RT-ACL: Identification of High-Risk Youth Patients and their Most Significant Risk Factors to Reduce Anterior Cruciate Ligament Reinjury Risk

Amanda Watson*, Pengyuan Lu*, Elliot Greenberg+, J. Todd R. Lawrence+,
Theodore J. Ganley+, Insup Lee*, James Weimer*
*Dept. of Computer and Information Science, University of Pennsylvania
{aawatson, pelu, lee, weimerj}@seas.upenn.edu
+Sports Medicine and Performance Center, Children's Hospital of Philadelphia
{greenberge, lawrencej, ganley}@chop.edu

***Abstract*—The anterior cruciate ligament (ACL) is the most commonly injured ligament in the body, accounting for more than 200,000 ACL tears occurring annually in the US and upwards of 90% of patients choosing to undergo reconstruction surgery. After the reconstruction surgery, approximately 30% of youth patients go on to re-tear their ACL, and it has been proven that repeated ACL reconstructions not only have inferior results but can be devastating for all those involved. In this paper, we propose RT-ACL, a system that enhances patient outcomes by reducing their risk of an ACL re-tear by providing personalized recommendations of modifiable risk factors that can be altered during the patient's recovery process. Our system leverages the RT-ACL model that uses labeling functions designed by clinicians to classify the risk level of an ACL re-tear. Further, it identifies modifiable risk factors and suggests interventions to minimize adverse outcomes and complications. We evaluated our system on a dataset of 441 youth patients, 8-21 years of age, that underwent an ACL reconstruction at the Children's Hospital of Philadelphia. The results indicate patients classified as low risk re-tear at a rate of 12%, medium risk at a rate of 30%, and high risk re-tear at a rate of 59%. This demonstrates those classified by our system as high risk are 4.6 times as likely to re-tear their ACL than those classified as low risk.**

I. Introduction

The anterior cruciate ligament (ACL) functions to stabilize the knee by resisting the combined motions of anterior tibial translation and internal tibial rotation, thus providing rotational stability to the knee [1]. Injuries to the ACL are common, and if torn, the injured knee has a significant risk of instability, increased risk of injury to other parts of the knee, and future osteoarthritis [2]. Reconstruction of a torn ACL is a surgical procedure that restores stability to the knee and decreases the risk of subsequent injury. Approximately 200,000 ACL reconstructions are performed annually in the United States, costing more than two billion dollars; inclusive of surgery, rehabilitation, physical therapy, etc. [3]. Recovery from an ACL reconstruction generally takes more than six months and requires a significant financial and time investment from the patient [4]. Therefore, a repeat tear of the ACL is devastating for the patient as well as their family, trainer, and surgeon. re-tears occur in approximately 20% of patients [5], and it has been demonstrated that repeated ACL reconstructions have inferior results compared with the original [6].

ACL tears in youth patients have been rising [7] due to increased and earlier participation in year-round sports [8], higher clinical awareness of and recognition of the symptoms of ACL tears [9], and diagnostic technologies such as magnetic resonance imaging [10], [11]. Additionally, age is a recognized risk factor for the re-tear of an ACL, with younger patients being at a higher risk, approximately 30% compared to 4-17% [12]–[15]. Further top-level athletes, such as those competing at the National Collegiate Association Division 1 level, have even higher re-tear rates at 37% [16]. The high reinjury rates in younger patients have drawn significant concerns from the research community prompting further exploration to determine risk factors and the design mitigation strategies [17].

To date, researchers have identified risk factors [18]–[20], created clinical protocols [18], [21], [22], and developed technical algorithms [23], [24] to decrease the number of re-tears after ACL reconstruction. Risk factors aid in the identification of weaknesses in individual patients. However, interactions between the risk factors still need to be explored to gain a holistic understanding of the overall risk of specific patients. Clinical protocols have been developed to determine patients risk at the time of discharge. Similar to risk factors, the clinical discharge protocols are highly variable and specialized to specific subgroups of patients and clinical environments. To our knowledge, a system that holistically evaluates a patient's overall risk, identifies modifiable risk factors, and suggests the most significant risk factors to clinicians to develop intervention methods from does not exist.

In this paper, we propose to identify patients at high risk for re-tear after ACL reconstruction by developing the RT-ACL model that leverages domain knowledge to classify the risk level of a patient. Domain knowledge is collected from clinicians and encoded in the form of labeling functions that classify high vs. low risk of re-tear. These labeling functions are weighted based on their calculated risk of re-tear. The RT-ACL model outputs an overall vote used to determine a patient's risk level: high, medium, or low.

978-1-6654-3965-7/21/$31.00 ©2021 IEEE
DOI 10.1109/CHASE52844.2021.00012

Next, we evaluate the proposed classifier on a dataset obtained from the Children's Hospital of Philadelphia. The dataset includes 442 youth patients, 8-21 years of age, that have undergone an ACL reconstruction surgery. Consistent with previous studies [12], [13], 27% of the patients in our dataset have gone on to re-tear their ACL after their reconstruction surgery. The dataset includes patient demographics, injury information, family history, surgery information, recovery information, re-tear information, and rehabilitation information. The results indicate patients classified as high risk re-tear at a rate of 59%, medium risk at a rate of 30%, and low risk at a rate of 12%. This demonstrates high-risk classifications are 4.6 times as likely to re-tear than low-risk classifications.

Finally, we create a feedback system that provides personalized recommendations for modifiable risk factors. These risk factors can be altered during the patient's recovery period to decrease their risk of re-tear. The most significant risk factors are chosen based on their impact and the ease with which they can be modified. The weight of each risk factor and category from the RT-ACL model determines the impact, and the clinicians set the ease with which each risk factor can be modified. The top five risk factors are chosen and displayed to the clinicians for interpretation and conveyance back to the patients.

The goal of this paper is to develop a system that can be utilized to reduce the risk of ACL re-tears after reconstruction. To achieve this goal, this system should holistically evaluate a patient's overall risk, identify modifiable risk factors, and suggest the most significant risk factors to clinicians to develop intervention methods from to reduce the risk of complications.

Specifically, the contributions of this work are as follows:

1) Development of the RT-ACL model that identifies youth patients at high risk for re-tear by leveraging domain knowledge in the form of labeling functions designed by clinicians.
2) An evaluation of the RT-ACL model on a dataset obtained from the Children's Hospital of Philadelphia and a comparison with other algorithms.
3) A clinical decision support system that provides personalized recommendations for modifiable risk factors to lessen the risk of an ACL re-tear.

The remainder of this paper is structured as follows: Section II summarizes the work related to this paper. In Section III, we formulate our problem. Then, Section IV describes the the RT-ACLSystem. In Section V, we evaluate our method and compare it to other state-of-the-art methods. Finally, we conclude our paper in Section VII.

II. Related Work

In recent years, the high reinjury rate of patients who have undergone ACL reconstruction surgery has drawn the interest of the research community, particularly in younger patients. This section discusses the results of that interest, including the identification of risk factors, the establishment of clinical protocols, and the development of technical algorithms.

A. Risk Factor Identification

The first step to understanding the high reinjury rate of ACL reconstruction patients is to identify risk factors. A systematic review of risk factors associated with ACL reconstructions over 66 studies identified risks factors over three categories: technical factors, patient-related factors, and the factors associated with the status of the knee joint [25]. Technical factors are factors related to the graft itself, such as graft size and tension. Patient-related factors are factors related to a patient's demographics, such as age and BMI. Status of the knee joint takes into account other parts of the knee, such as meniscus damage. Further, research has been done into kinematic factors that are tested throughout a patient's recovery process, such as strength [26], mobility [27], [28], and stability [29], [30]. These risk factors help identify shortcomings in individual patients. Additional analysis should combine risk factors and analyze their interactions to fully understand the patient's overall risk. RT-ACLleverages risk factors such as those discussed here to determine a patient's overall risk. Additionally, we combine multiple risk factors such as age and sex to further capture a patient's risk.

B. Clinical Clearance Protocols

Once risk factors have been established, clinical protocols for clearance to return to full activity can be developed. Surveys [31]–[33] have highlighted this problem and noted that protocols use factors such as isokinetic testing, time since surgery, and functional performance. While clinical protocols lead to better outcomes, they can be highly variable and specific. Grindem et al. [24] created a discharge protocol that leveraged quadriceps strength, hop testing, and symmetry scores to determine if a patient was at high risk for re-tear or other injuries at the time of return to sport. In a study of male professional soccer players [22], a discharge protocol was established that tested players' isokinetic strength, dynamic running, and functional hopping. These studies found that a patient who did not meet all the requirements of these protocols would be at four times higher risk than those who did. As with risk factors, these clinical discharge protocols are highly variable and are frequently specialized to specific subgroups of patients or clinical equipment. RT-ACLseeks to account for subgroups of patients with additional features and risk factors. For example, we use sport played as a risk factor. This allows for a more versatile re-tear prediction system.

C. Technical Algorithms

Further, technical contributions have been made via machine learning and data analytics. In a study of 503 athletes, Nguyen et al. [34] used a multivariate logistic regression over sex, age, and graft size while keeping other variables such as surgical technique and rehabilitation protocol stable. They found that female patients younger than 25 with a graft size less than 8 mm were at higher risk of a re-tear. Paterno et al. [23] used classification and regression tree to classify the risk of re-tear for patients based on clinical measures such as age, sex, knee-related confidence, and performance of triple hop

for distance. Their results show that their high-risk group is five times more likely to suffer a second ACL injury. In a study of over 700 high school and college-age athletes (14-22 years), the MOON knee group used a multivariable regression model to demonstrate a hamstring autograft was 2.1 times more likely to fail than a bone tendon bone autograft [35]. While their study shows the predictive ability of non-modifiable risk factors, it does not account for postoperative factors that can be modified to support a better rehabilitation outcome, such as neuromuscular training, changes in activity level, and time to return to sport. In RT-ACL, we account for these factors allowing us to present a more holistic view of the patient's recovery which can have an even greater contribution to the risk prediction of graft rupture. These models establish groupings of patients that are at higher risk for re-tear. While this information is useful and can be incorporated into RT-ACLas additional risk factors, it is not always helpful to each individual patient. RT-ACLassesses each patient's risk of re-tear based on their individual statistics on their risk factors.

III. Preliminaries and Problem Formulation

This section describes the dataset used for training and evaluation in this paper, the labeling functions used to encode clinical knowledge, and the formulation of the problem we solve. First, we discuss our dataset and its limitations. Next, we describe the expert knowledge on the risk of ACL re-tear collected from the clinicians and our encoding method. Lastly, we present the problem we address in this research.

TABLE I: Dataset Information

Category	#	Missing	Example
Demographics	6	5%	Age, DOB
Injury Inforation	2	2%	Date, Sport Played
Family History	2	43%	Relative with ACL Tear?
Surgery Information	20	9%	Type of Reconstruction
Recovery Information	2	14%	Date of Release to Activity
Re-tear Information	7	59%	Time to Repeat ACL Tear
Rehab Information	213	79%	Triple Hop LSI

A. Dataset

This study was approved by the Institutional Review Board. Our dataset was collected in a retrospective analysis of 441 youth patients, 8-21 years of age, that have undergone a primary ACL reconstruction surgery. Overall, in our entire dataset, 27% of patients went on to re-tear their ACL. This data was sampled from a more extensive database of patients from the Children's Hospital of Philadelphia, over which there was approximately a 15% re-tear rate. We discuss the disparity in these numbers in the following paragraph. Overall, there are 223 males and 218 females in this dataset. In total, there are 252 features that we break down into seven categories as shown in Table I. Individual patients had different approaches to their rehabilitation, causing many patients in our dataset to have missing data points. This missing data will need to be accounted for in our system. We show the amount missing data per category in Table I.

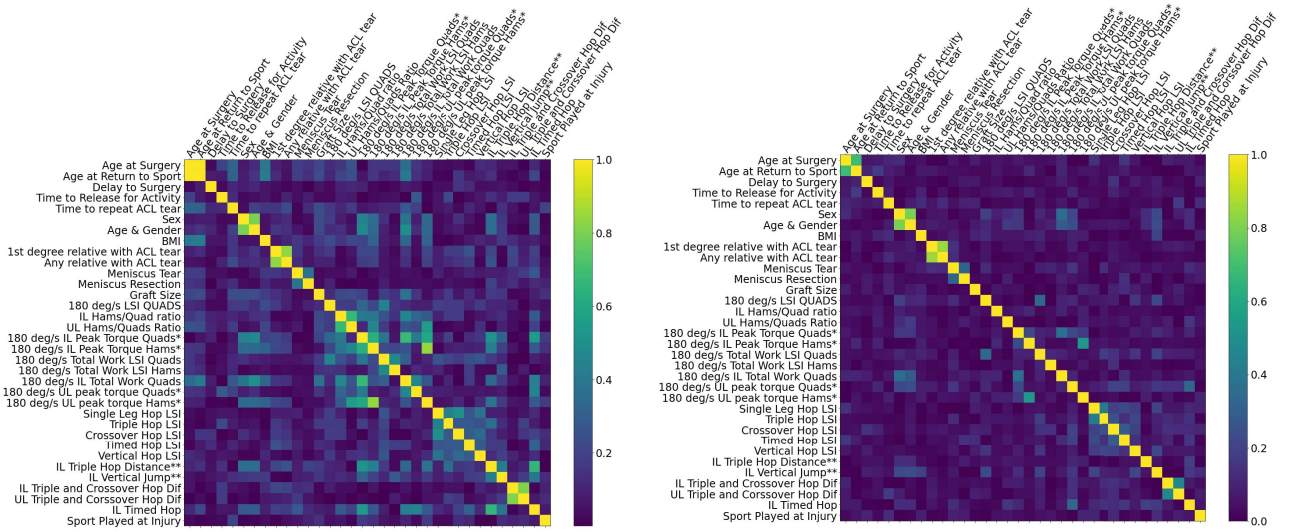
We split our data $Z = (X, Y)$ into three portions $Z_1 = (X_1, Y_1)$ and $Z_2 = (X_2, Y_2)$, and $Z_3 = (X_3, Y_3)$. Z_1 is the training dataset comprised of 60% of our total data. Z_2 is the testing dataset comprised of 20% of our dataset. Z_3 is comprised of 20% of our dataset and will be used in our evaluation as a hold-out set. Since there is a class imbalance between patients who re-tear and those who did not, we asked for a larger portion of re-tears for our training dataset, Z_1. The percent of re-tears in each dataset are as follows- Z_1: 32%, Z_2: 32%, Z_3: 6%. Overall, 32% of our test set, Z_2, or 28 patients go on to re-tear. This ensures that we will have re-tears in the high, medium, and low categories. Our hold-out set has a re-tear rate of 6%; this gives us 5 patients for evaluation between the three risk categories.

B. Labeling Functions

As shown above, our dataset includes a high number of features. We collected a list of the most important features or combinations of features that we will refer to as risk factors from our clinicians. Next, we asked clinicians to develop a set of labeling functions such that for each risk factor, labeling functions were developed to distinguish between high and low-risk patients. Additionally, any patient not encompassed in the high or low risk is considered unknown. This can be due to missing data or a patient being outside of the bounds set by the high and low-risk labeling functions.

In total, the clinicians distinguished 34 risk factors across all seven categories of data. To understand the correlations between these risk factors, we calculate the correlation coefficient between each pair of risk factors. The results of this are shown in Figure 1a. In this figure, negative one indicates a negative linear correlation, zero indicates no linear correlation, and one indicates a positive linear correlation. Most pairs of risk factors show no to low linear correlation. There are three types of correlated pairs in our system: timing pairs, encapsulated pairs, and symmetry pairs.

- **Timing Pairs** are pairs of correlated risk factors that occur in the same time frame. For example, *Age at Surgery* and *Age at Return to Sport* are expected to occur between six months and a year of each other. This is due to recovery processes being approximately six months.
- **Encapsulated Pairs** are pairs of correlated risk factors where one risk factor is a subset of another. For example, *Sex* and *Age & Gender* or *1st degree relative with ACL tear* and *Any relative with ACL tear* encompass one other so their correlation is expected to be high.
- **Symmetry Pairs** are pairs of correlated risk factors where the patient targets symmetry between the risk factor during their recovery process. The recovery process aims to develop symmetry in strength and flexibility between the Involved (IL) and Uninvolved limb (UL). As seen in figure 1a, this occurs in a few pairs including *180 deg/s UL peak torque Hams Normalized to Body Weight*



(a) Risk Factor (b) Labeling Function

Fig. 1: Correlation Matrices. IL: Involved Limb, UL: Uninvolved Limb, LSI: Limb Symmetry Index, *: Normalized to Body Weight, **: Normalized to Body Height

and *180 deg/s IL peak torque Hams Normalized to Body Weight*.

While the risk factors themselves may be correlated, the labeling functions tell a different story. For example, the high-risk for *Age at Surgery* is 12-17 years of age, and low-risk is less than 12 and greater than 17 years of age. For, *Age at Return to Sport*, high risk is between 13 and 18 years of age, and low-risk is over 18. It is unknown what risk those under the age of 13 are at. Since these labeling functions provide more information than the data itself, we do not remove any risk factors from our dataset based only on this. We further evaluate correlations based on the labels set by our labeling functions where one is a high risk, zero is unknown, and a negative one is a low risk. Correlations between pairs of datapoints drop drastically using this comparison. The results are shown in Figure 1b.

From these figures, we determine that we should drop a few risk factors. A single timing pair remains highly correlated: *Age at Surgery* and *Age at Return to Sport*. *Age at Surgery* is also encapsulated in *Age & Gender*. Additionally, we see that *Sex* and *Age & Gender* are still highly correlated. To resolve these high correlations, we remove *Age at Surgery* and *Sex* from our risk factors. While no symmetry pairs remain highly correlated, they should not be removed as they are a goal of rehabilitation. Instead, new features should be devised that encompass this symmetry. For example, many of the risk factors already do this with a Limb Symmetry Index.

C. Problem Formulation

Our dataset described above gives us our input space $\mathcal{X}$ of patient information, such that each $x \in \mathcal{X}$ describes a feature related to a patient's demographics, injury information, family history, etc. The features are limited to those encompassed by the list of risk factors obtained from the clinicians. Additionally, we have a label space $\mathcal{Y} = \{-1, 1\}$ that represents whether a patient has gone on to re-tear their ACL after their reconstruction. Thus, the ground truth labeling function can be described as $f_{gt} : \mathcal{X} \mapsto \mathcal{Y}$.

Additionally, we collected a set of labeling functions from clinicians that distinguish high and low risk factors for re-peat ACL tears. These labeling functions use a label space $\bar{\mathcal{Y}} = \{high, low\}$ corresponding to two conclusive and mutually exclusive risk levels of ACL re-tear: high and low. To make the system intuitive, we extend our label space to $\hat{\mathcal{Y}} = \{high, mid, low\}$ by adding a label for medium-level risk. Thus, we have two sets of binary partial label functions: (1) a set H such that $\forall f_h \in H$, $f_h : \mathcal{X} \mapsto \hat{\mathcal{Y}}_h$ where $\hat{\mathcal{Y}}_h = \{high, \neg high\}$ such that $\neg high = \{low, mid\}$ and (2) a set L such that $\forall f_l \in L$, $f_l : \mathcal{X} \mapsto \hat{\mathcal{Y}}_l$ where $\hat{\mathcal{Y}}_l = \{low, \neg low\}$ such that $\neg low = \{mid, high\}$.

The goal of this paper is to evaluate each patients' risk for re-tearing their ACL. By using X, Y, H and L, we seek a labeling function $f^* : \mathcal{X} \mapsto \hat{\mathcal{Y}}$ such that f^* the probability of a re-tear occurring decreases from the high to medium to low-risk category. From the literature, we found that ACL re-tears occur in approximately 30% of patients. In our dataset, we saw a similar occurrence at 32% of patients. We define the rate of occurrence as r. Thus, optimally, we seek to find a labeling function such that

$$\begin{aligned} \Pr[f_{gt}(x) = 1 | f^*(x) = high] &> 2r \\ \Pr[f_{gt}(x) = 1 | f^*(x) = mid] &= r \\ \Pr[f_{gt}(x) = 1 | f^*(x) = low] &< \frac{1}{2}r \end{aligned} \tag{1}$$

In the following section, we introduce RT-ACL as a solution to the problem statement above.

IV. RT-ACL System

The RT-ACL System leverages domain knowledge from clinicians and machine learning techniques to identify a patient's risk level for ACL re-tear. Once the risk level has been determined, it recommends modifications that can be made to lessen this risk level. It is comprised of three components: the RT-ACL model, risk factor identification, and feedback system. The RT-ACL model classifies the risk level of each patient and outputs impact scores for each risk factor. Risk factor identification determines which risk factors most impact a patient's risk level by using the ease and impact of each risk factor combined with a label of modifiable/not modifiable. The feedback system displays the risk level and most significant risk factors back to the clinician for interpretation and conveyance back to the patient. An overview of the RT-ACL system is depicted in Figure 2.

A. RT-ACL Model

The RT-ACL model uses majority vote to classify the risk level of each patient and outputs impact scores for each risk factor as shown in Algorithm 1. It is trained using training data, Z_1, from which weighting is derived. This weighting determines the impact score for each risk factor. Once trained, it takes as input a patient's data and outputs their risk level: high, medium, or low.

First, we split our data $Z = (X, Y)$ into three portions $Z_1 = (X_1, Y_1)$ and $Z_2 = (X_2, Y_2)$, and $Z_3 = (X_3, Y_3)$. Z_1 is the training dataset comprised of 60% of our total data. Z_2 is the testing dataset comprised of 20% of our dataset. Z_3 is comprised of 20% of our dataset and will be used in our Evaluation Section in a hold out set evaluation. Since 32% of our dataset or 114 patients go on to re-tear, we must assure that enough re-tears are in the test set so that can be test high, medium, and low categories. Using 20% of our patients as a test set gives us approximately 20 re-tear patients. By inputting all $x \in X_1$ into the labeling functions in H and L, we can compute a weight corresponding to each function as follows:

Given a labeling function $f_h \in H$, its weight w_{f_h} is

$$w_{f_h} = \frac{\sum_{x \in X_1} \mathbf{1}(f_h(x) = high \wedge f_{gt}(x) = 1)}{\sum_{x \in X_1} \mathbf{1}(f_h(x) = high)} \tag{2}$$

i.e. the percentage of patients that actually go on to re-tear out of the patients labeled as high risk by f_h.

Similarly, given a labeling function $f_l \in L$, its weight w_{f_l} is

$$w_{f_l} = \frac{\sum_{x \in X_1} \mathbf{1}(f_l(x) = low \wedge f_{gt}(x) = 1)}{\sum_{x \in X_1} \mathbf{1}(f_l(x) = low)} \tag{3}$$

i.e. the percentage of patients that go on to re-tear out of the patients labeled as low risk by f_l. We can see that the weights w_{f_h}, $w_{f_l} \in [0, 1]$ as they are percentages. Moreover, the weights measure a labeling functions actual risk using the ground truth f_{gt} in the dataset X_1. For every $f_h \in H$, as w_{f_h} approaches 1, the calculated risk level increases. On the other hand, for every $f_l \in L$, as w_{f_l} approaches 0, the calculated risk level decreases.

Algorithm 1 RT-ACL Model Algorithm

Input: Training Dataset $Z_1 = (X_1, Y_1)$, High-Risk Labeling Functions H, Low-Risk Labeling Functions L, Hyperparameters c_h, c_l, Thresholds t_{lm}, t_{mh}
Output: Risk Estimation Function f^*
Obtain the ground truth labeling function f_{gt} from Z_1
Initialize empty lists w_L, w_H, s
for f_h in H **do**
  $w_{f_h} = \frac{\sum_{x \in X_1} \mathbf{1}(f_h(x) = high \wedge f_{gt}(x) = 1)}{\sum_{x \in X_1} \mathbf{1}(f_h(x) = high)}$
  $w_H.append(w_{fh})$
end for
for f_l in L **do**
  $w_{f_l} = \frac{\sum_{x \in X_1} \mathbf{1}(f_l(x) = low \wedge f_{gt}(x) = 1)}{\sum_{x \in X_1} \mathbf{1}(f_l(x) = low)}$
  $w_L.append(w_{fl})$
end for
for x in X_1 **do**
  $s_x = c_h \sum_{w_{fh} \in w_H} w_{f_h} \mathbf{1}(f_h(x) = high) + c_l \sum_{w_{fl} \in w_L} w_{f_l} \mathbf{1}(f_l(x) = low)$
  s.append(s_x)
end for
for s_x in s **do**
  Define $f^*(x)$ as $\begin{cases} low & s_x < t_{lm} \\ mid & t_{lm} \le s_x < t_{mh} \\ high & s_x \ge t_{mh} \end{cases}$
end for
return f^*

Next, we compute the majority vote by applying the following equation to each $x \in X$:

$$\begin{aligned} s(x) = c_h \sum_{f_h \in H} w_{f_h} \mathbf{1}(f_h(x) = high) \\ + c_l \sum_{f_l \in L} w_{f_l} \mathbf{1}(f_l(x) = low) \end{aligned} \tag{4}$$

where c_h and c_l are hyperparameters chosen as coefficients. The sum $s(x) \in \mathbb{R}$ shows an estimated numerical value of ACL re-tear risk: the larger $s(x)$, the higher risk the patient has. Ideally, $c_h > 0$ and $c_l < 0$ for a system given highly accurate labeling functions.

Finally, we will pick two thresholds t_{lm} and $t_{mh} \in \mathbb{R}$ to split the low, medium and high risk levels based on the $s(x)$ computed. The final ACL re-tear risk estimation function $f^* : \mathcal{X} \mapsto \hat{\mathcal{Y}}$ is

$$f^*(x) = \begin{cases} low & s(x) < t_{lm} \\ mid & t_{lm} \le s(x) < t_{mh} \\ high & s(x) \ge t_{mh} \end{cases} \tag{5}$$

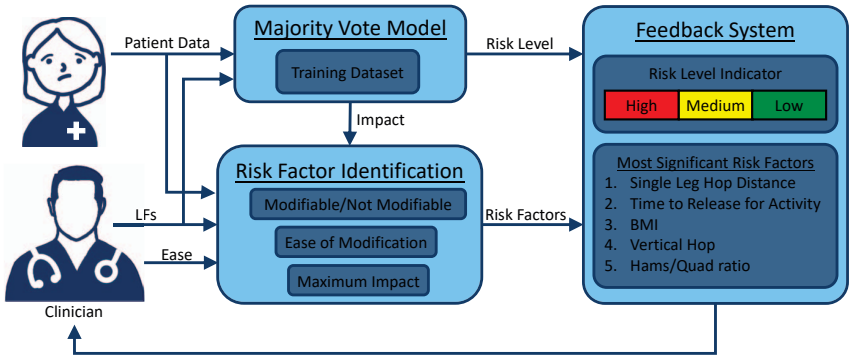


Fig. 2: System Architecture

B. Risk Factor Identification

Once a patient's risk level is determined, our system identifies the most significant risk factors that can be altered during the patient's recovery process to decrease their risk of re-tear. First, we must determine whether a risk factor is modifiable or not. Then, we rank the risk factors by their impact on their risk level and the ease at which they can be modified. The most important modifiable risk factors are chosen based on their impact and ease of modification. These risk factors are then presented to the clinicians for interpretation and conveyance back to the patient.

Most Significant Risk Factors Performance Metrics: For each risk factor, we consider the following performance metrics

- **Modifiable** (m) The clinicians labeled each of the risk factors used in our system as modifiable(1) or not modifiable(0). Thus, $m = \{0, 1\}$ Examples of modifiable risk factors include body mass index(BMI), single-leg hop limb symmetry index, and time to release for activity. Examples of non-modifiable risk factors are first-degree relative with an ACL tear, Meniscus tear, and sex.
- **Impact** (i) is determined by the weight of the risk factor used in the RT-ACL model. Thus $i = w_{f_h}$ or w_{f_l} based on whether a low risk or high risk labeling function was used.
- **Ease of Modification** (e) is set by the clinicians, where risk factors are ranked from one to five, with one being the easiest to modify and five being the most difficult.

Leveraging these metrics, for a given patient, RT-ACLaims to solve the following optimization problem,

$$j^* = \arg \max_{j \in dim(x)} m(x[j]) * (i(x[j]) + e(x[j])) \qquad (6)$$

Where $dim(x)$ is dimensions of x that have been normalized and j^* is the top risk factor identified for our patient. For each patient, we identify the top k risk factors. Once the top risk factor is identified, we remove it from the set of possible risk factors and repeat this process $k-1$ additional times. Once the most significant risk factors have been determined, we display them to the clinician for interpretation and conveyance to the patient.

C. Feedback System

Once the risk level has been determined and the most significant risk factors have been identified, we convey this information to the clinicians. Currently, we display the five most significant risk factors based on their modifiability, impact, and ease of modification. They are ordered from most to least significant. Each significant risk factor is displayed with the labeling functions showing the high and low risk categories along with their absolute risk. They use this information to formulate a rehabilitation program specifically tailored to each individual patient. Patients are then monitored continuously throughout their rehabilitation process. Based on the ongoing data collection, risk levels and risk factors can be reassessed as the patient progresses. As this occurs, rehabilitation programs can be updated to provide an optimal patient outcome.

V. Evaluation

In this section, we evaluate the performance of our RT-ACL model. First, we analyze our labeling functions and their predictive capabilities over the dataset. Then, we consider four weighting methods and two hyperparameter schemes for our RT-ACL model. The weighting methods allow us to assign greater importance to risk factors or labeling functions that are more predictive of re-tear. Hyperparameters allow for exploration of the impact of high vs. low-risk factors. Once weights and hyperparameters are chosen, we perform cross-validation over our dataset obtained from the Children's Hospital of Philadelphia. Finally, we further evaluate the RT-ACL algorithm over a hold-out set.

A. Labeling Function Accuracy

The labeling functions encoded by the clinicians allow us to incorporate professional insight into the RT-ACL model. The purpose of the labeling functions used in this research is to distinguish high and low-risk patients based on individual risk factors. For example, a high-risk labeling function for the feature *Age at Surgery* would be an age between 12 and

TABLE II: Labeling Function Analysis. W: Weight, #: Total Number of Patients in the Risk Level, IL: Involved Limb, UL: Uninvolved Limb, LSI: Limb Symmetry Index, PT: Peak Torque, *: Normalized to Body Weight, **: Normalized to Body Height

Risk Factor	High Risk		Low Risk		Unlabeled
	W	#	W	#	#
Age at Return to Sport	37	188	11	27	139
Delay to Surgery	32	31	14	14	323
Time to Release for Activity	33	227	23	51	76
Time to repeat ACL tear	92	36	69	22	296
Age and Sex	29	193	37	161	0
BMI	32	284	25	8	62
1st Degree Relative ACL Tear	23	40	31	87	227
Any relative ACL tear	26	31	30	96	227
Meniscus tear	28	226	37	71	57
Meniscus resection	23	92	33	205	57
Graft Size	18	11	38	126	57
180 deg/s LSI Quads	31	101	41	130	123
IL Hams/Quads Ratio	32	307	34	44	3
UL Hams/Quads Ratio	32	304	36	47	3
180 deg/s IL PT Quads*	28	110	35	241	3
180 deg/s IL PT Hams*	26	109	36	242	3
180 deg/s Ttl Work LSI Quads	28	139	36	129	86
180 deg/s Ttl Work LSI Hams	34	173	29	80	101
180 deg/s IL Ttl Work Quads	36	110	31	240	4
180 deg/s UL PT Quads*	26	221	43	131	2
180 deg/s UL PT Hams*	33	109	32	242	3
Single Leg Hop LSI	19	32	43	129	193
Triple Hop LSI	31	29	36	140	185
Crossover Hop LSI	54	13	53	69	272
Timed Hop LSI	18	11	53	86	257
Vertical Hop LSI	34	62	40	111	181
IL Triple Hop Distance(cm)**	43	101	39	106	147
IL Vertical Jump(cm)**	44	18	47	82	254
IL Triple & Crossover Hop Dif	52	42	54	82	230
UL Triple & Crossover Hop Dif	54	46	53	78	230
IL Timed Hop	49	102	55	45	207
Sport Played at Injury	37	253	11	34	67

17. Given a data point a, we would encode this as $12 \leq a < 17$. A complete list of the labeling functions used in our analysis is given in the Appendix in Table VI. We calculate the weight of these risk factors, as shown in Equation 2 and Equation 3 which is used in our RT-ACL model. The weights for our labeling functions are shown in Table II. We display the number of patients in each category in the # columns. Further, we present the number of data points that remain unlabeled by our high and low-risk labeling functions in the unlabeled column of this table. These data points can either be missing or do not satisfy the conditions of high or low risk.

We limited our experts to a single opportunity to design our labeling functions. This creates an environment in which the experts are able to impart their domain specific knowledge without biasing it based off of labeling function performance on our specific dataset. This creates labeling functions that should be more easily extrapolated to more generalized datasets. Individually, our labeling functions do not always perform as expected. 20 out of 32 total labeling functions can be considered reversed where the low-risk labeling function has a higher accuracy or predictive capability than the high-risk labeling function. For example, *Age and Sex* has a higher predictive capability for the low-risk labeling function at 37% compared to the high-risk labeling function at 29%. Additionally, some labeling functions cover a low number of data points. Four of the thirty-two total labeling functions do not cover at least two-thirds of our dataset.

B. Weighting Methods

Weighting of risk factors or individual labeling functions allows greater importance to be assigned to those that are more predictive of re-tear. We evaluated four different weighting methods for our RT-ACL model: equal weighting, risk factor weighting, calculated risk weighting, and clinician weighting. As stated in Section IV, our algorithm has two weights, w_{f_h} for the high risk labeling functions and w_{f_l} for the low risk labeling functions. The calculation as well as a discussion of the intuition of each weighting method are as follows:

- **Equal Weighting** is the simplest weighting method. Each labeling function is given an equal vote, i.e. $w_{f_h} = w_{f_l} = 1$ where each labeling function gets one vote. This method puts equal weight on each risk factor and risk category. This indicates that being classified as high risk in BMI is just as important as being classified low risk in graft size or high risk in delay to surgery.
- **Risk Factor Weighting** weights each risk factor individually by calculating the overall risk of those classified in either the high or low risk labeling functions. The weight for each risk factor is calculated by averaging the weights for the high and low risk labeling function given by Equation 2 and Equation 3. In this method, the importance is placed on the risk factor. If a patient moves from a high risk category to a low risk category for a single risk factor, the high risk is assigned the same importance as the low risk.
- **Labeling Function Weighting** weights each labeling function individually. We calculate the weight for the high-risk labeling functions using by Equation 2 and low-risk labeling functions using Equation 3. The values for the weights are shown in the W columns in Table II. In this method, the importance is placed on the risk category. It allows for a distinction between the importance of being in a high or a low risk category in a single risk factor. For example, high risk weight is 37 and low risk weight is 11 for sport played at injury. In the case of this risk factor, it is more important to not be in the high risk category than to be in the low risk category.
- **Clinician Weighting** allows for additional domain knowledge to be incorporated into the algorithm. We collected weights for our labeling functions from two clinicians. The weights ranged from one to five, where five is the most predictive of re-tear, and one is the least. Overall, the clinicians exactly agreed on 27% of the weights with an average difference of 1.2 per risk factor. This shows that even though there wasn't an exact agreement on a high number of risk factors, clinician

TABLE III: Weighting Methods and Hyperparameter Tuning Results

Weighting	Hyperparameter	High Risk #	High Risk %	Mid Risk #	Mid Risk %	Low Risk #	Low Risk %
Equal	$c_h = 1, c_l = -1$	26	11.54	102	32.30	101	37.62
Equal	$c_h = c_l = 1$	103	50.49	230	25.33	21	14.28
Risk Factor	$c_h = 1, c_l = -1$	37	13.89	233	31.76	84	41.67
Risk Factor	$c_h = c_l = 1$	85	56.47	195	28.35	74	12.51
Labeling Function	$c_h = 1, c_l = -1$	17.50	31.22	237	31.22	76	43.42
Labeling Function	$c_h = c_l = 1$	**83**	**59.04**	**184**	**29.51**	**87**	**11.49**
Clinician 1	$c_h = 1, c_l = -1$	52	30.77	252	31.35	49	38.78
Clinician 1	$c_h = c_l = 1$	89	49.44	214	28.50	50	16.00
Clinician 2	$c_h = 1, c_l = -1$	27	25.93	242	31.41	84	36.91
Clinician 2	$c_h = c_l = 1$	115	48.96	219	25.57	19	5.25
Clinician Average	$c_h = 1, c_l = -1$	30	20.00	243	31.82	81	38.27
Clinician Average	$c_h = c_l = 1$	99	50.50	222	27.03	33	12.12
Goal			$>$64		$\approx$32		$<$16

weighting was similar. We evaluated each clinician's weights individually as well as averaged.

C. Hyperparameter Tuning

Hyperparameters allow for exploration of the impact of high vs. low-risk factors. Intuitively, high-risk factors should add to the overall vote while low-risk factors should subtract. However, in practice, this is not always the case. As discussed above, we showed that many of our labeling functions have a reversed predictive capability. To determine how this affects our system, we evaluate an additional hyperparameter scheme. Our RT-ACL algorithm has two hyperparameters, c_h and c_l, as discussed in Section IV. The most intuitive setting for these hyperparameters is $c_h = 1$ and $c_l = -1$, where the high-risk labeling function adds to the total vote, and low-risk labels subtract from the total vote. An additional setting is $c_h = c_l = 1$, where all labels add to the vote, and the amount is solely determined by the weight.

TABLE IV: Demographics of Evaluation Sets

Evaluation	Dataset	% Re-tear	Age Range	Gender(M/F)
CV1	Train	32	8.3-20.7	145/137
CV1	Test	34	8.5-21.0	32/39
CV2	Train	32	8.3-21.0	139/143
CV2	Test	32	9.8-19.7	38/33
CV3	Train	31	8.3-21.0	143/139
CV3	Test	37	8.7-20.7	34/37
CV4	Train	35	8.5-21.0	139/144
CV4	Test	21	8.3-19.8	38/32
CV5	Train	31	8.3-21.0	143/141
CV5	Test	31	10.6-18.5	35/35
Hold Out	Validation	6	9.8-21.5	46/42

D. Performance Analysis

The results from our four weighting methods combined with our two hyperparameters are shown in Table III. Additionally, we show the goal for the different risk level classifications as defined in the Problem Formulation. Overall, the $c_h = c_l = 1$ hyperparameter outperformed the $c_h = 1, c_l = -1$ hyperparameter, further exemplifying the high number of reversed labeling functions. The best overall performance came from the labeling function weighting with $c_h = c_l = 1$ for hyperparameters. It has the highest risk prediction of all of the high-risk categorizations at 59.04%. While this does not achieve the goal of twice the dataset re-tear rate of 64%, it comes the closest of all the high-risk categories. Its medium risk and low-risk categorizations meet the goals stated in the problem formulation.

E. Cross Validation

We perform further evaluation on the RT-ACL model that leverages labeling function weighting and $c_h = c_l = 1$ for hyperparameters. In order to evaluate the performance of our weighted RT-ACL model, we perform a 5-fold patient cross-validation on the dataset obtained from the Children's Hospital of Philadelphia. We use datasets Z_1 and Z_2 in our cross validation. We combine the datasets and randomly partitioned into five equal-sized partitions of patients. In each fold; four partitions are used for training and one for testing. The process is repeated five times such that each subgroup is used for validation exactly once.

TABLE V: Cross Validation Results

Fold	HR #	HR %	MR #	MR %	LR #	LR %
1	20	55.00	28	18.18	22	18.18
2	22	54.55	31	45.16	18	5.56
3	27	48.14	30	10.00	13	7.69
4	10	80.0	30	33.33	30	10.00
5	15	46.67	26	44.00	29	20.68

The results of the cross validation are shown in Table V. In each fold, we see that the high-risk category has the highest percentage of re-tears, followed by mid-risk and low risk. On average, the high-risk category has a 53% chance of re-tear, the medium risk category has a 30% chance of re-tear, and the low-risk category has a 16% chance of re-tear.

Additionally, we present the demographics for each of our training and test sets for the cross validation in Table IV. From this we see that the fourth partition's test set has a much lower re-tear percentage, 21% compare to an average of 31.6%. When we look to our results in Table V, we see that our RT-ACLmodel still performs as expected. Fold two and five had a high percentage of re-tears in their medium risk bins. Optimally, we expect to be closer to 32%. When looking into the demographics, these folds had a smaller age range within the test set. This should be explored in future work to determine if age should play a greater factor in our RT-ACLmodel.

F. Holdout Evaluation

Additionally, we tested our model using the holdout dataset, Z_3. This dataset was not used in any of the previously described analysis or model development. Overall, our holdout dataset contained 88 patients of which only five went on to re-tear. Of these five patients two were high risk, two were medium risk, and one was low risk. This is a very low number of re-tears but the re-tear percentages for each of these categories were 8.0%, 5.4%, and 4% respectively. The results are promising as consistent with our goals for the system we seek to accomplish twice the re-tear rate(12%) of the entire dataset in the high risk category, a similar re-tear rate(6%) in the medium risk category, and half the re-tear rate(3%) in the low risk category. While our system did not perfectly meet these goals, it does show the stratification between the categories and with more data, can be evaluated further in the future.

VI. Discussion and Future Work

Future work in this research area should encompass the following: First, additional risk factors can be identified with targeted studies and literature reviews. Second, new algorithms and updates to the existing algorithms can be examined to increase the performance of the model. Finally, confidence metrics can be added to our risk prediction to make our predictions more robust.

A. Risk Factors

In this work, we focused on risk factors provided by our clinicians. Since our clinicians collected our dataset and developed our labeling functions, it creates a preference for our RT-ACL model towards our data. While this is expected, to extend the model to handle other datasets, we can add additional risk factors and re-weight our algorithm. Additionally, a literature review can be done to collect already known risk factors that have been evaluated by others. If these risk factors are included in our dataset, we can evaluate the new risk factors on our dataset to set the weights for our RT-ACL model.

As features are added, it is possible that correlations may develop. While we do not seek to overweight the importance of correlated features, we can use them interchangeably, allowing for a more robust model. This would allow our model to be used more effectively on new datasets with a different feature space. For example, *age and sex* could be exchanged for *age* and/or *sex*. Further, this would allow our model to be tailored to specific surgeons, clinics, and rehab facilities.

B. Model Updates

In this work, we saw that many labeling functions did not perform as expected on our dataset, i.e., the low-risk category had a higher re-tear rate than the high-risk category. Regardless of this, we were able to create an accurate model that could predict a patient's risk of re-tear. To further improve our model, an investigation into these labeling functions should be done including a literature and dataset review. In future work, more intelligent labeling functions can be designed to promote the use of new models and hyperparameters. Further, more sophisticated machine learning models can be investigated to potentially increase the performance of RT-ACL.

C. Confidence

Currently, our majority vote model handles missing data by considering the risk as unlabeled; thus, a high or low-risk category is not assigned. To take the missing data into consideration, a confidence score could be developed. In its simplest form, it could be based on the number of risk factors for which a patient has data recorded. For example, if a patient has 80% of the risk factors, we could give a confidence score of 80% for their risk level. Since 20% of the data was missing, we cannot evaluate their risk for these risk factors. This could also be adapted to express the percentage of the total weight of the risk factors that could be evaluated.

If a patient is missing information, the system could recommend that this information be gathered. As discussed above, similar risk factors that are highly correlated could be identified as replacements for those targeted in our system. With this in mind, the confidence score can be updated based on how similar these metrics are. Additionally, a threshold should be established for the minimum number of risk factors that can lead to an accurate risk level prediction. If a patient does not reach that threshold, then no risk level prediction should be given, just the risk factor data that should be collected.

VII. Conclusion

In this paper, we presented RT-ACL, a system that identifies high-risk patients and determines their most significant risk factors to reduce ACL reinjury risk. RT-ACLuses a RT-ACL model that leverages labeling functions designed by clinicians to classify the risk level of an ACL re-tear. Once a risk level has been determined, it identifies modifiable risk factors, and suggests the most significant risk factors to clinicians to develop intervention methods from to minimize adverse outcomes and complications. We evaluated our system on a dataset of 441 youth patients, 8-21 years of age that underwent an ACL reconstruction at the Children's Hospital of Philadelphia. The results indicate patients classified as high risk re-tear at a rate of 59%, medium risk at a rate of 30%, and low risk at a rate of 12%. This demonstrates that those classified by our system as high risk are 4.6 times as likely to re-tear their ACL than those classified as low risk.

ACKNOWLEDGMENT

This research was supported in part by the U.S. National Science Foundation under grant 1915398 and the National Institute of Health under grant R01EB029363.

REFERENCES

[1] F. R. Noyes, "The function of the human anterior cruciate ligament and analysis of single-and double-bundle graft reconstructions," *Sports Health*, vol. 1, no. 1, pp. 66–75, 2009.

[2] N. A. Friel and C. R. Chu, "The role of acl injury in the development of posttraumatic knee osteoarthritis," *Clinics in sports medicine*, vol. 32, no. 1, pp. 1–12, 2013.

[3] N. A. Bates, A. L. McPherson, M. B. Rao, G. D. Myer, and T. E. Hewett, "Characteristics of inpatient anterior cruciate ligament reconstructions and concomitant injuries," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 24, no. 9, pp. 2778–2786, 2016.

[4] R. M. Queen, "Infographic: Acl injury reconstruction and recovery," *Bone & joint research*, vol. 6, no. 11, pp. 621–622, 2017.

[5] N. D. Schilaty, C. Nagelli, N. A. Bates, T. L. Sanders, A. J. Krych, M. J. Stuart, and T. E. Hewett, "Incidence of second anterior cruciate ligament tears and identification of associated risk factors from 2001 to 2010 using a geographic database," *Orthopaedic journal of sports medicine*, vol. 5, no. 8, p. 2325967117724196, 2017.

[6] J. Wegrzyn, J. Chouteau, R. Philippot, M.-H. Fessy, and B. Moyen, "Repeat revision of anterior cruciate ligament reconstruction: a retrospective review of management and outcome of 10 patients with an average 3-year follow-up," *The American journal of sports medicine*, vol. 37, no. 4, pp. 776–785, 2009.

[7] N. A. Beck, J. T. R. Lawrence, J. D. Nordin, T. A. DeFor, and M. Tompkins, "Acl tears in school-aged children and adolescents over 20 years," *Pediatrics*, vol. 139, no. 3, 2017.

[8] J. S. Frank and P. L. Gambacorta, "Anterior cruciate ligament injuries in the skeletally immature athlete: diagnosis and management," *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*, vol. 21, no. 2, pp. 78–87, 2013.

[9] C. R. LaBella, W. Hennrikus, T. E. Hewett *et al.*, "Anterior cruciate ligament injuries: diagnosis, treatment, and prevention," *Pediatrics*, vol. 133, no. 5, pp. e1437–e1450, 2014.

[10] N. A. Beck, N. M. Patel, and T. J. Ganley, "The pediatric knee: current concepts in sports medicine," *Journal of Pediatric Orthopaedics B*, vol. 23, no. 1, pp. 59–66, 2014.

[11] R. Crawford, G. Walley, S. Bridgman, and N. Maffulli, "Magnetic resonance imaging versus arthroscopy in the diagnosis of knee pathology, concentrating on meniscal lesions and acl tears: a systematic review," *British medical bulletin*, vol. 84, no. 1, pp. 5–23, 2007.

[12] K. E. Webster, J. A. Feller, W. B. Leigh, and A. K. Richmond, "Younger patients are at increased risk for graft rupture and contralateral injury after anterior cruciate ligament reconstruction," *The American journal of sports medicine*, vol. 42, no. 3, pp. 641–647, 2014.

[13] M. V. Paterno, M. J. Rauh, L. C. Schmitt, K. R. Ford, and T. E. Hewett, "Incidence of second acl injuries 2 years after primary acl reconstruction and return to sport," *The American journal of sports medicine*, vol. 42, no. 7, pp. 1567–1573, 2014.

[14] T. L. Sanders, A. Pareek, T. E. Hewett, B. A. Levy, D. L. Dahm, M. J. Stuart, and A. J. Krych, "Long-term rate of graft failure after acl reconstruction: a geographic population cohort analysis," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 25, no. 1, pp. 222–228, 2017.

[15] R. W. Wright, R. A. Magnussen, W. R. Dunn, and K. P. Spindler, "Ipsilateral graft and contralateral acl rupture at five years or more following acl reconstruction: a systematic review," *The Journal of Bone and Joint Surgery. American volume.*, vol. 93, no. 12, p. 1159, 2011.

[16] G. V. Kamath, T. Murphy, R. A. Creighton, N. Viradia, T. N. Taft, and J. T. Spang, "Anterior cruciate ligament injury, return to play, and reinjury in the elite collegiate athlete: analysis of an ncaa division i cohort," *The American journal of sports medicine*, vol. 42, no. 7, pp. 1638–1643, 2014.

[17] K. E. Webster and J. A. Feller, "Exploring the high reinjury rate in younger patients undergoing anterior cruciate ligament reconstruction," *The American journal of sports medicine*, vol. 44, no. 11, pp. 2827–2832, 2016.

[18] M. Buckthorpe, "Recommendations for movement re-training after acl reconstruction," *Sports Medicine*, pp. 1–18, 2021.

[19] K. E. Webster, C. V. Nagelli, T. E. Hewett, and J. A. Feller, "Factors associated with psychological readiness to return to sport after anterior cruciate ligament reconstruction surgery," *The American journal of sports medicine*, vol. 46, no. 7, pp. 1545–1550, 2018.

[20] A. Kotsifaki, R. Whiteley, S. Van Rossom, V. Korakakis, R. Bahr, V. Sideris, P. Graham-Smith, and I. Jonkers, "Single leg hop for distance symmetry masks lower limb biomechanics: time to discuss hop distance as decision criterion for return to sport after acl reconstruction?" *British Journal of Sports Medicine*, 2021.

[21] A. Irfan, S. Kerr, G. Hopper, W. Wilson, L. Wilson, and G. Mackay, "A criterion based rehabilitation protocol for acl repair with internal brace augmentation," *International Journal of Sports Physical Therapy*, vol. 16, no. 3, pp. 870–878, 2021.

[22] P. Kyritsis, R. Bahr, P. Landreau, R. Miladi, and E. Witvrouw, "Likelihood of acl graft rupture: not meeting six clinical discharge criteria before return to sport is associated with a four times greater risk of rupture," *British journal of sports medicine*, vol. 50, no. 15, pp. 946–951, 2016.

[23] M. V. Paterno, B. Huang, S. Thomas, T. E. Hewett, and L. C. Schmitt, "Clinical factors that predict a second acl injury after acl reconstruction and return to sport: preliminary development of a clinical decision algorithm," *Orthopaedic journal of sports medicine*, vol. 5, no. 12, p. 2325967117745279, 2017.

[24] H. Grindem, L. Snyder-Mackler, H. Moksnes, L. Engebretsen, and M. A. Risberg, "Simple decision rules can reduce reinjury risk by 84% after acl reconstruction: the delaware-oslo acl cohort study," *British journal of sports medicine*, vol. 50, no. 13, pp. 804–808, 2016.

[25] X. Shen, Y. Qin, J. Zuo, T. Liu, and J. Xiao, "A systematic review of risk factors for anterior cruciate ligament reconstruction failure," *International Journal of Sports Medicine*, 2021.

[26] P. L. Sousa, A. J. Krych, R. A. Cates, B. A. Levy, M. J. Stuart, and D. L. Dahm, "Return to sport: does excellent 6-month strength and function following acl reconstruction predict midterm outcomes?" *Knee surgery, sports traumatology, arthroscopy*, vol. 25, no. 5, pp. 1356–1363, 2017.

[27] J. R. Ebert, P. Edwards, L. Yi, B. Joss, T. Ackland, R. Carey-Smith, J.-U. Buelow, and B. Hewitt, "Strength and functional symmetry is associated with post-operative rehabilitation in patients following anterior cruciate ligament reconstruction," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 26, no. 8, pp. 2353–2361, 2018.

[28] Y. Tanaka, Y. Yonetani, Y. Shiozaki, T. Kitaguchi, N. Sato, S. Takeshita, and S. Horibe, "Retear of anterior cruciate ligament grafts in female basketball players: a case series," *BMC Sports Science, Medicine and Rehabilitation*, vol. 2, no. 1, pp. 1–6, 2010.

[29] M. V. Paterno, L. C. Schmitt, K. R. Ford, M. J. Rauh, G. D. Myer, B. Huang, and T. E. Hewett, "Biomechanical measures during landing and postural stability predict second anterior cruciate ligament injury after anterior cruciate ligament reconstruction and return to sport," *The American journal of sports medicine*, vol. 38, no. 10, pp. 1968–1978, 2010.

[30] R. Papalia, F. Franceschi, A. Tecame, S. D'Adamio, N. Maffulli, and V. Denaro, "Anterior cruciate ligament reconstruction and return to sport activity: postural control as the key to success," *International orthopaedics*, vol. 39, no. 3, pp. 527–534, 2015.

[31] S. D. Barber-Westin and F. R. Noyes, "Factors used to determine return to unrestricted sports activities after anterior cruciate ligament reconstruction," *Arthroscopy: the journal of arthroscopic & related surgery*, vol. 27, no. 12, pp. 1697–1705, 2011.

[32] S. D. BarberWestin and F. R. Noyes, "Objective criteria for return to athletics after anterior cruciate ligament reconstruction and subsequent reinjury rates: a systematic review," *The Physician and sportsmedicine*, vol. 39, no. 3, pp. 100–110, 2011.

[33] J. Kvist, "Rehabilitation following anterior cruciate ligament injury," *Sports Medicine*, vol. 34, no. 4, pp. 269–280, 2004.

[34] D. Nguyen, "Sex, age, and graft size as predictors of acl re-tear: a multivariate logistic regression of a cohort of 503 athletes," *Orthopaedic Journal of Sports Medicine*, vol. 4, no. 7_suppl4, p. 2325967116S00164, 2016.

[35] M. K. Group, K. P. Spindler, L. J. Huston, A. Zajichek, E. K. Reinke, A. Amendola, J. T. Andrish, R. H. Brophy, W. R. Dunn, D. C. Flanigan *et al.*, "Anterior cruciate ligament reconstruction in high school and college-aged athletes: Does autograft choice influence anterior cruciate ligament revision rates?" *The American journal of sports medicine*, vol. 48, no. 2, pp. 298–309, 2020.

APPENDIX A
EXTENDED LABELING FUNCTIONS

In the following table we present the high and low risk labeling functions as provided by the clinicians. If the high and low risk categories did not encompass all the patients, they were put in the Unknown category. This includes null values.

TABLE VI: **normalized to height, * normalized to weight, std: standard deviation

Feature	High Risk	Low Risk	Unknown
Age at Surgery	$12 \leq x < 17$	$(x < 12) \vert (x \geq 17)$	
Age at Return to Sport	$13 \leq x < 17$	$x \geq 18$	$(17 \geq x > 18) \vert (x < 13)$
Delay to Surgery	$x \leq 21$		$x > 21$
Time to Release for Activity	$x \leq 10.5$	$x > 10.5$	
Time to repeat ACL tear	$1 \leq x \leq 3$	$x > 3$	$x < 365$
Sex	$x == F$	$x == M$	
Age(a) & Sex(s)	$(a \geq 12) \& (s == F)$ $(a < 12) \& (s == M)$	$(a < 12) \& (s == F)$ $(a \geq 12) \& (s == M)$	
BMI	$x < 25$	$x \geq 35$	$(x \geq 25)$ & $(x < 35)$
1st degree relative with ACL tear	$x == Yes$	$x == No$	
Any relative with ACL tear	$x == Yes$	$x == No$	
Meniscus tear	$x == Yes$	$x == No$	
Meniscus resection	$x == Yes$	$x == No$	
Graft Size	$x \leq 7.5$	$x \geq 8$	$(x > 7.5) \& (x < 8)$
180 deg/s LSI QUADS	$x \leq 90$	$x > 90$	
Involvled Limb Hams/Quad ratio	$(x < 60) \vert (x > 65$	$(x >= 60) \& (x <= 65)$	
Uninvolvled Limb Hams/Quads Ratio	$(x < 60) \vert (x > 65)$	$(x >= 60) \& (x <= 65)$	
180 deg/s involved limb peak torque quads*	$(x > std) \vert (x < std)$	$(x <= SH) \& (x >= SL)$	
180 deg/s involved limb peak torque hams*	$(x > std) \vert (x < std)$	$(x <= SH) \& (x >= SL)$	
180 deg/s Total Work LSI Quads	$(x < 90)$	$(x >= 90) \& (x <= 100)$	$x > 100$
180 deg/s Total Work LSI Hams	$(x < 90)$	$(x >= 90) \& (x <= 100)$	$x > 100$
180 deg/s Involved limb total work Quads	$(x > std) \vert (x < std)$	$(x <= SH) \& (x >= SL)$	
180 deg/s Involved limb total work Hams	$(x > std) \vert (x < std)$	$(x <= SH) \& (x >= SL)$	
Single Leg Hop LSI	$x < 90$	$x \geq 90$ & $x \leq 100$	$x > 100$
Triple Hop LSI	$x < 90$	$(x >= 90) \& (x <= 100)$	$x > 100$
Crossover Hop LSI	$x < 90$	$(x >= 90) \& (x <= 100)$	$x > 100$
Timed Hop LSI	$x < 90$	$(x >= 90) \& (x <= 100)$	$x > 100$
Vertical Hop LSI	$x < 90$	$(x >= 90) \& (x <= 100)$	$x > 100$
Involved Limb Triple Hop Distance(cm)**	$((2.5 < x \leq 3) \vert (x < 1.9)$	$(x \geq 1.9) \& (x \leq 2.5)$	$x > 3$
Involved Limb Vertical Jump*	$x < .1$	$x > .15$	
Triple and Crossover Hop Involved Difference	$(x < 0) \vert (x > 24)$	$x \geq 0$	$0 < x \leq 24$
Triple and Crossover Hop Uninvolved Difference	$(x < 0) \vert (x > 24)$	$x \geq 0$	$x \leq 24$
Involved Limb Timed Hop	$x \geq 2$	$x < 2$	
Sport Played at Injury	Basketball Field Hockey Football Lacrosse Rugby Skiing Soccer Volleyball	Baseball Softball Running Swimming Non-Sport Injuries	Cheerleading Gymnastics Ice Hockey Tennis Squash Snowboarding Wrestling