.REC.1402.986. We took a cross sectional design, focusing on a diverse population within East Azerbaijan Province, Iran. We recruited a broad age range of participants 20 to 75 years from the general public, ensuring a representative sample regardless of social, economic, or cognitive background. After obtaining written informed consent, we collected detailed demographics and health histories. This included recording musculoskeletal issues such as low back pain, osteoarthritis, or history of lower-limb surgery as well as assessing their lifetime engagement in sports and current activity levels. To ensure precision, height and weight were directly measured using a calibrated stadiometer and a high precision digital scale. This comprehensive data collection spanned 16 months, reaching participants in various community settings, from recreational centers to commercial hubs.

2-2- Data Preprocessing

After completing all assessments, individuals who reported experiencing low back pain for more than six months were identified and separated for further analysis.

COP parameters were extracted using a custom-written code developed in MATLAB v21, based on the mathematical formulation of each parameter and got normalized to each foot dimension [1]. This procedure enabled precise quantification of postural control indices.

A detailed list of all extracted COP variables, along with their corresponding equations and database abbreviations, is presented in **Error! Reference source not found.**1 in Appendix.

2-3- Categorizing Sports Activities

When we look at different sports, we can categorize them in several ways whether by the type of activity, how intense they are, or even the risk of injury they carry. Traditionally, we split these activities into two main buckets: static and dynamic.

In dynamic sports, your muscles are constantly lengthening and shortening, driving joint movement. While this keeps you moving, it actually generates relatively low muscular force. Static activities are the opposite; they involve high force contractions where the muscles stay under tension without much change in length or joint position.

However, the reality of sports is rarely that "black and white." Most activities are a hybrid of both. Think of long-distance running: it's primarily dynamic, yet there's a small, constant static element involved in maintaining posture. On the other hand, something like water skiing leans heavily into static strength with only a minor dynamic component. Because of this overlap, it's more accurate to view sports on a spectrum [2]. By looking at the intensity of both static and dynamic demands, we can classify sports into the nine distinct groups shown in

Table **1**.

Table 1 Categorizing Sports by their Static and Dynamic Demands

Low Static, Moderate Dynamic	**Low Static, Low Dynamic**
Moderate Static, Low Dynamic	Low Static, High Dynamic
Moderate Static, High Dynamic	Moderate Static, Moderate Dynamic
High Static, Moderate Dynamic	High Static, Low Dynamic
	High Static, High Dynamic

2-4- Data Analysis and Neural Networks

Since this study focused exclusively on individuals with low back pain without a separate healthy control group we established a benchmark for "poor balance" using a distribution-based approach [3]. Specifically, we categorized participants in the upper 25th percentile of Sway Area as the weak balance group. Once the target labels were set, we began preparing the data for the modeling phase. All quantitative variables were normalized using z-score standardization [9]. To ensure our model evaluations were both stable and reliable, we employed a 5 fold stratified cross validation strategy.

A critical step in our pipeline was the prevention of data leakage. To achieve this, the Standard Scaler was integrated directly into the cross validation loops; the scaler was "fit" only on the training folds and then "transformed" onto the test folds in each cycle. The entire experimental framework was implemented step-by-step using Python 3.10.

2-4- Logistic Regression Modeling

We employed a Logistic Regression model, a probabilistic linear framework that estimates the likelihood of a participant belonging to the weak balance group on a scale from 0 to 1. Central to this model is the Logit function, defined for the dependent variable as shown in equation 1.

Equation 1 p is the probability of weak equilibrium and X_i are the input features. The coefficients β_i were estimated using Maximum Likelihood Estimation.

$$logit(p) = ln\left(\frac{p}{p-1}\right) = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \cdots + X_n\beta_n \quad (1)$$

To ensure a robust evaluation, the model was trained independently across each of the five folds, with performance measured against the respective test sets. We took strict precautions against data leakage by calculating normalization parameters solely on the training data and then applying that same transformation to the test data. For every fold, we tracked a comprehensive suite of metrics like accuracy, precision, recall, specificity, F1-score, and ROC-AUC. The final results are presented as the mean and standard deviation across all folds.

Once the cross-validation was complete, we averaged the standardized coefficients to quantify the impact of each feature on balance impairment. We also calculated Odds Ratios for each variable; features with an OR greater than 1 were identified as risk factors, while those below 1 were interpreted as protective factors. To visualize these relationships, we generated a bar chart illustrating the regression coefficients for the top fifteen features.

2-5- Random Forest Modeling

To compare our linear results with a non-linear model capable of capturing complex feature interactions, we utilized the Random Forest algorithm. This ensemble method is particularly robust in identifying non-linear patterns and handling noise within biomechanical datasets. The model was constructed using 200 decision trees. In this architecture, each tree is trained on a random subset of data and features, with the final classification determined by a majority vote across the ensemble.

The primary hyperparameters were configured as follows: n estimators = 200, max depth = None allowing trees to grow until all nodes are pure, and random state = 42 to ensure the reproducibility of our results. Unlike the logistic regression model, Random Forest was trained on each of the five folds without the need for prior data normalization.

For each fold, we calculated the same performance metrics used for the regression model: Accuracy, Precision, Recall, Specificity, F1-score, and ROC-AUC. To visualize the model's ability to distinguish between the two classes, we plotted the ROC curve for the first fold.

Feature importance was determined based on the Mean Decrease in Gini Impurity. We extracted the mean and standard deviation of these importance scores across all folds, visualizing the top fifteen features in a bar chart. For a deeper analysis, PDP (Partial Dependence Plots) were generated for the three most influential features, illustrating the specific relationship between feature variations and the probability of impaired balance.

To rigorously check for data leakage or overfitting, we performed a y-randomization permutation test. The output labels were randomly shuffled, and the models were retrained. The significant drop in performance metrics during this test confirmed the validity of our original model and ensured no "shortcuts" or leakage existed between the features and the target labels. Finally, the mean and standard deviation of performance metrics across all five iterations were compared to identify the superior model in terms of accuracy, sensitivity, specificity, and generalizability.

2-6- Artificial Neural Networks and UMAP

The dataset included 63 features extracted from Center of Pressure signals and one feature representing exercise intensity. Data were initially imported from Excel, using the Sway_Area_B column as the reference index for balance status. A classification label was defined based on the third quartile (Q3) of this variable: samples with a Sway_Area_B value above the Q3 were categorized as the "unstable" group (1), while the remainder were classified as "stable" (0).

To prevent bias and improve model stability, zero-variance features were removed using Variance Threshold, and any missing values were imputed using the median. We analyzed feature correlation to reduce redundancy, removing variables with a correlation coefficient exceeding r > 0.95. This step effectively reduced the feature space dimensions while preserving essential information.

In each iteration of the validation process, features were normalized using Standard Scaler. We then applied PCA (Principal Component Analysis) to retain 95% of the total variance, which served to eliminate noise while enhancing computational speed and stability. For hyperparameter optimization, we utilized the Optuna Bayesian search algorithm with a TPE sampler. The base model was a Support Vector Machine (SVM) with an RBF kernel. The two primary parameters—C (the regularization coefficient) and γ (the kernel coefficient)—were searched across logarithmic scales ranging from 10−3 to 103 and 10−4 to 101, respectively.

The model was evaluated using 5-fold Stratified Cross-Validation, with the AUC (Area Under the Curve) as the primary optimization metric. To address potential class imbalance, automatic class weighting was applied. For more accurate probability estimates, the final model was adjusted using Calibrated Classifier CV with sigmoid calibration. During each validation fold, metrics including precision, recall, specificity, F1-score, and ROC-AUC were calculated and averaged.

To investigate the structure of the feature space and visualize the separation between stable and unstable samples, we employed

Table 2 Performance metrics (mean values) of logistic regression models for predicting balance status.

ROC_AUC	F1	Precision	Specificity	Sensitivity	Accuracy	Fold
0.985	0.891	0.915	0.972	0.869	0.947	**3**

Table 3 Performance metrics (standard deviation values) of logistic regression models for predicting balance status.

ROC_AUC	F1	Precision	Specificity	Sensitivity	Accuracy	Fold
0.008	0.044	0.034	0.013	0.058	0.022	**1.581**

UMAP (Uniform Manifold Approximation and Projection). Unlike linear methods, UMAP is a non-linear dimensionality reduction technique that excels at preserving both local and global data structures. We implemented a Supervised UMAP approach, where the class labels were used to guide the embedding process using a target weight of 0.8 and n neighbors=15. This mapped the high-dimensional COP and exercise data into a two-dimensional space. The resulting visualization displays each participant as a distinct point; green points represent Class 0 stable balance, and red points represent Class 1 unstable balance, providing a clear graphical representation of the model's ability to differentiate between the two groups.

3- Results and Model Performance

Following correlation filtering and final feature selection, the Logistic Regression model demonstrated high efficacy in classifying participants based on their balance status. Based on the 5-fold cross-validation, the model achieved a mean Accuracy of 0.947 ± 0.022 and a mean Sensitivity of 0.870 ± 0.059. Furthermore, the Specificity was 0.972 ± 0.014, with a mean ROC-AUC of 0.985 ± 0.009. These metrics highlight the model's robust capability in distinguishing between the "impaired" and "normal" balance groups. Comprehensive performance details are provided in

2 and **Table**

Table .

To quantify the impact of individual features on the likelihood of impaired balance, we calculated the Coefficients, Standardized Coefficients OR and their respective 95% CI (Confidence Intervals). The full list of these parameters is available in Table 5 in appendix.

According to the results in Table 5, the parameters with the most significant positive impact on the probability of poor balance were MVELO_ML_B (OR = 6.53), MVELO_B (OR = 3.61), and MVELO_AP_B (OR = 2.84). The distribution of these model coefficients across all features is visualized in figure 1, while the ROC curve in figure 2 further illustrates the model's precise separability between the two study groups.

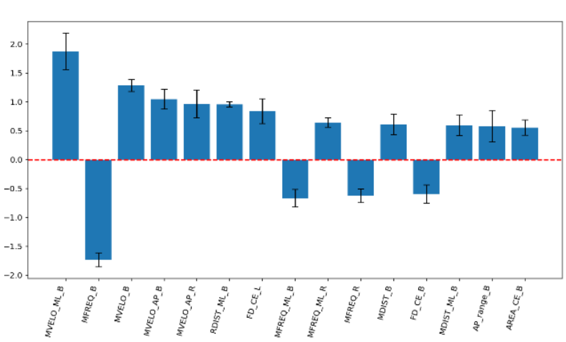


Fig. 1 Standardized coefficients of the logistic regression model. Positive coefficients indicate an increased likelihood of poor balance as the feature value increases.

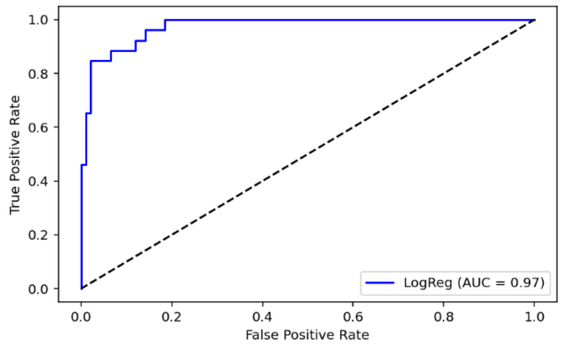


Fig. 2 ROC curve of the logistic regression model on the test data.The high area under the curve (AUC) demonstrates the model's strong discriminative ability.

Table 4 Performance metrics (mean values) of random forest models for predicting balance status.

ROC_AUC	F1	Precision	Specificity	Sensitivity	Accuracy	Fold
0.984	0.848	0.91	0.974	0.789	0.928	**3**

Table 5 Performance metrics (standard deviation values) of random forest models for predicting balance status.

ROC_AUC	F1	Precision	Specificity	Sensitivity	Accuracy	Fold
0.007	0.033	0.046	0.020	0.040	0.023	**1.581**

The Random Forest model also showed strong performance in differentiating individuals with weak balance from those that has stable balance. Across the cross-validation folds, the model achieved a mean accuracy of 0.928 ± 0.024, along with a sensitivity of 0.790 ± 0.040 and a specificity of 0.974 ± 0.021. The high mean ROC-AUC value like 0.984 ± 0.007 further indicates excellent discriminative ability. Compared with the Logistic Regression model, Random Forest shows more stable performance across all folds, as summarized in Table 4.

The most important features in the model were MVELO_ML_B, MVELO_AP_B, RDIST_ML_B, and AREA_CE_B as shown in figure 3.

To provide a better interpretation of how these features influence balance prediction, PDP (Partial Dependence Plots) were generated for the three most significant variables shown in figure 4.

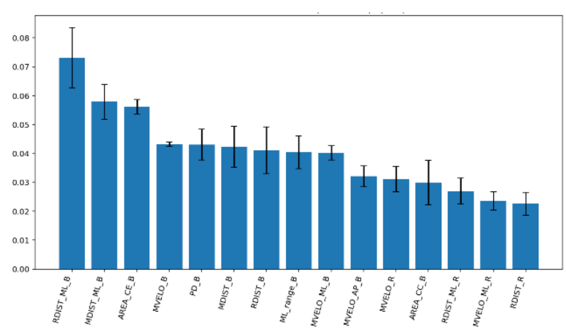


Fig. 1 Feature importance ranking from the Random Forest model. Features with higher importance contribute more to the prediction of balance status

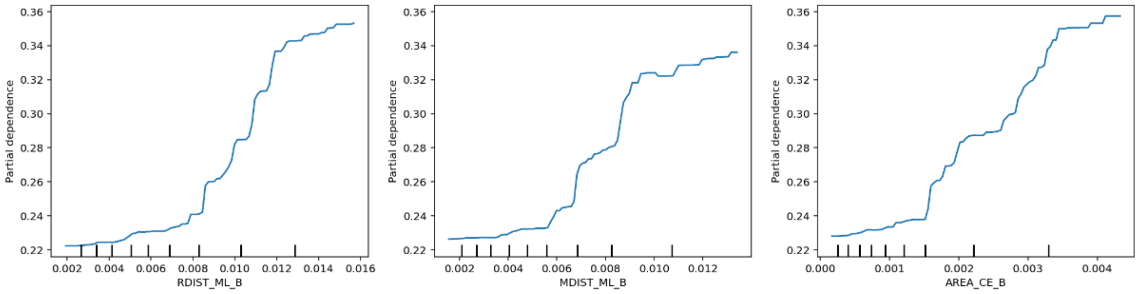


Fig. 4 Partial dependence plots of the three most important features in the Random Forest model.These plots illustrate how each feature influences the model's predicted probability of poor balance

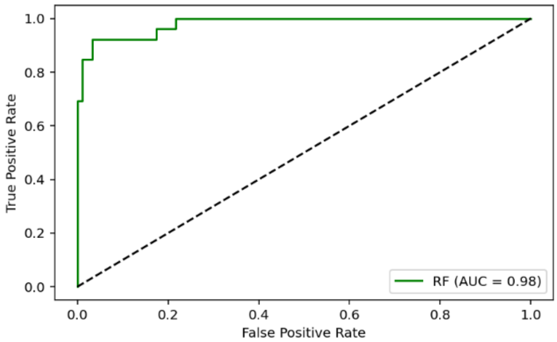


Fig. 2 ROC curve of the Random Forest model. A high AUC value confirms the strong performance of the Random Forest classifier in distinguishing between balanced and unbalanced individuals.

Additionally, the ROC curve for the Random Forest model is presented in figure 5, confirming its high classification accuracy.
The SVM model, optimized via Bayesian search for the hyperparameters C and γ demonstrated stable performance across the five-fold cross-validation iterations shown in Table 6. The model achieved a mean Accuracy of 0.923 ±0.022, with a Sensitivity (Recall) of 0.782 ±0.076 and a Specificity of 0.970 ±0.017. These results underscore the model's strong capability in correctly identifying individuals with stable balance versus those with impaired balance. The mean AUC of 0.977 ±0.013 further indicates high discriminative power within the SVM decision space. A summary of these results can be found in

Table 6 Performance metrics of the SVM model for each cross-validation fold

	accuracy	precision	recall	specificity	f1	auc	tn	fp	fn	tp
Fold1	0.940	0.960	0.800	0.988	0.872	0.985	87	1	6	**24**
Fold2	0.923	0.888	0.800	0.965	0.842	0.984	84	3	6	**24**
Fold3	0.905	0.821	0.793	0.943	0.807	0.964	83	5	6	**23**
Fold4	0.897	0.904	0.655	0.977	0.760	0.961	86	2	10	**19**
Fold5	0.948	0.925	0.862	0.977	0.892	0.989	86	2	4	**25**

Table 7 Summary of SVM performance metrics: mean and standard deviation across folds.

metric	mean	std
accuracy	0.923	**0.021**
precision	0.900	**0.051**
recall	0.782	**0.076**
specificity	0.970	**0.017**
f1	0.834	**0.052**
auc	0.977	**0.013**
tn	85.2	**1.643**
fp	2.6	**1.516**
fn	6.4	**2.190**
tp	23	**2.345**

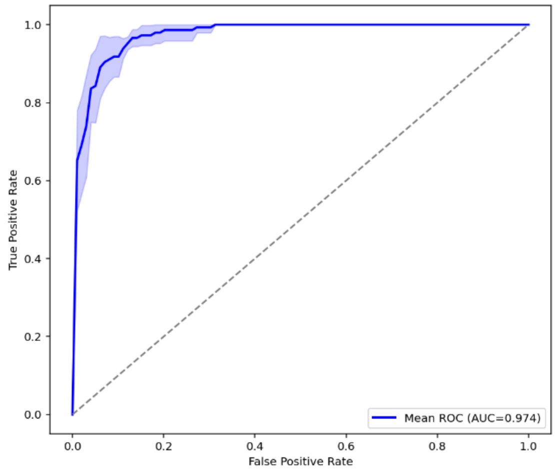


Fig. 6 Mean ROC curve of the SVM model averaged across all folds

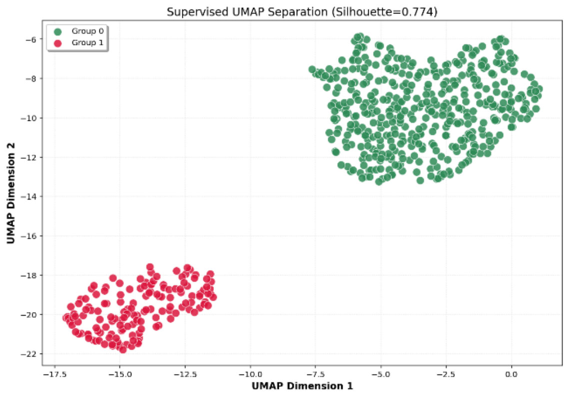


Fig. 3 UMAP visualization of the feature space

Furthermore, the mean ROC curve shown in figure 6 demonstrates the model's high stability across all iterations. The UMAP visualizations, color-coded by actual labels figure 7, reveal a distinct and near-perfect separation between the balanced and imbalanced groups within the two-dimensional feature space.

4- Discussion

Our results indicate that combining COP metrics with demographic data and athletic history enables the identification of balance-impaired individuals with a high degree of precision ROC-AUC nearly equal to 0.98. Interestingly, the proposed Neural Network, along with the Random Forest and Logistic Regression models consistently outperformed established benchmarks in the field. This suggests that a multimodal framework integrating raw biomechanical variables with specific patient profiles is essential for a truly nuanced analysis of postural control. These findings align with previous literature linking chronic low back pain to stability deficits. Specifically, meta-analytic data has shown that conservative pain management can lead to a 45% improvement in static and dynamic balance particularly across COP parameters like area and velocity [10].

While our findings strongly support velocity and range as dominant predictors, not all studies align perfectly. For instance Rojas et al. [13] achieved only 76% accuracy using approximate entropy features from COP data, significantly lower than our 94-95% range. That difference likely stems from their limited feature set entropy only versus our 63 feature multimodal approach, which captures spatial, temporal, and frequency domains. More directly a study by Park et al [14] found that sway area not velocity, was the most sensitive discriminator between LBP patients and healthy controls during eyes closed standing. The discrepancy may arise from population differences their meta analysis included heterogeneous LBP subgroups some with radiculopathy whereas our sample was restricted to chronic, nonspecific LBP without lower-limb involvement, potentially amplifying velocity based deficits.

Regarding sport intensity, our model assigned it a relatively low feature importance like OR = 0.803 and 95% CI is 0.957 to 0.674. This contradicts the hypothesis that lifetime exercise history would strongly predict balance. However, our results actually align with Zheng et al [10], who found that pain intensity changes correlate weakly with balance improvements unless accompanied by specific proprioceptive retraining. Simply having a history of high intensity sport does not guarantee good postural control if LBP has disrupted sensorimotor integration. The null effect here suggests that current pain status and COP metrics override historical physical activity as predictors, a finding that deserves direct testing in future prospective studies.

Although Logistic Regression was marginally less accurate than the Neural Network and Random Forest models, its clinical interpretability proved indispensable. The coefficient analysis (figure 1), identified several COP parameters as dominant predictors of instability. Notably, increases in mediolateral mean velocity, total mean velocity, and anteroposterior range were all strongly associated with impaired balance. In contrast, we observed a negative correlation between mean frequency and balance deficits. This suggests that higher frequency components might actually signal a more refined or active level of postural control. a finding that fits well with current physiological models of stability [11]. This high level of reliability underscores why COP metrics remain a cornerstone of predictive studies, including our own. However, stability is rarely just a physical byproduct; psychological state also significantly influences COP behavior. For instance, previous research involving female athletes with back pain has shown that pain-related anxiety can notably prolong stabilization time during dynamic tasks. This highlights that what we measure as "sway" is often an intersection of biomechanical capacity and the patient's psychological response to their condition. [12]. These insights are consistent with our own findings. by incorporating demographic and athletic history alongside mechanical COP features, our models likely accounted for the subtle, indirect influences that psychological and behavioral factors exert on postural control. Rather than looking at biomechanics in a vacuum, this integrated approach allowed the models to reflect the more complex, real-world nature of stability in patients suffering from back pain.

Our Random Forest findings complemented the Logistic Regression results by providing a non-linear perspective on feature importance. According to figure 3, the most critical predictors included mediolateral radial distance, ellipse area, and mean distance parameters. These variables reflect both the spatial dispersion of displacement and the overall magnitude of COP movement, both of which are biomechanical hallmarks of instability.

Our findings are further supported by broader meta-analyses, which consistently show that back pain patients exhibit significantly greater sway than healthy controls. This disparity is often most pronounced during sensory challenging tasks such as standing on unstable surfaces or with eyes closed where the body's reliance on proprioceptive and vestibular feedback is pushed to its limit [14]. This is consistent with our identification of COP velocity and range as the strongest predictors, suggesting that shifts in sensory integration and motor strategies are quantitatively reflected in COP data. In the machine learning area, prior studies have used COP data for balance classification; for example, a model based on approximate entropy features achieved roughly 76% accuracy [13]. While lower than our results, it demonstrates the predictive value of even limited feature sets. Our superior performance likely stems from the use of a more extensive variable set and a multimodal approach. Furthermore, while deep learning architectures like CNNs and LSTMs have reached F1 score is mostly equal to 0.98 in fall risk detection [15], our Neural Network achieved comparable performance using static, limited data. Our results also align with reports that COP changes are predictable even in the moments preceding a fall [16], confirming that COP is a reliable source for modeling instability risk in both static and dynamic conditions.

From a biomechanical standpoint, COP parameters such as displacement velocity, mediolateral and anteroposterior range, and displacement directly reflect postural control strategies. Increased velocity or displacement range typically indicates inefficient feedback control. In contrast, the presence of higher frequency components may signal more precise neuromuscular regulation. From a practical perspective in sports and rehabilitation, these findings are highly significant. First, these models can serve as representation tools to identify athletes or back pain patients at risk of falls. Second, identified key variables like velocity and range can become specific targets for stability training. Finally, combining Explainable frameworks like Logistic Regression with powerful non-linear models like Random Forest provides a more complete clinical picture.

In summary, this research emphasizes that balance impairment in athletes with back pain is a multifactorial phenomenon arising from the interaction of COP psychological, and demographic components. The core part lies in our multimodal approach and the testing of four distinct models, showing that data fusion leads to high accuracy and stability. The convergence of results across different machine learning methods reduces the likelihood of chance findings and underscores the robustness of our conclusions.

5- study limitations

Several limitations must be acknowledged. this was a cross sectional, single site study conducted in East Azerbaijan Province, which limits generalizability to other populations and geographic regions. we did not include a healthy control group without low back pain instead, we used a distribution based definition of "poor balance" within the LBP sample. While this approach is statistically valid for establishing relative risk, it does not provide absolute comparisons to asymptomatic individuals. our sport intensity categorization relied on self reported lifetime athletic history, which is subject to recall bias. the COP measurements were taken during static bipedal stance only, dynamic balance tasks for example perturbed standing or gait initiation might reveal different deficits. we did not account for pain medication use or recent exacerbations, both of which could transiently affect postural control. And Finally our sample size, while adequate for cross validation, was not powered for external validation on an independent group.

6- Future recommendations

Future studies should address these limitations through several directions. Longitudinal designs tracking LBP patients over 12 to 24 months could determine whether COP velocity increases precede or follow balance deterioration. External validation on multi site datasets ideally from different countries and ethnic groups is essential before clinical deployment. Integrating wearable inertial sensors for example lumbar or shin mounted IMUs would allow for continuous real world monitoring outside the laboratory, capturing balance during activities of daily living rather than isolated force plate trials. Additionally randomized controlled trials testing targeted velocity reduction training for example mediolateral stability exercises, against standard core stabilization would directly validate the causal role of our identified predictors. Finally, clinical validation studies comparing our ML models against expert physical therapist assessments in real-time fall prediction would establish practical utility.

7- conclusion

This study demonstrates that combining COP derived metrics with machine learning achieves exceptionally high accuracy like 94.7% with logistic regression, 98.5% ROC-AUC, in identifying balance impairment in individuals with chronic low back pain. Mediolateral mean velocity and total mean velocity emerged as the most potent biomechanical predictors, while sport intensity surprisingly showed minimal independent contribution when COP features were included. The convergence of three distinct models across five fold cross validation confirms that postural instability in LBP is not a random phenomenon but a quantifiable predictable deficit rooted in specific COP signatures. Clinically these findings provide a framework for objective fall risk screening in LBP patients using only a force plate and 60 seconds of quiet standing. Rather than relying on subjective clinical scales, rehabilitation specialists can now target velocity and range parameters as modifiable biomarkers of instability. Ultimately, this multimodal machine learning approach transforms postural control assessment from a descriptive art into a predictive science.

8- acknowledgement

The authors would like to express their gratitude to the Ministry of Science and Research and to the Education, Research, Technology, and Innovation Committee (APFAN) of East Azerbaijan Province for their support. Appreciation is also extended to East Azerbaijan Industrial Estate Company for its collaboration. In addition, we acknowledge Sahand University of Technology for its institutional assistance. We are further thankful to Sharif University of Technology for providing access to SPSS 27 software used in this work.

9- Funding

This research was financially supported by East Azerbaijan Industrial Estate Company (Grant No. 11-5227), the Ministry of Science and Research (Grant No. m/81227), and the Education, Research, Technology, and Innovation Committee (APFAN) of East Azerbaijan Province (Grant No. 105/3376), which facilitated the conduct and dissemination of the study findings.

10- AI-Assisted Language Revision and Author Verification

During the preparation of this work, the authors used chat GPT to improve the grammar and writing of the English text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article

11- Authorship Contribution Statement

M.R. Azghani conceptualized the study and proposed the core research idea. P. Dorostkar Rouhani developed the neural network codes, curated the primary dataset, and prepared the original draft of the manuscript. A. Partovifard contributed to data acquisition and performed the formal statistical analyses applied to the sports-related data. M. Pouryosef Miandoab implemented the UMAP-related software components. All authors reviewed and edited the final manuscript and accept full responsibility for its content and integrity.

12- Conflict of Interest Declaration

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. No funding was received for conducting this study, and all authors confirm that their independence and objectivity were maintained throughout the research process.

References

[1] K. F. Anthropometry, “Extraction of Foot Anthropometric Dimensions Using a Markerless 3D Scanner Based on Big Data Analysis in East Azerbaijan Province,” vol. 25, no. 03, pp. 155–161, 2025.

[2] A. Partovi, M. Azghani, S. Jalali, E. Hazrati, and S. Asghari, “Modares Mechanical Engineering Effect of Athletic Status on Medial Longitudinal Arch (MLA) in Individuals Under 20 Years of Age,” vol. 25, no. 05, pp. 261–269, 2025, doi: 10.48311/MME.25.5.261.

[3] A. J. Strang, A. DiDomenico, W. P. Berg, and R. W. McGorry, “Assessment of differenced center of pressure time series improves detection of age-related changes in postural coordination.,” Gait Posture, vol. 38, no. 2, pp. 345–348, Jun. 2013, doi: 10.1016/j.gaitpost.2012.11.006.

[4] A. Partovifard, mahmood. R. Azghani, S. jalali, E. Hazrati, and S. Asghari, “Effects of Athletic Status on Plantar Pressure Distribution and Biomechanical Foot Health in Children and Adolescents,” 2025. doi: 10.1109/ICBME68496.2025.11392560.

[5] A. Partovifard, mahmood. R. Azghani, S. jalali, and S. Asghari, “Impact of Dynamic and Static Sports on Growth and Anthropometric Characteristics (Height, Weight, BMI) in Children and Adolescents,” 2025. doi: 10.1109/ICBME68496.2025.11392550.

[6] T. Paillard, “Relationship Between Sport Expertise and Postural Skills.,” Front. Psychol., vol. 10, p. 1428, 2019, doi: 10.3389/fpsyg.2019.01428.

[7] E. Zemková and Z. Kováčiková, “Sport-specific training induced adaptations in postural control and their relationship with athletic performance.,” Front. Hum. Neurosci., vol. 16, p. 1007804, 2022, doi: 10.3389/fnhum.2022.1007804.

[8] A. P. Fard, M. Azghani, S. Jalali, and S. Asghari, “Impact of Dynamic and Static Sports on Growth and Anthropometric Characteristics (Height, Weight, BMI) in Children and Adolescents,” in 2025 32nd National and 10th International Iranian Conference on Biomedical Engineering (ICBME), IEEE, 2025, pp. 895–901.

[9] M. P. MIANDOAB, M. N. ASHTIANI, R. ABEDINI-NASSAB, and S. M. R. AKRAMI, “DETECTION OF KNEE OSTEOARTHRITIS BASED ON CENTER OF PRESSURE DATA AND THE BAT ALGORITHM,” J. Mech. Med. Biol., vol. 24, no. 03, p. 2350054, 2024, doi: 10.1142/S0219519423500549.

[10] D. K. Y. Zheng et al., “Are changes in pain intensity related to changes in balance control in individuals with chronic non-specific low back pain? A systematic review and meta-analysis.,” J. Sport Heal. Sci., vol. 14, p. 100989, Dec. 2025, doi: 10.1016/j.jshs.2024.100989.

[11] S. Ansari and S. Sharma, “Steady or Swaying? Assessing the Reliability of the HUMAC Balance System in Athletes With Chronic Low Back Pain,” Sports Health, p. 19417381251334640, May 2025, doi: 10.1177/19417381251334638.

[12] Z. Amerian, S. ShahAli, Z. S. Rezaeian, and S. Shanbehzadeh, “Dynamic postural control in women athletes with and without nonspecific low back pain with high and low pain-related anxiety- A case-control study.,” BMC Sport. Sci. Med. Rehabil., vol. 15, no. 1, p. 149, Nov. 2023, doi: 10.1186/s13102-023-00764-7.

[13] F. Rojas, I. K. Niazi, P. Maturana-Russel, and D. Taylor, “Exploring the Potential of Machine Learning for the Diagnosis of Balance Disorders Based on Centre of

Pressure Analyses," Sensors, vol. 22, no. 23. p. 9200, 2022. doi: 10.3390/s22239200.
[14] J. Park, V. Q. Nguyen, R. L. M. Ho, and S. A. Coombes, "The effect of chronic low back pain on postural control during quiet standing: A meta-analysis.," Sci. Rep., vol. 13, no. 1, p. 7928, May 2023, doi: 10.1038/s41598-023-34692-w.
[15] H.-L. Chan et al., "Deep Neural Network for the Detections of Fall and Physical Activities Using Foot Pressures and Inertial Sensing," Sensors, vol. 23, no. 1. p. 495, 2023. doi: 10.3390/s23010495.
[16] M. Rodríguez Márquez, E. D. Gutiérrez, J. S. Medina Álvarez, J. G. Milton, and J. L. Cabrera, "Machine learning forecasting of extreme fluctuations in a human balancing task," Knowledge-Based Syst., vol. 280, p. 111000, 2023, doi: 10.1016/j.knosys.2023.111000.

Appendix

Table 8 The Center of Pressure indices were extracted using the characteristic equation of each index. In this context, the letter L denotes the left foot, R denotes the right foot, and B denotes both feet. N represents the total number of stored COP points.

Parameter name	Short form	Equation
Mean distance	MDIST_L MDIST_R MDIST_B	$MD = \frac{1}{N}\sum_n \lvert R_n \rvert$
Mean distance anterior posterior	MDIST_AP_L MDIST_AP_R MDIST_AP_B	$MD\ AP = \frac{1}{N}\sum_n \lvert Y_n \rvert$
Mean distance medial-lateral	MDIST_ML_L MDIST_ML_R MDIST_ML_B	$MD\ ML = \frac{1}{N}\sum_n \lvert X_n \rvert$
Root mean square distance	RDIST_L RDIST_R RDIST_B	$RD = \sqrt{\frac{\sum_{n=1}^{N}(ML^2[n] + AP^2[n])}{N}}$
Root mean square distance anterior posterior	RDIST_AP_L RDIST_AP_R RDIST_AP_B	$RMS\ AP = \sqrt{\frac{1}{N}\sum_{n=1}^{N} Y_n^2}$
Root mean square distance medial-lateral	RDIST_ML_L RDIST_ML_R RDIST_ML_B	$RMS\ ML = \sqrt{\frac{1}{N}\sum_{n=1}^{N} X_n^2}$
anterior posterior Range or amplitude	AP_range_L AP_range_R AP_range_B	$RANGE\ AP = max_{1\le n\le m\le N}\lvert Y_n - Y_m \rvert$
medial-lateral Range or amplitude	ML_range_L ML_range_R ML_range_B	$RANGE\ ML = max_{1\le n\le m\le N}\lvert X_n - X_m \rvert$
Planar deviation	PD_L PD_R PD_B	$Planar\ dev = \sqrt{RMS\ ML^2 + RMS\ AP^2}$
Mean velocity	MVELO_L MVELO_R MVELO_B	$MEAN\ VELOCITY = \frac{SWAY\ LENGTH}{T}$
Mean velocity anterior posterior	MVELO_AP_L MVELO_AP_R MVELO_AP_B	$M\ VELOCITY\ AP = \frac{SWAY\ LENGTH\ AP}{T}$
Mean velocity medial-lateral	MVELO_ML_L MVELO_ML_R MVELO_ML_B	$M\ VELOCITY\ ML = \frac{SWAY\ LENGTH\ ML}{T}$
Fractal dimension of Confidence Ellipse	FD_CE_L FD_CE_R FD_CE_B	$FD = \lim_{\epsilon\to 0}\frac{\log N(\epsilon)}{\log\frac{1}{\epsilon}}$
Fractal dimension of Confidence circle	FD_CC_L FD_CC_R FD_CC_B	$FD = \lim_{\epsilon\to 0}\frac{\log N(\epsilon)}{\log\frac{1}{\epsilon}}$
Sway area	sway_Area_L sway_Area_R sway_Area_B	$A = \pi \times a \times b$
Mean frequency	MFREQ_L MFREQ_R MFREQ_B	$MF = \frac{\sum f_i . P(f_i)}{\sum P(f_i)}$
anterior posterior Mean frequency	MFREQ_AP_L MFREQ_AP_R MFREQ_AP_B	$MFap = \frac{\sum f_i . P_{AP}(f_i)}{\sum P_{AP}(f_i)}$
medial-lateral Mean frequency	MFREQ_ML_L MFREQ_ML_R MFREQ_ML_B	$MFml = \frac{\sum f_i . P_{ML}(f_i)}{\sum P_{ML}(f_i)}$

Table 9 The full list of all 59 coefficients and 95% confidence intervals is provided in Supplementary.

Feature	Mean Coef	Std Coef	Odds Ratio	95% CI Lower	95% CI Upper
MVELO_ML_B	1.876	0.3177	6.532	3.503	12.181
MVELO_B	1.283	0.106	3.610	2.931	4.447
MVELO_AP_B	1.045	0.172	2.844	2.029	3.986
MVELO_AP_R	0.963	0.243	2.621	1.626	4.225
RDIST_ML_B	0.958	0.0499	2.606	2.364	2.874
FD_CE_L	0.841	0.214	2.318	1.522	3.530
MFREQ_ML_R	0.643	0.083	1.902	1.615	2.239
MDIST_B	0.607	0.180	1.836	1.289	2.615
MDIST_ML_B	0.593	0.180	1.810	1.270	2.581
AP_range_B	0.577	0.275	1.781	1.0384	3.057
AREA_CE_B	0.555	0.131	1.743	1.347	2.254
RDIST_B	0.484	0.117	1.622	1.287	2.0441
PD_B	0.483	0.118	1.621	1.286	2.0438
MFREQ_AP_B	0.332	0.223	1.394	0.899	2.160
MVELO_R	0.269	0.118	1.309	1.037	1.651
ML_range_B	0.265	0.143	1.304	0.983	1.729
RDIST_ML_L	0.240	0.129	1.271	0.987	1.638
ML_range_L	0.214	0.112	1.239	0.993	1.545
MDIST_ML_L	0.170	0.084	1.186	1.004	1.400
AREA_CC_B	0.169	0.223	1.185	0.764	1.835
RDIST_ML_R	0.168	0.116	1.183	0.941	1.488
RDIST_R	0.165	0.027	1.179	1.117	1.246
PD_R	0.165	0.027	1.179	1.117	1.245
AREA_CC_R	0.161	0.187	1.174	0.814	1.695
MFREQ_AP_R	0.147	0.192	1.158	0.794	1.688
RDIST_L	0.139	0.098	1.149	0.947	1.395
PD_L	0.139	0.098	1.149	0.946	1.395
MDIST_L	0.137	0.108	1.147	0.927	1.418
MVELO_AP_L	0.124	0.122	1.132	0.890	1.439
FD_CC_B	0.107	0.145	1.113	0.837	1.480
MDIST_R	0.099	0.160	1.104	0.807	1.512
MDIST_AP_B	0.062	0.175	1.064	0.755	1.500
AREA_CC_L	0.060	0.166	1.062	0.766	1.471
MFREQ_L	0.057	0.120	1.058	0.836	1.339
MDIST_AP_L	0.044	0.140	1.045	0.794	1.376
MDIST_AP_R	0.040	0.051	1.041	0.942	1.151
FD_CC_L	0.035	0.256	1.036	0.626	1.713
RDIST_AP_B	0.024	0.140	1.024	0.778	1.348
ML_range_R	-0.006	0.145	0.993	0.746	1.321
AREA_CE_R	-0.008	0.067	0.991	0.868	1.131
MFREQ_AP_L	-0.023	0.134	0.976	0.750	1.270
MFREQ_ML_L	-0.046	0.110	0.954	0.768	1.184
MVELO_ML_R	-0.056	0.176	0.945	0.669	1.335
MVELO_L	-0.069	0.123	0.932	0.732	1.187
MVELO_ML_L	-0.088	0.172	0.915	0.652	1.285
RDIST_AP_L	-0.099	0.049	0.905	0.821	0.998
MDIST_ML_R	-0.101	0.102	0.903	0.739	1.104
AREA_CE_L	-0.150	0.106	0.860	0.699	1.058

AP_range_R	-0.187	0.109	0.828	0.668	1.027
AP_range_L	-0.190	0.215	0.826	0.541	1.261
sport intensity	-0.218	0.089	0.803	0.674	0.957
FD_CC_R	-0.296	0.177	0.743	0.524	1.052
FD_CE_R	-0.325	0.178	0.721	0.508	1.024
RDIST_AP_R	-0.436	0.104	0.646	0.526	0.792
FD_CE_B	-0.599	0.156	0.549	0.403	0.746
MFREQ_R	-0.627	0.116	0.534	0.424	0.671
MFREQ_ML_B	-0.668	0.151	0.512	0.380	0.690
MFREQ_B	-1.734	0.122	0.176	0.138	0.224s