

Article

Leveraging Multivariable Linear Regression Analysis to Identify Patients with Anterior Cruciate Ligament Deficiency Using a Composite Index of the Knee Flexion and Muscle Force

Haoran Li ¹ , Hongshi Huang ² , Shuang Ren ^{2,*} and Qiguo Rong ^{1,*} 

¹ Department of Mechanics and Engineering Science, College of Engineering, Peking University, Beijing 100871, China

² Department of Sports Medicine, Peking University Third Hospital, Institute of Sports Medicine of Peking University, Beijing 100871, China

* Correspondence: xixishuang123@126.com (S.R.); qrong@pku.edu.cn (Q.R.)

Abstract: Patients with anterior cruciate ligament (ACL) deficiency (ACLD) tend to have altered lower extremity kinematics and dynamics. Clinical diagnosis of ACLD requires more objective and convenient evaluation criteria. Twenty-five patients with ACLD before ACL reconstruction and nine healthy volunteers were recruited. Five experimental jogging data sets of each participant were collected and calculated using a musculoskeletal model. The resulting knee flexion and muscle force data were analyzed using a *t*-test for characteristic points, which were the time points in the gait cycle when the most significant difference between the two groups was observed. The data of the characteristic points were processed with principal component analysis to generate a composite index for multivariable linear regression. The accuracy rate of the regression model in diagnosing patients with ACLD was 81.4%. This study demonstrates that the multivariable linear regression model and composite index can be used to diagnose patients with ACLD. The composite index and characteristic points can be clinically objective and can be used to extract effective information quickly and conveniently.

Keywords: composite index; characteristic points; multivariable linear regression; anterior cruciate ligament deficiency



Citation: Li, H.; Huang, H.; Ren, S.; Rong, Q. Leveraging Multivariable Linear Regression Analysis to Identify Patients with Anterior Cruciate Ligament Deficiency Using a Composite Index of the Knee Flexion and Muscle Force.

Bioengineering **2023**, *10*, 284.

<https://doi.org/10.3390/bioengineering10030284>

Academic Editors: Christina Zong-Hao Ma, Zhengrong Li, Chen He and Aurélien Courvoisier

Received: 9 January 2023

Revised: 10 February 2023

Accepted: 17 February 2023

Published: 22 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Anterior cruciate ligament (ACL) deficiency (ACLD) is a common injury in people who play sports. The ACL plays an important role in maintaining the stability of the knee joint. However, because of the complexity of the knee joint and ACL, it is difficult to conduct kinematic and dynamic research on patients with ACLD. Studies have focused on building a mechanical model of ACL in vitro [1,2]. Since the establishment of a muscle model by Zajac et al. [3], musculoskeletal models have improved [4,5]. With the help of musculoskeletal models, many studies have investigated the kinematics and dynamics in ACLD-affected knees. Some studies have shown that patients with ACLD adopted quadricep avoidance [6,7] and a stiffening strategy [8], resulting in reductions in the knee flexion moment and peak knee flexion angle. ACLD affects a patient's gait patterns and further kinematics and dynamics [9]. Ren et al. [10] and Yin et al. [11], respectively, studied the kinematics and dynamics in patients with ACLD. Shelburne et al. [7] considered the role of muscles, explaining that in patients with ACLD, quadricep avoidance occurred to restore anterior tibial translation. Furthermore, increasing hamstring force was also sufficient, implying muscle compensation in the knee instability. Even after ACL reconstruction, patients still have a high risk of osteoarthritis [12,13] because of the loss of normal muscle compensation in patients with ACL reconstruction [14,15].

The clinical diagnosis of ACLD is complicated and expensive, and the diagnosis process requires the subjective judgment of clinicians. For auxiliary diagnosis, many studies have used statistics to study the gait of ACLD. Christian et al. [16] trained the gait trajectory points of patients with ACLD using support vector machines (SVMs) to extract trajectory features. Berruto et al. [17] counted the fluctuation range of the acceleration of the patient's legs with one ACL reconstruction in a pivot-shift test and demonstrated a significant difference between the ACLD-affected and contralateral sides. Zeng et al. [18] used kinematic data extracted by a motion capture system as features for neural network training. Kokkotics et al. [19] used different machine learning methods to identify patients with ACLD and ACL reconstruction from kinematics and dynamics data. However, only a few studies have used kinematics and dynamics data to diagnose patients with ACLD. There are even fewer studies that can be directly reproduced and can rapidly diagnose patients.

The feature choice is the most important variable, regardless of the statistical method. Reinbolt et al. [20] performed *t*-tests on the entire gait cycle and selected the peak points of the statistics as features to predict outcomes of rectus femoris transfer surgeries. Principal component analysis (PCA) has been widely used for dimensionality reduction and feature extraction [21]. Armstrong et al. [22] used PCA to extract the feature points of kinematics and reconstruct the kinematics process. Based on multiple parameters extracted from gait data, some indexes were developed to identify walking patterns of normal [23] and abnormal [24–27] gait. Schutte et al. [26] proposed a normalcy index to reflect gait deviations from the mean of normal gait. Liu et al. [28] assessed the abnormal gait in patients with ACLD using the normalcy index calculated by PCA based on kinematics and dynamics data. Similarly, Rozumalski et al. [29] combined a single muscle strength score using PCA to describe the overall lower body joint strength. Hicks et al. [30] used this score as a variable for multivariable regression to study crouch gait. Their regression model had 71% classification accuracy when the parameters were analyzed in detail. However, few studies have combined kinematics and muscle forces to extract features.

This study was performed to identify patients with ACLD using multivariable linear regression through a composite index that combined kinematics and muscle forces.

2. Materials and Methods

2.1. Participants

Twenty-five patients with unilateral chronic ACLD (the contralateral side was intact) were recruited before ACL reconstruction (ACLD group). Their knees had been injured 6 months to 4 years before testing. Most injuries occurred during basketball. Exclusion criteria were that the patient had no prior ACL and concomitant meniscal and ligament rupture and no history of musculoskeletal disease of the hip or ankle. Their physical activity levels were assessed by the Tegner score, which is a reliable and valid tool for assessing the activity level of patients with ACLD [31]. Average activity level of all patients was normal before knee injuries (score range 3.0–6.0). A control group comprising nine healthy volunteers with no history of musculoskeletal injury or surgery in the lower extremities was selected (Control group). All participants were young males to rule out biomechanical differences between sexes [32]. Ethical approval was obtained from the university's ethics committee, and written informed consent was obtained from all participants. The morphological data are shown in Table 1, and the participants' characteristics of the groups were not significantly different.

Table 1. Characteristics of the participants in each group.

Parameters	Control	ACL D
Age (years)	29.22 ± 5.61	27.20 ± 4.19
Height (cm)	173.67 ± 1.95	178.36 ± 7.23
Weight (kg)	74.06 ± 4.73	82.04 ± 11.55
BMI (kg/m ²)	24.56 ± 1.62	25.75 ± 2.93
Pace (m/s)	2.32 ± 0.17	2.36 ± 0.25
Time since injury (months)	/	11.10 ± 6.87
Tegner score	/	4.16 ± 1.72

Data are presented as mean ± standard deviation. ACLD, anterior cruciate ligament deficiency; BMI, body mass index.

2.2. Data Collection and Modeling Analysis

From January 2014 to December 2016, the experimental 3D data were collected while the patients were jogging using an optical motion capture system (Vicon MX; Oxford Metrics, Yarnton, Oxfordshire, UK). The marker trajectory data were filtered at 12 Hz, and the force data were filtered at 100 Hz using a low-pass Butterworth filter. To track the segmental motion during jogging, all participants had a set of markers attached to their anatomical lower limbs at specific locations based on the plug-in-gait model. The participants were asked to run along a 10-m path at a self-selected speed, and the kinematic data were recorded by eight cameras. No participants reported pain during jogging. Ground reaction forces were collected using two embedded force plates at a sampling rate of 1000 Hz (AMTI, Advanced Mechanical Technology Inc., Watertown, MA, USA). Each participant stepped on the force plates at their self-selected speed. For each participant, five successful jogging trials were recorded, and these results were imported into multi-body dynamics software, AnyBody Modeling System version 6.0.5 (AnyBody Technology, Aalborg, Denmark), to estimate the kinetics of the knee joint.

A lower extremity model [33] implemented in the AnyBody Modeling System was used for the analysis. The model comprised 12 body segments, and 11 joints were used to connect the segments. Six joint degrees of freedom were considered for each leg, with a spherical joint with three degrees of freedom for the hip joint and a universal joint with two degrees of freedom for the ankle joint. The knee joint was modeled as a hinge joint with one degree of freedom because of the soft tissue artifact error [34]. Based on the morphological parameters measured from each subject, each model was scaled with a mass-fat scaling algorithm to perform the subject-specific jogging simulation. The min/max recruitment principle solver based on the optimization of the objective function [35,36], which has good numerical convergence and physiological representation, was used to predict the muscle force during the inverse dynamics analysis. The objective function is generally formulated as follows [5]:

$$\text{Minimize } \max \left(\frac{f_i^{(M)}}{N_i} \right) \quad (1)$$

Subject to

$$\mathbf{Cf} = \mathbf{d}, \quad 0 \leq f_i^{(M)} \leq N_i, \quad i \in \{1, \dots, n^{(M)}\} \quad (2)$$

where $n^{(M)}$ is the number of muscles, $f_i^{(M)}$ is the respective muscle force, and N_i is the strength of the muscle. \mathbf{f} contains all unknown forces in the optimization problem. \mathbf{C} is the coefficient-matrix for the unknown forces. \mathbf{d} contains all known applied loads and inertia forces. Muscle parameters were obtained from a comprehensive musculoskeletal geometry dataset [37]. Some studies have validated the ability of computational muscle forces [38,39].

2.3. Muscle Data Processing

According to the characteristics of the model and anatomy related to the knee muscles, force data were output from 13 muscles: rectus femoris (RF), popliteus (POP), vastus (VAS), gastrocnemius lateralis (GL), gastrocnemius medialis (GM), soleus medialis (SOLm), soleus lateralis (SOLl), semitendinosus (ST), semimembranosus (SM), proximal sartorius (SAp), distal sartorius (SAd), biceps femoris long head (BF_{lh}), and biceps femoris short head (BF_{sh}). For each participant, separate simulations were performed based on the data from five different jogging trials. The average values of the five calculations were used to perform dynamics analysis using MATLAB version 2019b (MathWorks, Natick, MA, USA).

The acquired muscle force data was processed in a dimensionless manner, and the nondimensionalization of the force data was divided by the subject's gravity (mass \times 9.8) [40]. To investigate one gait cycle, the kinematics and dynamics data were interpolated to a 0%–100% gait cycle. Additionally, to more intuitively study the patterns of knee flexion and muscle strength, all flexion and muscle data were normalized to their maximal muscle force within that cycle, leading to a normalized amplitude between 0 and 1 [41].

2.4. Extracting Features

In this study, PCA was adopted as a statistical method for data dimensionality reduction. The main algorithm of PCA is to map the original n -dimensional data to a new k -dimensional feature that retains the largest variance. Thus, the other parts where the variance is close to zero can be ignored and the loss of information is guaranteed to be small. The flow of PCA is as follows:

1. Collect an $m \times n$ matrix G , where m is the sample size and n is the n -dimensional variable.
2. Subtract the respective mean from each variable.
3. Compute the covariance matrix of the de-averaged matrix.
4. Calculate the eigenvalues and eigenvectors of the covariance matrix by singular value decomposition.
5. Sort the eigenvalues from large to small, and select the largest k eigenvalues among them. In this study, the ratio of selected eigenvalues to the sum of all eigenvalues was used to assess the information content. Arrange the eigenvectors in the same order as the eigenvalues to form a matrix of principal component coefficients (PCcoeff).
6. Transform the data into a new space, i.e., the new data samples = $G \times$ PCcoeff. The first k columns are the required features.

Therefore, data are reduced to k dimensions. If a new sample ($1 \times n$) needs to be predicted, perform the same de-average operation first (the centered sample = the new sample—the variable mean of G from Step 2). Then, the centered sample \times PCcoeff produces the predicted value of the sample after the same processing, and the first k columns can be selected as the features of the new sample.

Inspired by the NI index [28] and the strength score [29,30], the first three features used in the calculation are the average force of each muscle during the stance/swing phase of each participant and the value of each person's knee flexion during the swing phase. The specific process was to first obtain the average value of the stance/swing phase of 13 muscle forces and then perform PCA on the average muscle force to obtain a column variable. PCA was also used to process the knee flexion data in the swing phase to obtain a column variable. Only one column variable for each feature after PCA was used because the information content was sufficiently large. For the data of knee flexion in the stance phase and other columns after PCA, their regression parameters in the following multivariable regression were not significant and did not affect the final accuracy.

2.5. Composite Index

In addition to the above features, a composite index containing the data of the knee flexion and muscle forces' characteristic points, which were the time points in the gait cycle when the most significant difference was observed between the two groups, was

used in this study. Comparing the ACLD and Control groups based on knee flexion and muscle force data for all participants, *t*-tests were performed at each point during a 0–100% gait cycle. The data of the characteristic points ($p < 0.05$ and p values were minimal) were finally filtered out as a matrix to calculate the composite index. The selection method of the characteristic points is shown in Figure 1 (using the rectus femoris as an example), and the filtered characteristic points are shown in Figure 2. The muscle force/knee flexion at these characteristic points were filtered out to form a matrix, where the rows of the matrix were the number of participants and the columns of the matrix were the number of characteristic points. Finally, PCA was used to process this matrix to select the representative columns as features.

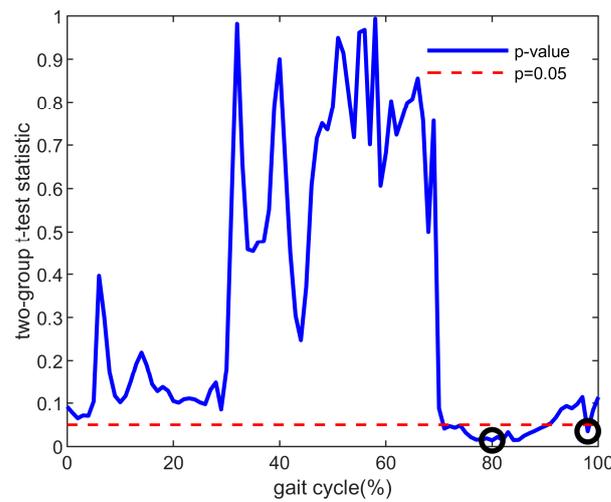


Figure 1. Example of p values between the ACLD and Control groups (taking the muscle force of rectus femoris as an example). The red dotted line means that the p value equals 0.05. The black circle is the selected characteristic point used to calculate the composite index, which is the minimum value in the range of significant differences ($p < 0.05$). ACLD, anterior cruciate ligament deficiency.

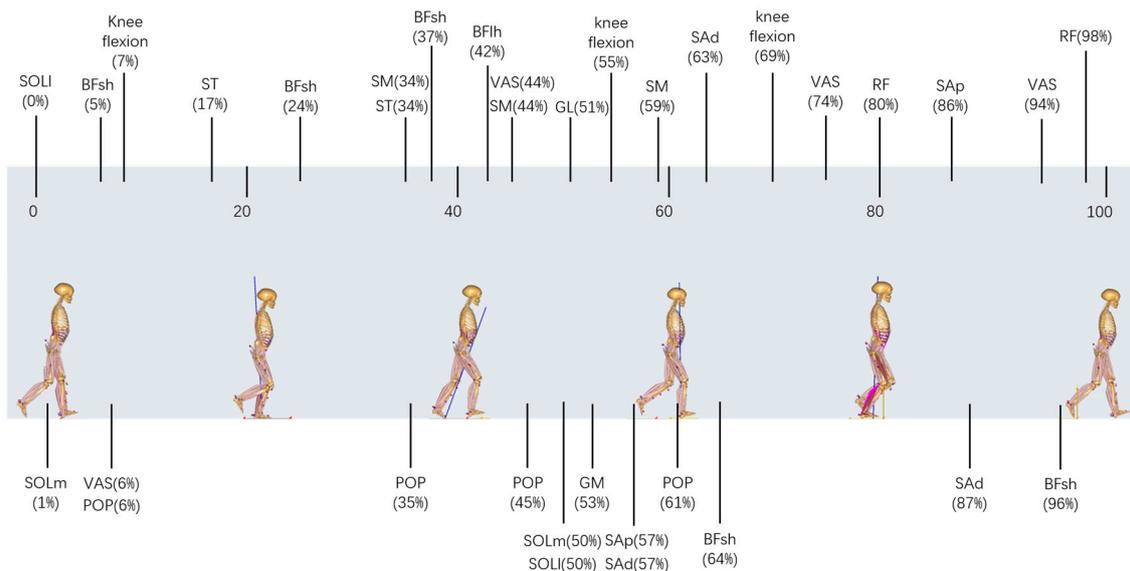


Figure 2. The filtered characteristic points, which were the time points in the gait cycle when the most significant difference ($p < 0.05$ and p values were minimal) occurred between the ACLD and Control groups according to the two-sample *t*-test method. Characteristic points are shown at their corresponding points during the 0–100% gait cycle. ACLD, anterior cruciate ligament deficiency.

2.6. Statistical Analysis

Using the above variables and samples, we built a multivariable linear regression model. The general form of this model is:

$$Y_{diagnosis} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_q X_q \quad (3)$$

The outcome variable $Y_{diagnosis}$ was a diagnosis of whether the participant was a patient with ACLD, such that positive values correspond with patients with ACLD and negative values correspond to the Control group. In the data used for training, the ACLD group had $Y_{diagnosis} = 1$ and the Control group had $Y_{diagnosis} = -1$. X_i are the predictive features in the data obtained above and β_i are the linear weighting coefficients for the predictive features. The formulation and prediction of the model were conducted in MATLAB.

3. Results

For these 43 samples (the affected legs of the 25 patients in the ACLD group and both legs of the 9 participants in the Control group), the number of variables selected will affect the final accuracy. The final prediction accuracy changes with feature selection changes are shown in Table 2, where 5-fold cross-validation was used to estimate the predictive ability of the regression model. The second column showed whether only the composite index was used as features. If not, the first column showed the first three features (knee flexion and mean values of muscle force during the stance/swing phase) + the number of features retained in the composite index. If yes, the first column showed only the number of features retained in the composite index. The composite index produced eight features when 90% of the PCA information content was preserved. Therefore, when just using the composite index as features and considering all eight features produced by the composite index, the maximum accuracy achieved was 81.4%. The last column showed the p values in the t -test for the coefficients of the features of the composite index during regression. The smaller the p value, the more significant the corresponding feature. $p < 0.001$ indicated very significant findings and was replaced by 0.001 in Table 2. When using the first 3 features + the composite index, the accuracy gradually increased as the features produced by the composite index increased. When the features produced by the composite index were more than three, the accuracy remained the same and the p value of the newly introduced features increased and was not significant. When the features produced by the composite index were equal to five, the p value of the last feature was 0.999, indicating that the newly introduced feature had no new information. For a comprehensive comparison, the optimal condition was to select six features (the first three features + three composite index features), and the accuracy rate after 5-fold cross-validation was 81.4%. For comparison and validation, under the condition of using only three composite index features, the accuracy was 79.1%.

The classification ability evaluation of the optimal condition is shown in Table 3. The actual results of classification and the accuracy, precision, recall, specificity, and F1-score were used to evaluate the classification ability of the regression model under the optimal condition. Most of the actual results were correctly classified. All evaluation criteria were above 80%, which proved the good performance of the regression model.

Finally, multivariable linear regression was performed on all samples, and the resulting model is shown in Table 4. In Table 4, the coefficients of the average muscle force during the swing phase, Composite Index 1, and Composite Index 2 were negative and their absolute values were the largest among all coefficients. The p value of Composite Index 1 was less than 0.001, the p value of Composite Index 2 was 0.006, and the p value of Composite Index 3 was 0.208. The overall p value of the regression model was less than 0.001.

Table 2. Accuracy changes as the number of features change.

Number of Features	Use of the Composite Index Only	Accuracy	<i>p</i> Value for the <i>t</i> -Test of Regression Coefficients for the Composite Index
3 + 1	No	67.4%	0.001
3 + 2	No	74.4%	0.001 0.001
3 + 3	No	81.4%	0.001 0.006 0.208
3 + 4	No	81.4%	0.001 0.008 0.209 0.686
3 + 5	No	81.4%	0.001 0.009 0.225 0.692 0.999
0 + 1	Yes	72.1%	0.001
0 + 3	Yes	79.1%	0.001 0.001 0.75
0 + 8	Yes	81.4%	0.001 0.001 0.74 0.66 0.45 0.01 0.3 0.48

If the composite index is used as features only, the first column shows 0 + the number of features retained in the composite index. If not, the first column shows three features (knee flexion and mean values of muscle force during the stance/swing phase) + the number of features retained in the composite index. *p* < 0.001 indicates very significant results and is replaced by 0.001 in the last column.

Table 3. Evaluation criteria for the classification ability.

TP	FP	FN	TN	Accuracy	Precision	Recall	Specificity	F1-Score
20	3	5	15	81.4%	87.0%	80.0%	83.3%	83.3%

TP = true positive, samples are classified as positive (ACLD group) and the judgment is correct. FN = false negative, samples are classified as negative (Control group) and the judgment is wrong. FP = false positive, samples are classified as positive and the judgment is wrong. TN = true negative, samples are classified as negative and the judgment is correct. Accuracy = (TP + TN)/(TP + TN + FP + FN). Precision = TP/(TP + FP). Recall = TP/(TP + FN). F1-score represents the harmonic mean of precision and recall. ACLD, anterior cruciate ligament deficiency.

Table 4. Multivariable linear regression model of all samples.

Features	Coefficient	Standard Error of the Coefficients	<i>p</i> Value ^a
Constant	0.1628	0.1113	0.152
Mean muscle force during the stance phase	0.2384	0.1850	0.205
Average muscle force during the swing phase	−2.5339	0.1968	0.206
Knee flexion during the swing phase	0.7346	0.5518	0.191
Composite Index 1	−2.3055	0.5040	<0.001
Composite Index 2	−1.5697	0.5468	0.006
Composite Index 3	1.2417	0.9698	0.208
RMSE		0.73	
R-squared		0.542	
<i>p</i> value ^b		<0.001	

^a *p* value for the *t*-test of each regression coefficient. ^b *p* value for the F-test on the model. Composite index 1–3 represent the first three features of the composite index, respectively. RMSE, root mean squared error.

4. Discussion

The multivariable linear regression model using the composite index was able to predict, with 81.4% accuracy, whether participants had ACLD. Under the optimal condition (Table 3), data were well classified, and the evaluation criteria were greater than 80%. Among them, the value of precision was high (87.0%), meaning that the correct proportion of the samples classified as the ACLD group was high. Our model was very capable in diagnosing patients with ACLD. The F1-score was high (83.3%), indicating that our model was effective.

As shown in Table 2, when the composite index was used as features only, the best accuracy of 81.4% was achieved by retaining all eight variables. With only one variable of the composite index, there was still 72.1% accuracy. After importing the first three features

and retaining three features of the composite index, the best accuracy of 81.4% was achieved. Therefore, this composite index characterized the information of kinematics and dynamics. Using only the three features of the composite index can also achieve an accuracy of 79.1%. Based on the optimal condition of six features, when more composite index features were imported, the accuracy remained unchanged and the p value of the coefficient continued to increase closer to 1, indicating that the introduction of more features was no longer significant. Therefore, the composite index contained more information in the model. Most information in knee flexion and muscle force can be covered in the composite index.

Interpretation of the model must be taken with caution (Table 4). The R-squared value of the regression was 0.542, which indicated that the model was able to explain 54.2% of the variance in the diagnosis of patients with ACLD [30]. For all samples, the optimal features used for regression were significant at the $p = 0.21$ level by the t -test. For the F-test on the model, $p < 0.001$ indicated that the fitting process of the model was very significant. The root mean squared error, which estimated the standard deviation of the error distribution, was equal to 0.73, indicating that the model fit well. As shown in Table 4, there was a significant ($p < 0.001$) negative relationship between Composite Index 1 and the diagnostic outcome, with an expected 2.3055 decline in the final outcome for each one-point increase in the first variable of the composite index. Additionally, there was also a significant ($p < 0.05$) negative relationship between Composite Index 2 and the outcome, indicating that the composite index, especially the first two variables, played the most important role in the regression model.

The composite index was determined by the data of the characteristic points in muscle force and knee flexion (Figure 2). Each muscle selected was associated with the knee joint, which aids in understanding gait pathology and planning treatment using gait analysis and biomechanical models [20,30]. As shown in Figure 2, most of the characteristic points in thigh muscles were concentrated at the terminal of the stance phase, which also corresponded to the previous studies, especially decreased quadriceps [6,7] and increased hamstring [7,42,43]. Alternatively, although tibialis triceps were active during the mid-stance phase, they had more of an impact on ankle dorsiflexion during this period and thus were not significantly different in patients with ACLD [44]. In this study, the muscle force and knee flexion were normalized. Therefore, in further research or clinical diagnosis and treatment, even if the muscle force or knee flexion is obtained in different principles, the characteristic points in Figure 2 can still be directly selected.

With further validation, the regression model can be used to aid clinical practice [30]. Table 5 describes the characteristics of two hypothetical subjects from the ACLD and Control groups. These two subjects have feature values close to the mean of their respective groups. Their expected final outcomes are 0.6164 and -0.4672 , respectively, which can be clearly classified into the ACLD group and Control group. Of all the features, the expected improvements in Composite Index 1 and Composite Index 2 have the most impact on the final outcomes. Notably, the subject values of the ACLD group are all less than 0, while the subject values of the Control group are all greater than 0. With t -tests between the two groups on the composite index, we were able to obtain $p < 0.001$ for the first variable of the composite index and $p < 0.05$ for the second variable, verifying the validity of the composite index in the regression model and demonstrating that using only the composite index is also a successful evaluation index, similar to the normalcy index [28] and strength score [29,30].

Some limitations of this study should be noted. First, some patients with ACLD also had meniscus injuries. One study [45] has shown that about 40% to 80% of patients with ACLD have a concurrent meniscal injury. Grouping the data more deeply will help improve our accuracy. Second, data quality can be further improved. The compensatory patterns in the knee joint change depending on the time after ACLD. Limited by clinical data, the period of patients' injuries in this study was not concentrated. Unconcentrated data may affect the accuracy of the final results. Third, electromyography (EMG) data can be introduced. EMG can assist in validating the muscle forces obtained from the

calculations [46]. In addition, EMG data can be directly involved in the calculation to obtain a composite index. Fourth, the results of the composite index need further validation for the assessment of walking.

Table 5. Characteristics of two hypothetical subjects from the ACLD and Control groups.

Features	ACLD Group		Control Group	
	Subject Value	Expected Improvement	Subject Value	Expected Improvement
Constant	1	0.1628	1	0.1628
Mean muscle force during the stance phase	−0.1757	−0.0419	0.2440	0.0582
Average muscle force during the swing phase	−0.0063	0.0160	0.0088	−0.0223
Knee flexion during the swing phase	−0.0051	−0.0037	0.0071	0.0052
Composite Index 1	−0.1490	0.3435	0.2069	−0.4771
Composite Index 2	−0.0937	0.1470	0.1301	−0.2042
Composite Index 3	−0.0059	−0.0073	0.0082	0.0102
Expected outcome		0.6164		−0.4672

Composite Index 1–3 are respectively the first three features of the composite index. ACLD, anterior cruciate ligament deficiency.

5. Conclusions

We built a multivariable linear regression model to diagnose patients with ACLD using a composite index that combined knee flexion and muscle forces. This statistical model and composite index can aid clinical diagnosis. The composite index and characteristic points can help avoid complex subjective diagnosis in clinical practice and can be used to extract effective information more quickly and conveniently for diagnosis.

Author Contributions: Conceptualization, Q.R.; methodology, H.L.; software, H.L.; validation, Q.R., H.H. and S.R.; formal analysis, H.L.; investigation, H.L.; resources, H.H. and S.R.; data curation, H.H. and S.R.; writing—original draft preparation, H.L.; writing—review and editing, Q.R. and H.L.; visualization, S.R.; supervision, Q.R.; project administration, Q.R.; funding acquisition, Q.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of China (Grant No: 11872074, 82202821, 31900943), Beijing Municipal Natural Science Foundation (Grant No: L222138), Peking University Third Hospital (Grant No: BYSYZHKC2022119, BYSY2022058, BYSYZD2021012).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of Peking University Third Hospital (IRB00006761-2012010).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kanamori, A.; Zeminski, J.; Rudy, T.W.; Li, G.; Fu, F.H.; Woo, S.L.-Y. The effect of axial tibial torque on the function of the anterior cruciate ligament: A biomechanical study of a simulated pivot shift test. *Arthrosc. J. Arthrosc. Relat. Surg.* **2002**, *18*, 394–398. [[CrossRef](#)] [[PubMed](#)]
2. Sakane, M.; Woo, S.L.-Y.; Hildebrand, K.A.; Fox, R.J. The contribution of the anterior cruciate ligament to knee joint kinematics: Evaluation of its in situ forces using a robot/universal force-moment sensor test system. *J. Orthop. Sci.* **1996**, *1*, 335–347. [[CrossRef](#)]
3. Zajac, F.E. Muscle and tendon: Properties, models, scaling, and application to biomechanics and motor control. *Crit. Rev. Biomed. Eng.* **1989**, *17*, 359–411.
4. Delp, S.L.; Anderson, F.C.; Arnold, A.S.; Loan, P.; Habib, A.; John, C.T.; Guendelman, E.; Thelen, D.G. OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Trans. Biomed. Eng.* **2007**, *54*, 1940–1950. [[CrossRef](#)]
5. Damsgaard, M.; Rasmussen, J.; Christensen, S.T.; Surma, E.; de Zee, M. Analysis of musculoskeletal systems in the AnyBody Modeling System. *Simul. Model. Pract. Theory* **2006**, *14*, 1100–1111. [[CrossRef](#)]

6. Berchuck, M.; Andriacchi, T.P.; Bach, B.R.; Reider, B. Gait adaptations by patients who have a deficient anterior cruciate ligament. *JBJS* **1990**, *72*, 871–877. [[CrossRef](#)]
7. Shelburne, K.B.; Torry, M.R.; Pandy, M.G. Effect of Muscle Compensation on Knee Instability during ACL-Deficient Gait. *Med. Sci. Sports Exerc.* **2005**, *37*, 642–648. [[CrossRef](#)]
8. Hurd, W.J.; Snyder-Mackler, L. Knee instability after acute ACL rupture affects movement patterns during the mid-stance phase of gait. *J. Orthop. Res.* **2007**, *25*, 1369–1377. [[CrossRef](#)]
9. Zhang, Y.; Huang, W.-H.; Yao, Z.; Ma, L.; Lin, Z.; Wang, S.; Huang, H. Anterior Cruciate Ligament Injuries Alter the Kinematics of Knees with or without Meniscal Deficiency. *Am. J. Sports Med.* **2016**, *44*, 3132–3139. [[CrossRef](#)]
10. Ren, S.; Yu, Y.; Shi, H.; Miao, X.; Jiang, Y.; Liang, Z.; Hu, X.; Huang, H.; Ao, Y. Three dimensional knee kinematics and kinetics in ACL-deficient patients with and without medial meniscus posterior horn tear during level walking. *Gait Posture* **2018**, *66*, 26–31. [[CrossRef](#)]
11. Yin, W.; Ren, S.; Huang, H.; Yu, Y.; Liang, Z.; Ao, Y.; Rong, Q. Dynamical Characteristics of Anterior Cruciate Ligament Deficiency Combined Meniscus Injury Knees. In *Advanced Computational Methods in Life System Modeling and Simulation*; Springer: Singapore, 2017; pp. 132–139.
12. Barenius, B.; Ponzer, S.; Shalabi, A.; Bujak, R.; Norlén, L.; Eriksson, K. Increased Risk of Osteoarthritis After Anterior Cruciate Ligament Reconstruction: A 14-year follow-up study of a randomized controlled trial. *Am. J. Sports Med.* **2014**, *42*, 1049–1057. [[CrossRef](#)]
13. Wellsandt, E.; Gardinier, E.S.; Manal, K.; Axe, M.J.; Buchanan, T.S.; Snyder-Mackler, L. Decreased Knee Joint Loading Associated with Early Knee Osteoarthritis After Anterior Cruciate Ligament Injury. *Am. J. Sports Med.* **2016**, *44*, 143–151. [[CrossRef](#)] [[PubMed](#)]
14. Makihara, Y.; Nishino, A.; Fukubayashi, T.; Kanamori, A. Decrease of knee flexion torque in patients with ACL reconstruction: Combined analysis of the architecture and function of the knee flexor muscles. *Knee Surg. Sports Traumatol. Arthrosc.* **2006**, *14*, 310–317. [[CrossRef](#)] [[PubMed](#)]
15. Smeets, A.; Willems, M.; Gilson, L.; Verschueren, S.; Staes, F.; Vandenuecker, H.; Claes, S.; Vanrenterghem, J. Neuromuscular and biomechanical landing alterations persist in athletes returning to sport after anterior cruciate ligament reconstruction. *Knee* **2021**, *33*, 305–317. [[CrossRef](#)]
16. Christian, J.; Kröll, J.; Strutzenberger, G.; Alexander, N.; Ofner, M.; Schwameder, H. Computer aided analysis of gait patterns in patients with acute anterior cruciate ligament injury. *Clin. Biomech.* **2016**, *33*, 55–60. [[CrossRef](#)]
17. Berruto, M.; Uboldi, F.; Gala, L.; Marelli, B.; Albisetti, W. Is triaxial accelerometer reliable in the evaluation and grading of knee pivot-shift phenomenon? *Knee Surg. Sports Traumatol. Arthrosc.* **2013**, *21*, 981–985. [[CrossRef](#)]
18. Zeng, W.; Ismail, S.A.; Pappas, E. Classification of gait patterns in patients with unilateral anterior cruciate ligament deficiency based on phase space reconstruction, Euclidean distance and neural networks. *Soft Comput.* **2020**, *24*, 1851–1868. [[CrossRef](#)]
19. Kokkotis, C.; Moustakidis, S.; Tsatalas, T.; Ntakolia, C.; Chalatsis, G.; Konstadakos, S.; Hantes, M.E.; Giakas, G.; Tsaopoulos, D. Leveraging explainable machine learning to identify gait biomechanical parameters associated with anterior cruciate ligament injury. *Sci. Rep.* **2022**, *12*, 6647. [[CrossRef](#)]
20. Reinbolt, J.A.; Fox, M.D.; Schwartz, M.H.; Delp, S.L. Predicting outcomes of rectus femoris transfer surgery. *Gait Posture* **2009**, *30*, 100–105. [[CrossRef](#)]
21. Deluzio, K.; Astephen, J. Biomechanical features of gait waveform data associated with knee osteoarthritis: An application of principal component analysis. *Gait Posture* **2007**, *25*, 86–93. [[CrossRef](#)]
22. Armstrong, D.P.; Pretty, S.P.; Weaver, T.B.; Fischer, S.L.; Laing, A.C. Application of Principal Component Analysis to Forward Reactive Stepping: Whole-body Movement Strategy Differs as a Function of Age and Sex. *Gait Posture* **2021**, *89*, 38–44. [[CrossRef](#)]
23. Tingley, M.; Wilson, C.; Biden, E.; Knight, W. An index to quantify normality of gait in young children. *Gait Posture* **2002**, *16*, 149–158. [[CrossRef](#)]
24. Leporace, G.; Batista, L.A.; Muniz, A.M.; Zeitoune, G.; Luciano, T.; Metsavaht, L.; Nadal, J. Classification of gait kinematics of anterior cruciate ligament reconstructed subjects using principal component analysis and regressions modelling. In *Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012*; pp. 6514–6517. [[CrossRef](#)]
25. Romei, M.; Galli, M.; Motta, F.; Schwartz, M.; Crivellini, M. Use of the normalcy index for the evaluation of gait pathology. *Gait Posture* **2004**, *19*, 85–90. [[CrossRef](#)]
26. Schutte, L.M.; Narayanan, U.; Stout, J.L.; Selber, P.; Gage, J.R.; Schwartz, M.H. An index for quantifying deviations from normal gait. *Gait Posture* **2000**, *11*, 25–31. [[CrossRef](#)]
27. Shin, K.-Y.; Rim, Y.H.; Kim, Y.S.; Kim, H.S.; Han, J.W.; Choi, C.H.; Lee, K.S.; Mun, J.H. A joint normalcy index to evaluate patients with gait pathologies in the functional aspects of joint mobility. *J. Mech. Sci. Technol.* **2010**, *24*, 1901–1909. [[CrossRef](#)]
28. Liu, X.; Huang, H.; Ren, S.; Rong, Q.; Ao, Y. Use of the normalcy index for the assessment of abnormal gait in the anterior cruciate ligament deficiency combined with meniscus injury. *Comput. Methods Biomech. Biomed. Eng.* **2020**, *23*, 1102–1108. [[CrossRef](#)]
29. Rozumalski, A.; Schwartz, M.H. Crouch gait patterns defined using k-means cluster analysis are related to underlying clinical pathology. *Gait Posture* **2009**, *30*, 155–160. [[CrossRef](#)]
30. Hicks, J.L.; Delp, S.L.; Schwartz, M.H. Can biomechanical variables predict improvement in crouch gait? *Gait Posture* **2011**, *34*, 197–201. [[CrossRef](#)]

31. Huang, H.; Zhang, D.; Jiang, Y.; Yang, J.; Feng, T.; Gong, X.; Wang, J.; Ao, Y. Translation, validation and cross-cultural adaptation of a simplified-Chinese version of the Tegner Activity Score in Chinese patients with anterior cruciate ligament injury. *PLoS ONE* **2016**, *11*, e0155463. [[CrossRef](#)]
32. Théoret, D.; Lamontagne, M. Study on three-dimensional kinematics and electromyography of ACL deficient knee participants wearing a functional knee brace during running. *Knee Surg. Sports Traumatol. Arthrosc.* **2006**, *14*, 555–563. [[CrossRef](#)]
33. Horsman, M.K.; Koopman, H.; van der Helm, F.; Prosé, L.P.; Veeger, H. Morphological muscle and joint parameters for musculoskeletal modelling of the lower extremity. *Clin. Biomech.* **2007**, *22*, 239–247. [[CrossRef](#)]
34. Andersen, M.S.; Benoit, D.L.; Damsgaard, M.; Ramsey, D.K.; Rasmussen, J. Do kinematic models reduce the effects of soft tissue artefacts in skin marker-based motion analysis? An in vivo study of knee kinematics. *J. Biomech.* **2010**, *43*, 268–273. [[CrossRef](#)]
35. Marra, M.A.; Vanheule, V.; Fluit, R.; Koopman, B.H.F.J.M.; Rasmussen, J.; Verdonschot, N.; Andersen, M.S. A Subject-Specific Musculoskeletal Modeling Framework to Predict In Vivo Mechanics of Total Knee Arthroplasty. *J. Biomech. Eng.* **2015**, *137*, 020904. [[CrossRef](#)]
36. Crowninshield, R. Use of optimization techniques to predict muscle forces. *J. Biomech.* **1979**, *12*, 627. [[CrossRef](#)]
37. Carbone, V.; Fluit, R.; Pellikaan, P.; van der Krogt, M.M.; Janssen, D.; Damsgaard, M.; Vigneron, L.; Feilkas, T.; Koopman, H.F.J.M.; Verdonschot, N. TLEM 2.0—A comprehensive musculoskeletal geometry dataset for subject-specific modeling of lower extremity. *J. Biomech.* **2015**, *48*, 734–741. [[CrossRef](#)]
38. Carbone, V.; van der Krogt, M.M.; Koopman, H.F.; Verdonschot, N. Sensitivity of subject-specific models to Hill muscle–tendon model parameters in simulations of gait. *J. Biomech.* **2016**, *49*, 1953–1960. [[CrossRef](#)] [[PubMed](#)]
39. Trinler, U.; Schwameder, H.; Baker, R.; Alexander, N. Muscle force estimation in clinical gait analysis using AnyBody and OpenSim. *J. Biomech.* **2019**, *86*, 55–63. [[CrossRef](#)]
40. Andriacchi, T.P.; Dyrby, C.O. Interactions between kinematics and loading during walking for the normal and ACL deficient knee. *J. Biomech.* **2005**, *38*, 293–298. [[CrossRef](#)]
41. Hug, F.; Vogel, C.; Tucker, K.; Dorel, S.; Deschamps, T.; Le Carpentier, É.; Lacourpaille, L. Individuals have unique muscle activation signatures as revealed during gait and pedaling. *J. Appl. Physiol.* **2019**, *127*, 1165–1174. [[CrossRef](#)] [[PubMed](#)]
42. Liu, W.; Maitland, M.E. The effect of hamstring muscle compensation for anterior laxity in the ACL-deficient knee during gait. *J. Biomech.* **2000**, *33*, 871–879. [[CrossRef](#)]
43. Catalfamo, P.F.; Aguiar, G.; Curi, J.; Braidot, A. Anterior Cruciate Ligament Injury: Compensation during Gait using Hamstring Muscle Activity. *Open Biomed. Eng. J.* **2010**, *4*, 99–106. [[CrossRef](#)]
44. Lenhart, R.L.; Francis, C.A.; Lenz, A.L.; Thelen, D.G. Empirical evaluation of gastrocnemius and soleus function during walking. *J. Biomech.* **2014**, *47*, 2969–2974. [[CrossRef](#)] [[PubMed](#)]
45. Granan, L.-P.; Inacio, M.C.; Maletis, G.B.; Funahashi, T.T.; Engebretsen, L. Sport-Specific Injury Pattern Recorded During Anterior Cruciate Ligament Reconstruction. *Am. J. Sports Med.* **2013**, *41*, 2814–2818. [[CrossRef](#)] [[PubMed](#)]
46. Lenhart, R.L.; Thelen, D.G.; Wille, C.M.; Chumanov, E.S.; Heiderscheit, B.C. Increasing Running Step Rate Reduces Patellofemoral Joint Forces. *Med. Sci. Sports Exerc.* **2014**, *46*, 557–564. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.