# Identifying Individuals with Anterior Cruciate Ligament Injury through Spectral and Linear Discriminant Analysis

Kristin Morgan[1], Carolyn Morgan[2], Heather Bush[3]
Brian Noehren[4]
[1]Department of Biomedical Engineering, University of Connecticut, Storrs, CT 06269
[2]Statistician, MECK Limited, LLC, Williamsburg, VA 23185
[3]Department of Biostatistics, University of Kentucky, Lexington, KY 40506
[4]Division of Physical Therapy, University of Kentucky, Lexington, KY 40506

**Abstract**
In human movement biomechanics, healthy and injured populations are often differentiated based on discrete measures, such as peak knee flexion angles. Discrete measures do not capture the changing dynamics that can characterize altered joint motion in injured populations. To capture these changing dynamics, spectral analyses will be employed. The results of these analyses will be used to develop a regression model to classify individuals into healthy and injured populations.

This study will use fast Fourier transform and control theory techniques to obtain frequency, phase, and stability metrics to characterize the gait in healthy individuals and individuals who have undergone an anterior cruciate ligament (ACL) reconstruction. Then, linear discriminant analysis will be used to reveal the critical variables that best classify the healthy and injured populations. It will also yield a model that describes the relationship between these variables that are very important to identifying individuals with adverse gait biomechanics. This analysis will allow us to potentially predict individuals with future risk of injury and to design more targeted ACL prevention and rehabilitation programs.

**Key Words:** Time series, Spectral Analysis, Regression, Linear Discriminant Analysis, Anterior Cruciate Ligament Reconstruction

## 1. Introduction

Anterior Cruciate Ligament (ACL) sprain or tear is encountered by one in 3,000 individuals annually (Boden et al. 2000). Despite advancements in research and ACL injury prevention programs, ACL injury rates have continued to rise (Donnelly et al. 2012). ACL injury results in loss of dynamic knee stability which is vital for movements like running and single-leg jump landing (Arden et al. 2014). Many studies have been conducted to understand the causes of ACL injury. In their research, Morgan et al. (2014) revealed how elevated gastrocnemius forces compensate for decreased hamstring forces during the weight-acceptance phase of single-leg jump landing and highlighted the implications for anterior cruciate ligament injury risk. Most dynamic knee stability data recorded on individuals during running and jump landing studies are in a time series (i.e., a sequence of data points, typically consisting of successive measurements made over a time interval.

However, quite often discrete measures are used to evaluate this data. Additional information, not unveiled in the time domain, could possibly provide valuable insight into alterations in knee gait patterns in post-ACL reconstructed (ACLR) individuals. De Fontenay et al. (2014) and Gao et al. (2010) have assessed dynamic gait stability via methods such as Lyapunov exponents. Also, Morgan et al. (2016) used the Nyquist and Bode stability criteria to assess changes in dynamic knee stability in healthy and anterior cruciate ligament reconstructed individuals during walking.

This work proposes the use of fast Fourier Transform (FFT) frequency-based techniques to capture and quantify changing knee dynamics using amplitude, frequency, and phase metrics. Past research has demonstrated that frequency domain information can be clinically relevant when analyzing gait patterns (e.g., Osgood, 2007; Shabani et al., 2015; Almekinders et al., 1995). Both Giakas et al. (1997) and Stergiou et al. (2002) successfully used frequency domain analyses to differentiate between healthy controls and individuals with scoliosis and elderly individuals, respectively, when the time domain variables failed to do so. This study uses FFT to compare differences in knee gait patterns between healthy and post-ACL reconstructed individuals. Unlike past studies that focused on the frequency component of FFT alone, this study focuses on the amplitude and phase metrics as well. Hence, multiple techniques will be used to evaluate dynamic knee stability from sagittal knee kinematics in individuals during running. We hypothesize that by using these techniques it will be possible to differentiate between stable and unstable knee biomechanics and the amplitude, frequency, and phase components will be employed to detect changes in gait patterns for the healthy individuals and post-ACL individuals. Linear discriminant analysis (LDA) will be employed to identify the key gait features that are associated with gait biomechanics and contribute to separating the healthy and injured groups. The results will be utilized to develop rule(s) to classify healthy and post-ACL reconstructed individuals into groups using the key gait features identified.

## 2. Methods

### 2.1 Human Subject Study Population

The were 32 participants in the study. Sixteen subjects (height $1.7\pm\mathbf{0}.1$ m; mass $66.7\pm13.5$ kg; age $20.88\pm3.9$ years) were healthy and sixteen injured subjects (height $1.7\pm\mathbf{0}.1$ m; mass $68.83\pm10.5$ kg; age $19.4\pm\mathbf{5.1}$ years) had experienced an ACL injury and surgery had been performed on the injured limb. All the participants were between the ages of 16 and 40 years old. The individuals were cleared by their physician to participate and they each provided written informed consent as required by the institutional review board. All subjects wore the same type of running sneaker.

Three population groups – **CTLR**, **ACLI** and **ACLNI** – were defined for this study. Based on the findings from prior research, after surgery for an ACL injury on either limb, the risk for injury (especially an ACL injury) increases. The **CTLR** group included the sixteen healthy control subjects with no ACL injury. For this group, sagittal knee joint signal time domain data measurements were captured on the right limb only. For the **ACLR** subjects, sixteen sagittal knee joint signal data measurements were captured on the injured limb **(ACLI** group**)** and sixteen measurements were captured on the non-injured **(ACLNI** group**)** limb.

## 2.2 Study Design and Measurement Protocol

Each participant ran at a self-selected speed ($control$ $2.7 \pm 0.4$ $m/s$; $ACLR$ $2.7 \pm 0.3$ $m/s$). Fifty-six reflective markers were placed on each subject using an established design protocol. Sagittal plane knee joint kinematics time domain data were extracted from the processed marker data and converted to a frequency domain representation using FFT. The FFT converts a discrete time domain signal into the frequency domain by representing the signal as a series of sinusoids. Converting the signal from the time to frequency domain does not alter the signal in any way but rather highlights certain characteristics; such as, the amplitude, frequency and phase of the signal that were not as apparent and easy to measure in the time domain. Figure 1 provides a plot of sagittal knee flexion (deg) measurements in the time domain and its related frequency domain data for a subject in the control group and the ACL reconstructed group. It is interesting to note the differences in the general patterns of the plots of the time domain and frequency domain measurements. Power and phase spectrums were generated for each signal using a custom MATLAB code. The power spectrum revealed that most of the signal energy was contained in the two dominant sinusoids. Dominant means the sinusoids with the largest amplitude were selected out. Differences in amplitude, frequency and phase represent kinematics deviations in gait biomechanics between healthy and ACLR individuals.

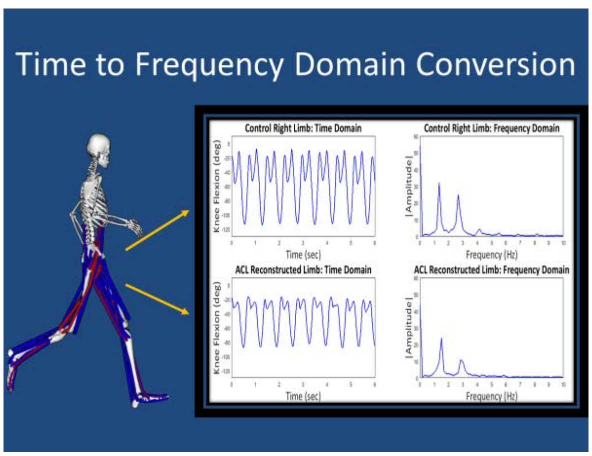


**Figure 1:** Plot of the time domain and frequency domain sagittal knee flexion measurements for subjects in the CTLR control group and the ACLI reconstructed group

Thus, the amplitude, frequency, and phase components of the ACLR and control limbs were chosen for this investigation. The amplitude computed from the FFT represents the magnitude of the oscillations of the knee flexion signal (Robertson et al. 2014). The FFT frequency describes how fast the signal oscillates, which in our study, reveals how fast the knee is oscillating during the gait cycle. The phase component of the FFT depicts when a time shift or a delay in the signal occurs, which illustrates when a time shift in gait events happens; for example, when a shift in peak knee flexion occurs.

Fifteen FFT or spectral measurements were captured on each member of the study groups. The measurements recorded were the frequency, phase, and amplitude measurements for the first five FFT signals. The candidate study measurements variables obtained are listed below in Table 1.

**Table 1:** Candidate FFT Predictor Variables for the Linear Discriminant Analysis Model

- 1st Frequency (*Hz*), 1st Phase (*Deg*), and 1st Amplitude $\left(\frac{Deg^2}{Hz}\right)$
- 2nd Frequency (*Hz*), 2nd Phase (*Deg*), and 2nd Amplitude $\left(\frac{Deg^2}{Hz}\right)$
- 3rd Frequency (*Hz*), 3rd Phase (*Deg*), and 3rd Amplitude $\left(\frac{Deg^2}{Hz}\right)$
- 4th Frequency (*Hz*), 4th Phase (*Deg*), and 4th Amplitude$\left(\frac{Deg^2}{Hz}\right)$
- 4th Frequency (*Hz*), 5th Phase (*Deg*), and 5th Amplitude $\left(\frac{Deg^2}{Hz}\right)$

The 1st Frequency and 1st Phase predictor variables were not used in the analyses because the variable values were the same constant value for each of the subjects in the three groups. Hence, only thirteen candidate predictor variables were used.

## 3. Linear Discriminant Analysis

### 3.1 Overview

Given that multiple measurements were obtained on each subject in the study, discriminant analysis, a multivariate analysis technique, was employed in this research exploration. Discriminant function analysis is a statistical method used to predict a categorical dependent variable or grouping variable using one or more continuous or categorical independent or predictor variables. The method is useful in determining whether a set of variables is effective in predicting category or group membership. Discriminant analysis is used in statistics, pattern recognition, supervised learning or supervised classification to find a linear combination of features that characterizes or separates two or more classes of objects or events. It is a classification technique, where two or more groups or clusters or populations are known *a priori* and one or more new observations are classified into one of the known populations based on the measured characteristics. The goal is to find a classification rule(s) that minimizes the total error of classification, i.e. makes the proportion of subjects misclassified as small as possible. The resulting combination may be used as a linear classifier, or, quite often for dimensionality reduction before later classification. If the groups are linearly separable, then the groups can be separated by a linear combination of features that describe the objects. If only two features are used, the separators between the object groups will become lines. If there are three features, the separator is a plane and if the number of features (i.e., independent variables) is more than 3, the separators become a hyper-plane.

With linear discriminant analysis, all groups are assumed to have the same covariance matrix. Hence, the assumption is the variance-covariance matrices are homogeneous for the g groups, i.e.,

$$\sum\nolimits_1 = \sum\nolimits_2 = \cdots = \sum\nolimits_g = \sum$$

In this case the variance-covariance matrix does not depend on the population from which the data are obtained. A quadratic discriminant analysis is used for heterogeneous variance-covariance matrices, i.e.,

$$\sum\nolimits_i \neq \sum\nolimits_j \quad for\ some\ i \neq j$$

This allows the variance-covariance matrices to depend on which population we are considering. The analysis is quite sensitive to outliers and the size of the smallest group must be larger than the number of predictor variables. The assumptions of discriminant analysis are

- Multivariate normality: Independent variables are normal for each level of the grouping variable.
- Homogeneity of variance/covariance (homoscedasticity): Variances among group variables are the same across levels of predictors. Equality of the variance-covariance matrices for the groups can be tested using Bartlett's test. Linear discriminant analysis may be used when the covariances are equal, and quadratic discriminant analysis may be used when covariances are not equal.
- Multicollinearity: Predictive power can decrease with an increased correlation between predictor variables.
- Independence: Participants are assumed to be randomly sampled, and a participant's score on one variable is assumed to be independent of scores on that variable for all other participants.

It has been suggested that discriminant analysis is relatively robust to slight violations of these assumptions, and it has also been shown that discriminant analysis may still be reliable when using dichotomous variables.

### 3.2 Discriminant Analysis Procedure

A key statistical assumption for linear discriminant analysis is that the predictor variables are normally distributed (i.e., each variable is shaped like a bell curve when plotted) and that each attribute has the same variance (i.e., the values of each variable vary around the mean by the same amount on average). LDA makes predictions by estimating the probability that a new set of inputs belongs to each class or group. The class that gets the highest probability is the output class and a prediction is made.

An exploratory data analysis was conducted to validate the assumptions that the predictor variables exhibit a Gaussian distribution. Figure 2 provides univariate histogram plots of the 1$^{st}$ amplitude, 2$^{nd}$ amplitude, 3$^{rd}$ amplitude, 4$^{th}$ amplitude, and 5$^{th}$ amplitude variables for each of the three groups. Figure 3 provides univariate histogram plots of the 2$^{nd}$ frequency, 3$^{rd}$ frequency, 4$^{th}$ frequency, and 5$^{th}$ frequency variables for each of the three groups. Figure 4 provides univariate histogram plots of the 2$^{nd}$ phase, 3$^{rd}$ phase, 4$^{th}$ phase, and 5$^{th}$ phase variables for each of the three groups. Histograms are not provided for the 1$^{st}$ Frequency and 1$^{st}$ Phase data since all the data values were the same for each of these variables for all groups. The variables appear to have a Gaussian distribution with comparable variances.

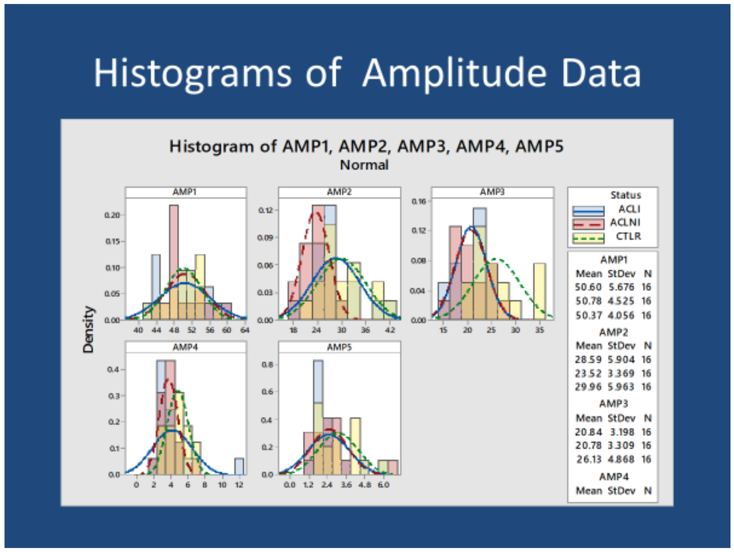


**Figure 2:** Histogram plots of the Amplitude Data for each of the three populations – ACLI, ACLNI and CTLR.

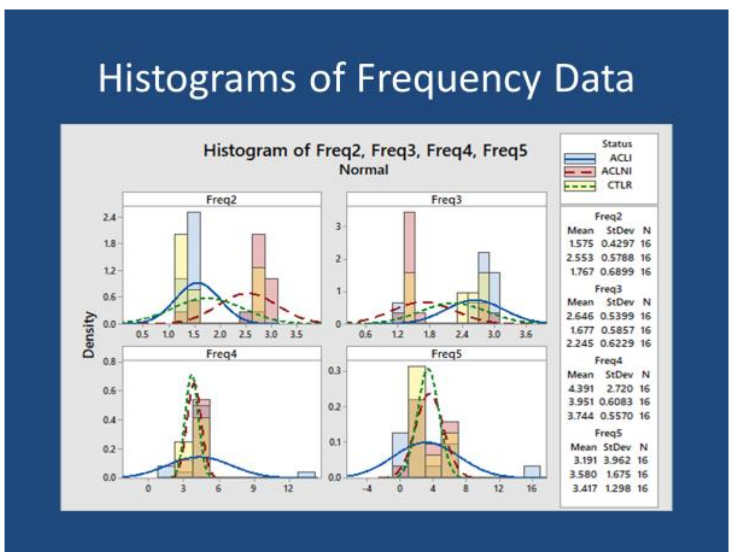


**Figure 3:** Histogram plots of the Frequency Data for each of the three populations – ACLI, ACLNI and CTLR.

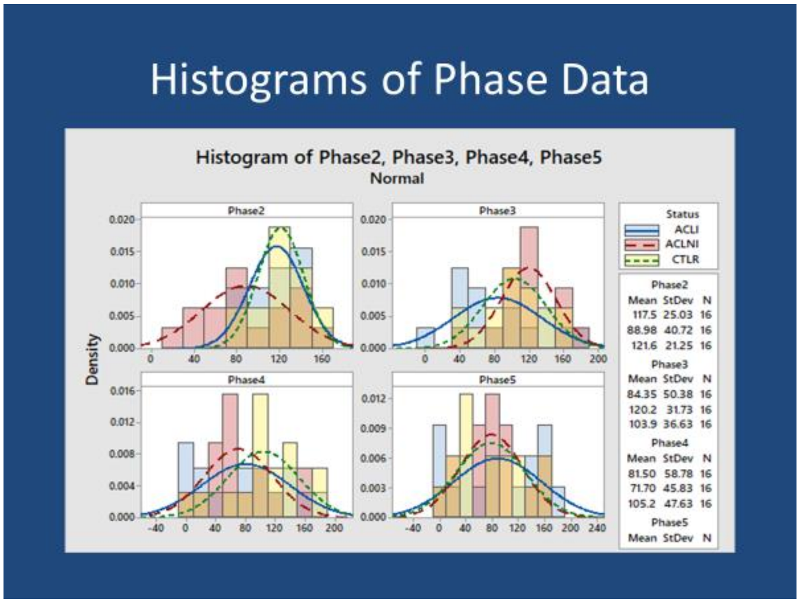


**Figure 4:** Histogram plots of the Phase Data for each of the three populations – ACLI, ACLNI and CTLR.

Since there were 13 candidate predictor variables, LDA investigations were conducted using all 13 of the predictor variables as well as selected subsets of the variables. The results of the investigations with the most promising findings will be presented in this paper. Bartlett's test was used to test the null hypothesis that the variance-covariance matrices for the predictor variables used in the model are homogeneous for the three population groups. If the null hypothesis is rejected, a quadratic discriminant analysis will be fit to the data rather than a linear one. The plots and the LDA analysis were performed in Minitab 18.

The description of the analysis method and procedures used in the Minitab 18 analyses are provided below. In linear discriminant analysis, an observation **x** is classified into a group if the squared distance, also called the Mahalanobis distance,

$$d_t^2(\mathbf{x}) = (\mathbf{x} - \mathbf{m}_t)'\mathbf{S}_p^{-1}(\mathbf{x} - \mathbf{m}_t)$$

of the observation **x** to the group **t** center or mean is the minimum. The generalized squared distance from **x** to group **t** for the linear discriminant function is given as

$$d_t^2(\mathbf{x}) = d_t^2(\mathbf{x}) - 2\ln(q_t)$$

In addition, there is a unique squared distance formula for each group and that is called the linear discriminant function for the group **t**. This function corresponds to the regression coefficients in multiple regression and is given by

$$\mathbf{m}_i'\mathbf{S}_p^{-1}\mathbf{x} - 0.5\mathbf{m}_i'\mathbf{S}_p^{-1}\mathbf{m}_i + \ln p_i$$

where

| | |
|---|---|
| **x** | column vector of length p containing the values of the predictors for this observation (this column vector is stored as one row) |
| p | number of predictors |
| n | total number of observations |
| t | group subscript |
| $n_t$ | number of observations in group t |
| $q_t$ | the prior probability of group t, which equals $n_t/n$ |
| $\mathbf{S}_p$ | pooled covariance matrix for linear discriminant analysis |
| $\mathbf{S}_i$ | covariance matrix of group i for quadratic discriminant analysis |
| $\mathbf{m}_t$ | column vector of length p containing the means of the predictors calculated from the data in group t |
| $\mathbf{S}_t$ | covariance matrix of group t |

For any given observation **x**, the group with the smallest squared distance has the largest linear discriminant function and the observation is then classified into this group. In some cases, subjects from different groups are encountered according to different probabilities. With the assumption that the data have a normal distribution, the linear discriminant function is increased by ln $(p_i)$ where $p_i$ is the prior probability of group i. Since the groups were of the same size (i.e., $n_i = 16, i = 1, 2, and\ 3\ and\ n = 48$), then $p_i = n_i/n = 1/3$ and

$\ln(p_i)$ is the natural log of the prior probability for group $i$

In comparing the performance of the proposed models, it will be important to assess how well the LDA model has performed in predicting the group for each observation. Thus, we will evaluate the proportion of observations correctly placed into their true group using the model. Quite often the percent of misclassified observations is optimistic because the data being classified are the same data used to build the classification function or model. The cross-validation technique works by omitting each observation one at a time, recalculating the classification function using the remaining data, and then classifying the omitted observations. This technique is used to compensate for a possible overly optimistic error rate or percent of misclassified observations. The cross-validation rates for each model will also be presented. We will also compare the distance values of the variables for the three groups and the magnitude of the regression coefficients for the variables in the LDA model.

## 4. Results

Our goal was to identify the best LDA model with the most parsimonious number of predictor variables. The first LDA analysis was conducted using all thirteen of the candidate predictor variables described above. Although linear discriminant analyses were conducted using other subsets of the thirteen variables, Table 2 only provides the results for a selected subset of the analyses performed.

**Table 2:** Summary of the Proportion of Correct Classifications – Overall and by Group – Without and With Cross-Validation for each Linear Discriminant Model

| **Predictor Variables Used in LDA** | **Proportion of Correctly Classified For all groups Without Cross-validation (with Cross-validation)** | **Proportion of ACLI Correctly Classified Without Cross-validation (with Cross-validation)** | **Proportion of ACLNI Correctly Classified Without Cross-validation (with Cross-validation)** | **Proportion of CTLR Correctly Classified Without Cross-validation (with Cross-validation)** |
|---|---|---|---|---|
| **All Thirteen Variables** | 0.854 (0.646) | 0.938 (0.563) | 0.813 (0.688) | 0.813 (0.688) |
| **2nd Frequency 3rd Frequency 4th Amplitude 5th Amplitude** | 0.708 (0.604) | 0.938 (0.813) | 0.750 (0.688) | 0.438 (0.313) |
| **2nd Frequency 3rd Frequency 3rd Phase** | 0.688 (0.604) | 0.875 (0.750) | 0.750 (0.750) | 0.438 (0.313) |
| **2nd Frequency 3rd Frequency** | 0.625 (0.625) | 0.750 (0.750) | 0.750 (0.750) | 0.375 (0.375) |

Note, when all thirteen variables were used in the LDA, 85.4% of all the observations were correctly placed, 93.8% of the observations in ACLI group were correctly placed, 81.3% of the observations in the ACLNI group were correctly placed, and 81.3% of the observations in the CTLR group were correctly placed. Thus, classifying observations into the ACLNI and CTLR groups has the most challenges when using all thirteen variables. However, when cross-validation was used, 64.6% of all the observations were correctly placed, 56.3% of the observations in ACLI group were correctly placed, 68.8% of the observations in the ACLNI group were correctly placed, and 68.8% of the observations in the CTLR group were correctly placed. Again, classifying observations into the ACLNI and CTLR groups has the most challenges when using all thirteen variables. This challenge was also noted when using the other LDA models. It is particularly interesting to see that the CTLR group has the lowest proportion of correct classifications. The most notable findings were the large difference in the proportion of correctly classified observations in each group using cross-validation when all thirteen variables were in the model. This supported the need to explore other possible subsets of the thirteen variables to be used for the LDA. The additional models that yielded the best results are listed in Table 2 and discussed further below.

For LDA it is also important to identify those variables that contribute most to the classification of observations into each group and review the squared distance from one group center (mean) to another group center (mean). It is important to compare the distances to see how different the groups are. Table 3 provides information on the differences between the squared distance for two groups and the variables that contributed the most for each LDA model. For example, when all thirteen of the variables are used in the LDA model, Table 3 highlights that the greatest distance is between the ACLI and CTLR groups (7.903). The difference between the ACLNI and CTLR groups is 7.350, and the smallest difference (5.469) is between the groups ACLI and ACLNI. This is not surprising given that the ACLI and CTLR populations should demonstrate vastly different gait patterns and the patterns for the ACLI and ACLNI groups should not differ as drastically. In reviewing the results in Table 3 for the LDA model with all thirteen variables, the ACLI group has the largest linear discriminant function for 2$^{nd}$ Amplitude, 3$^{rd}$ Frequency, 3$^{rd}$ Phase, 4$^{th}$ Amplitude, 5$^{th}$ Amplitude, 5$^{th}$ Frequency, and 5$^{th}$ Phase, which indicates that these variables for the ACLI group contribute more than those of the ACLNI group or CTLR group to the classification of group membership. Also, group ACLNI has the largest linear discriminant function for 1$^{st}$ Amplitude, 2$^{nd}$ Frequency, 4th Frequency, and 4$^{th}$ Phase which indicates that these variables for the ACLNI group contribute more than those of the ACLI group or CTLR group to the classification of group membership. The CTLR group has the largest linear discriminant function for 2$^{nd}$ Phase and 3$^{rd}$ Amplitude which indicates that these variables for the CTLR group contribute more than those of the ACLI group or ACLNI group to the classification of group membership.

**Table 3: Summary of the Difference for the Squared Distance between the Groups and the Variables that Contributed Most to the Classification of Group Membership**

| Predictor Variables Used in LDA | Difference For the Squared Distance Between Groups | Variable(s) that Contributed Most To Classification of Group Membership |
|---|---|---|
| **All Thirteen Variables** | ACLI and ACLNI (5.469)<br><br>ACLI and CTLR (7.903)<br><br>ACLNI and CTLR (7.350) | ACLI to<br>2nd Amplitude<br>3rd Frequency<br>3rd Phase<br>4th Amplitude<br>5th Amplitude<br>5th Frequency<br>5th Phase<br><br>ACLNI to<br>1st Amplitude<br>2nd Frequency<br>4th Frequency<br>4th Phase<br><br>CTLR to<br>2nd Phase<br>3rd Amplitude |
| **2nd Frequency<br>3rd Frequency<br>4th Amplitude<br>5th Amplitude** | ACLI and ACLNI (3.176)<br><br>ACLI and CTLR (1.926)<br><br>ACLNI and CTLR (2.977) | ACLNI to 2nd Frequency<br>ACLI to 3rd Frequency<br>ACLI to 4th Amplitude<br>ACLI to 5th Amplitude |
| **2nd Frequency<br>3rd Frequency<br>3rd Phase** | ACLI and ACLNI (2.966)<br><br>ACLI and CTLR (1.125)<br><br>ACLNI and CTLR (2.710) | ACLNI to 2nd Frequency<br>ACLI to 3rd Frequency<br>CTL to 3rd Phase |
| **2nd Frequency<br>3rd Frequency** | ACLI and ACLNI (2.934)<br><br>ACLI and CTLR (1.132)<br><br>ACLNI and CTLR (2.475) | ACLNI to 2nd Frequency<br>ACLI to 3rd Frequency |

For the other LDA models in the table, the largest two coefficient values in the model are for the 2nd Frequency and 3rd Frequency variables. In most instances, the magnitude of the two coefficient values are about the same. The 2nd Frequency variable contributed most for the ACLNI group and the 3rd Frequency variable is identified as the variable that contributed the most for the ACLI variable for all the models listed in Table 3.

The remainder of the discussion will focus primarily on the model with only the 2nd Frequency and 3rd Frequency variables since the proportion of correct classifications for all groups, the ACLI group, the ACLNI group, and the CTLR group with and Without Cross-validation are the same. Given that the 2nd Frequency and 3rd Frequency variables appeared to follow a normal distribution, Bartlett's test was used to determine if the variance-covariance matrices are homogeneous for the three populations involved. No significant difference between the variance-covariance matrices for the three populations was found (Bartlett's test statistic $L' = 2.38$ d.f. = 6; $p = 0.881$). Thus, linear discriminant analysis was appropriate for the data.

## 5. Discussion and Conclusions

This research was conducted to demonstrate the use of fast Fourier transform and control theory techniques to obtain frequency, phase, and stability metrics to characterize the gait in healthy individuals and individuals who have undergone an anterior cruciate ligament (ACL) reconstruction. In addition, the goal was to assess the use of linear discriminant analysis to identify the critical variables that best classify the healthy and injured populations and develop a model that describes the relationship between these variables that are very critical to identifying individuals with adverse gait biomechanics. Using control theory and FFT techniques, we could identify, quantify, and stratify gait factors that delineate the three groups of subjects in the study. It does indeed appear that the thirteen candidate FFT predictor variables did perform well in capturing important performance metrics for classifying the three study population groups.

It is particularly interesting to see that the CTLR group has the lowest proportion of correct classifications. It will be important to follow these subjects who are currently in the CTLR group and assumed to be "healthy" to see if our model has identified subjects who will potentially encounter a gait problem. We are interested in predicting future gait problems and this model may helpful in doing this.

The most notable findings were the large difference in the proportion of correctly classified observations in each group without and with cross-validation using all thirteen of the FFT variables. This supported the need to explore other possible subsets of the thirteen variables to be used for the LDA to see if this large difference would disappear with other subsets of the variables. Although this large difference did not reveal itself when the LDA model with only the 2nd Frequency and 3rd Frequency variables was used, the proportion of correctly classified subjects was low.

It was remarkable how the 2nd Frequency variable contributed most for the ACLNI group and the 3rd Frequency variable was identified as the variable that contributed the most for the ACLI variable for all the models listed in Table 3. In summary, linear discriminant analysis performed well to identify a classification model using FFT data. Also, the LDA model results are consistent with the findings obtained using the nominal logistic regression model in another study. Both statistical techniques selected the 2nd and 3rd Frequency metrics as key variables that contribute to group separation. The results of the research will allow us to potentially predict individuals with future injury risk and design more targeted ACL prevention and rehabilitation programs. Future work will involve conducting more studies and collecting data to validate and build on the LDA classification model results.

## Acknowledgements

"Portions of information contained in this publication/book are printed with permission of Minitab Inc. All such material remains the exclusive property and copyright of Minitab Inc. All rights reserved."

All data processing was performed off-line using a commercial software package (MATLAB 6.1, The MathWorks Inc., Natick, MA, 2000).

## References

Almekinders, L. C., Moore, T., Freedman, D., and Taft, T. N. (1995). Post-operative problems following anterior cruciate ligament reconstruction. *Knee Surgery, Sports Traumatology, Arthroscopy*, *3*(2), 78-82.

Ardern, C. L., Taylor, N. F., Feller, J. A., & Webster, K. E. (2014). Fifty-five per cent return to competitive sport following anterior cruciate ligament reconstruction surgery: an updated systematic review and meta-analysis including aspects of physical functioning and contextual factors. *Br J Sports Med*, *48*(21), 1543-1552.

Boden, B. P., Griffin, L. Y., & Garrett Jr, W. E. (2000). Etiology and prevention of noncontact ACL injury. *The Physician and sportsmedicine*, *28*(4), 53-60.

DeFontenay, B.P., Argaud, S., Blache, Y. and Monteil, K. Motion Alterations After Anterior Cruciate Ligament Reconstruction: Comparison of the Injured and Uninjured Lower Limbs During a Single-Legged Jump. *Journal of Athletic Training*, 49(3), 311-316, 2014.

Donnelly, C. J., Elliott, B. C., Ackland, T. R., Doyle, T. L. A., Beiser, T. F., Finch, C. F., ... & Lloyd, D. G. (2012). An anterior cruciate ligament injury prevention framework: incorporating the recent evidence. *Research in sports medicine*, *20*(3-4), 239-262.

Giakas, G., and Baltzopoulos, V. Time and Frequency domain analysis of ground reaction forces during walking: An investigation of variability and symmetry. *Gait and Posture*, 5 (3), 189-197, 1997.

Gao, B. and Zheng, N.N. Alterations in three-dimensional joint kinematics of anterior cruciate ligament-deficient and -reconstructed knees during walking. *Clinical Biomechanics*, 25, 222-229, 2010.

Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com)

Morgan, K. D., Zheng, Y., Bush, H., and Noehren, B. (2016). Nyquist and Bode stability criteria to assess changes in dynamic knee stability in healthy and anterior cruciate ligament reconstructed individuals during walking. *Journal of Biomechanics*, *49*(9), 1686-1691.

Morgan, K. D., Donnelly, C. J., and Reinbolt, J. A. (2014). Elevated gastrocnemius forces compensate for decreased hamstrings forces during the weight-acceptance phase of single-leg jump landing: implications for anterior cruciate ligament injury risk. *Journal of Biomechanics*, *47*(13), 3295-3302.

Osgood, Brad. “The Fourier Transform and its Applications.” Electrical Engineering 261 class at Stanford University, Stanford. 2007. Lecture notes.

Shabani, B., Bytyqi, D., Lustig, S., Cheze, L., Bytyqi, C., and Neyret, P. (2015). Gait knee kinematics after ACL reconstruction:3D assessment. *International Orthopedics*, 39(6), 1187-1193.

Stergiou, N., Giakas, G., Byrne, J.E., and Pomeroy, V. (2002). Frequency domain characteristics of ground reaction forces during walking of young and elderly females. *Clinical Biomechanics.* 17, 8, 615-617.