

Longitudinal Analysis of ACL Rehabilitation: Analyzing Adjusted Torque to Identify Early-Stage Recovery Plateaus

Terah Skillings
Marketing and Data Science
Master's of Data Science
Regis University, Denver, CO, USA
tskillings@regis.edu

Abstract

Even with large amounts of longitudinal data from ACL rehab, physical therapists follow return to play protocols using time-based guidelines and average recovery benchmarks. This does not account for how each individual athlete is recovering (Compagnin et al., 2025). Return to play decisions are currently affected by the “Limb Symmetry Index (LSI) Trap” where relative symmetry measures mask strength deficits, leading to unrecognized recovery plateaus and a secondary injury rate of nearly 23% in young athletes (Wiggins et al., 2016). This project addressed the traditional methods used by physical therapists by applying the full data science lifecycle to mixed messy clinical datasets.

The analysis uses real-world data from a physical therapist. Historical clinical data varied between each athlete, resolving major preprocessing hurdles, including normalizing column structure, creating a master file instead of individual athlete workbooks, and following HIPAA hiding athlete’s identities by creating athlete IDs. The features were also cleaned, including filtering (≥ 14 days post-operation) to eliminate acute surgical noise (Alqahtani et al., 2025). These engineered features were looked at from a longitudinal lens, resulting in an interpretable multivariate logistic regression model that was created to prioritize clinical interpretability (Zhu, Kelly, et al., 2025).

The model achieved a ROC-AUC of 0.71, demonstrated moderate discriminative performance and an improved predictive signal compared to symmetry-only metrics (AUC: 0.59) (Pietrosimone et al., 2016). A clinical threshold of 6.79 Nm/kg was observed as a potential red flag for early plateau detection. This data-driven framework offers PTs a scalable method to move beyond manual visual follow up. This approach helps PTs focus on metrics that can be more informative. Doing so can improve safety during return-to-play by better identifying recovery patterns that are not typical.

I. INTRODUCTION/BACKGROUND

ACL tears are not only a debilitating injury, but have a reinjury rate of nearly 23% in young, active populations (Wiggins et al., 2016). Even with demanding physical therapy (PT), it is normal for athletes to return to play with more underlying biomechanical issues. Although physical therapists collect large amounts of testing data, 92.9% rate their current testing as less than “excellent” due to a lack of objective, standardized frameworks (Compagnin et al., 2025). Recent surveys highlight a global shift toward criteria based progression. There remains a significant gap in the use of longitudinal data to predict long-term outcomes (Compagnin et al., 2025). This project applies the data science lifecycle to fix this gap by focusing on optimizing recovery trajectories and avoiding losing strength in the first 60 days after an operation (Alqahtani et al., 2025).

II. PROBLEM STATEMENT

Current clinical practices use the Limb Symmetric Index (LSI) as a key way to track the recovery of an athlete. LSI can create a “false positive” for readiness. If the uninjured limb loses strength from lack of training, an athlete may achieve 90% symmetry while remaining functionally weak (Sheenam et al., 2025).

There is currently no standard scalable framework for interpreting how multiple recovery metrics change over time and identify these “symmetrically weak” athletes in a clinical setting. This project addresses the “LSI Trap” by modeling adjusted torque metrics with symmetry to provide clearer decision support for return-to-play readiness.

III. RELATED WORK

New research suggests that quadricep thickness can decrease by 20% within the first 30 days after surgery (Sheenam et al., 2025). Although subjective pain scores often improve rapidly, surface EMG data indicate that neural firing capacity remains depressed for months (Sheenam et al., 2025). Along with this, adjusted torque Nm/Kg, defined as torque normalized to body mass, has been established as a key feature of function (AUC: 0.76) compared to LSI (AUC: 0.62) (Pietrosimone et al., 2016). Clinical benchmarks suggest a target of 3.10 Nm/kg for high-level function, although early-stage thresholds for plateau detection remain under-researched (Pietrosimone et al., 2016). Although Pietrosimone et al. (2016) established an example of the importance of adjusted torque over symmetry, their work focused primarily on late-stage outcomes using controlled laboratory data. This project builds on their work by applying similar metrics early stage detection (0–60 days post-op) where there is more variability and data sparsity making this clinical dataset more challenging. This study creates a data-based way to detect when progress could plateau by solving data engineering issues and focusing on the early signs of these recovery patterns.

IV. METHODOLOGY/APPROACH

The study followed the data science life cycle, beginning with the consolidation of individual athlete workbooks into a single de-identified master dataset. This standardized dataset was then analyzed in Python using the Pandas, NumPy, and Scikit-Learn libraries. Significant cleaning was required to resolve temporal data issues where initial deltas resulted in nanosecond-scale values (10^{15} units); these were mathematically converted to integer days to establish a valid longitudinal timeline. To ensure data quality and mitigate acute noise—such as post-operative swelling—records were filtered to include only entries ≥ 14 days after the operation (Alqahtani et al., 2025).

For the analytical approach, a Multivariate Logistic Regression was selected to prioritize clinical interpretability. Due to the limited features (two predictors) and the emphasis on interpretability, standard logistic regression was used over penalized alternatives to preserve the model transparency clinical interpretation. This allowed for a direct comparison between relative symmetry and adjusted torque metrics (Zhu, Kelly, et al., 2025). The model was trained on engineered features, specifically the mean values of “Quad LSI” and “Adj Quad Torque” within the 0–60 day window.

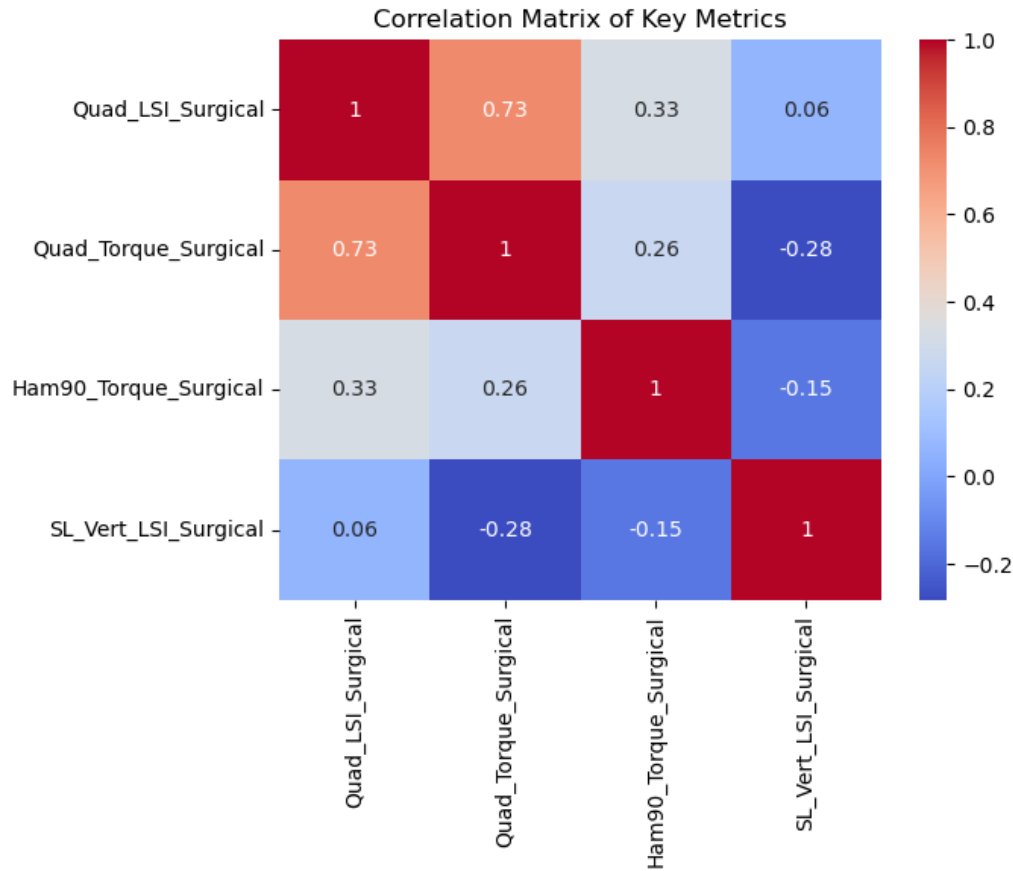


Fig. 1. Correlation Matrix of clinical features. The relationship between Quad LSI and Adjusted Torque ($r = 0.73$) justifies the use of a multivariate approach to isolate the predictive power of adjusted torque strength.

The target variable, a “Plateau,” was defined as an athlete who failed to reach a 90% peak LSI or showed $< 5\%$ improvement during their final three visits. Given the small sample size ($n = 18$), ROC-AUC was utilized as the primary evaluation metric to measure the effect size, as statistical theory suggests that a sample size of $n \geq 16$ is sufficient to detect robust clinical signals when the effect size is substantial (Vabalas et al., 2019). The Leave-One-Out (LOO) method was originally used to create a predictive model; however, the performance metrics varied substantially across folds, matching the high deviation expected in the small sample modeling. The addition or removal of a single subject greatly affected the model parameters, leading to the exploratory nature of this analysis.

V. DATA DESCRIPTION & ETHICS

The dataset consists of de-identified clinical performance records obtained from multiple athlete workbooks. Data governance and privacy was maintained by centralizing storage in a Master CSV file and utilizing version control for all Python transformation scripts. Ethics and privacy were prioritized by removing all personally identifiable information (PII) and assigning unique Athlete IDs to maintain HIPAA compliance.

Data quality assurance involved manual verification of column structures and a meticulous sparsity analysis to identify missingness throughout the longitudinal feature set. As shown in Figure 2, this visualization allowed the identification of testing gaps. It also informed the strategy for handling missing values during consolidation.

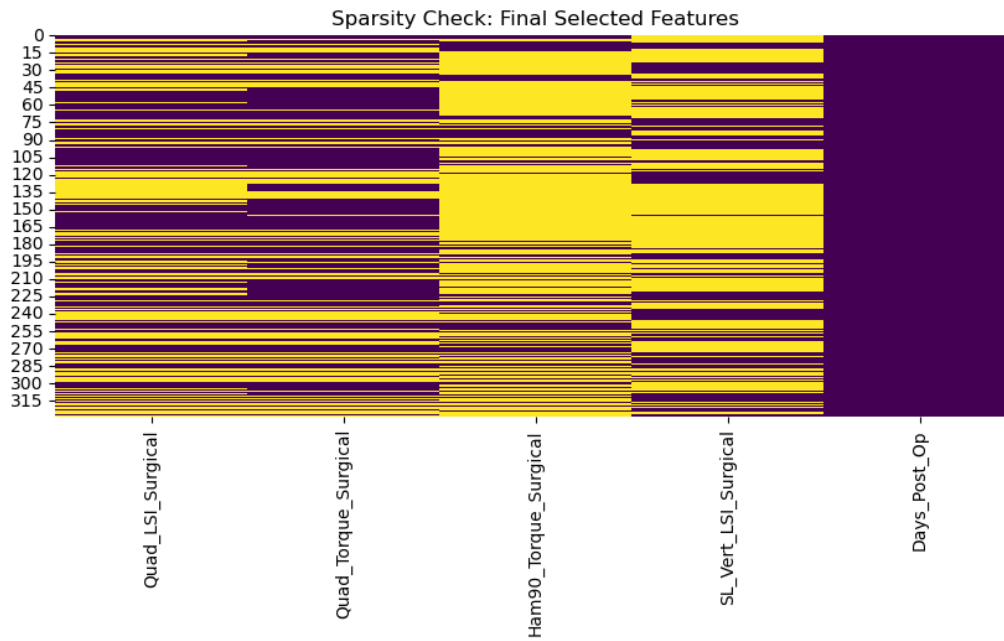


Fig. 2. Sparsity Matrix of clinical records. This visualization was used during the data cleaning phase to identify missing values and ensure the integrity of the longitudinal dataset prior to modeling. These are key features within all athletes testing

Potential biases, such as the small sample size ($n = 18$) and being only from one clinical setting, were acknowledged as limitations. The results of this project are intended to be interpreted as decision-support insights rather than diagnostic medical conclusions, following responsible research standards in healthcare data science.

VI. RESULTS AND DISCUSSION

Exploratory Data Analysis (EDA) revealed that although LSI scores were distributed throughout all recovery phases, adjusted torque showed a distinct stagnation in the “Plateau” group. As seen in Figure 3, the relationship between relative symmetry and adjusted torque highlights the “LSI Trap.” Many athletes who achieve the traditional 90% LSI benchmark still present with significantly low adjusted torque, linking them as functionally at risk despite meeting symmetry goals.

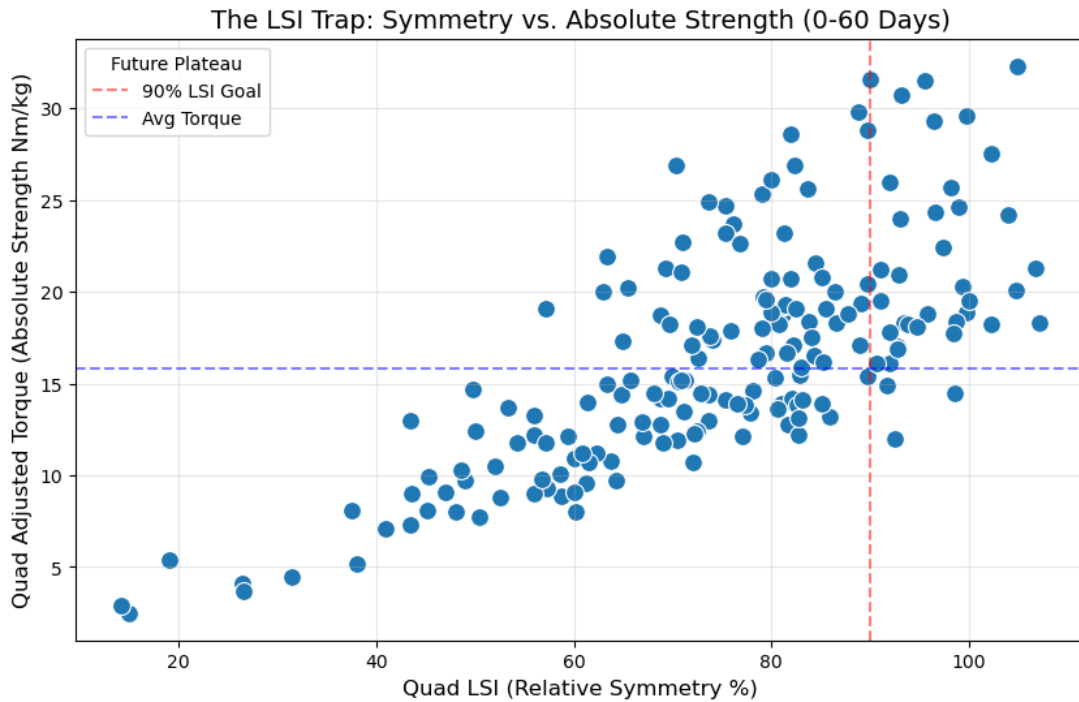


Fig. 3. The LSI Trap: Scatter plot of Quad LSI vs. Adjusted Torque. The bottom-right quadrant highlights athletes who meet symmetry benchmarks (> 90%) but fail to meet adjusted strength floors, hiding potential recovery plateaus.

Independent t-tests were performed. Neither variable reached statistical significance individually (LSI $p = 0.60$; Torque $p = 0.47$), consistent with the limited power in this group. The final multivariate model demonstrated stronger predictive power: the LSI Baseline achieved an ROC-AUC of 0.5952, while the Multivariate (LSI + Torque) model reached an ROC-AUC of 0.7143. Preformance remains exploratory because of the sample size. The improvement in AUC should be noted and suggest meaningful stronger signals when adjusted torque is incorporated.

Individual longitudinal trajectories (Figure 4) further demonstrate this variability. Some athletes have a stable recovery. Others show a clear stagnation. A threshold of 6.79 Nm/kg was observed for these athletes; those who did not reach this on Day 60 were more likely to be classified in the plateau group.

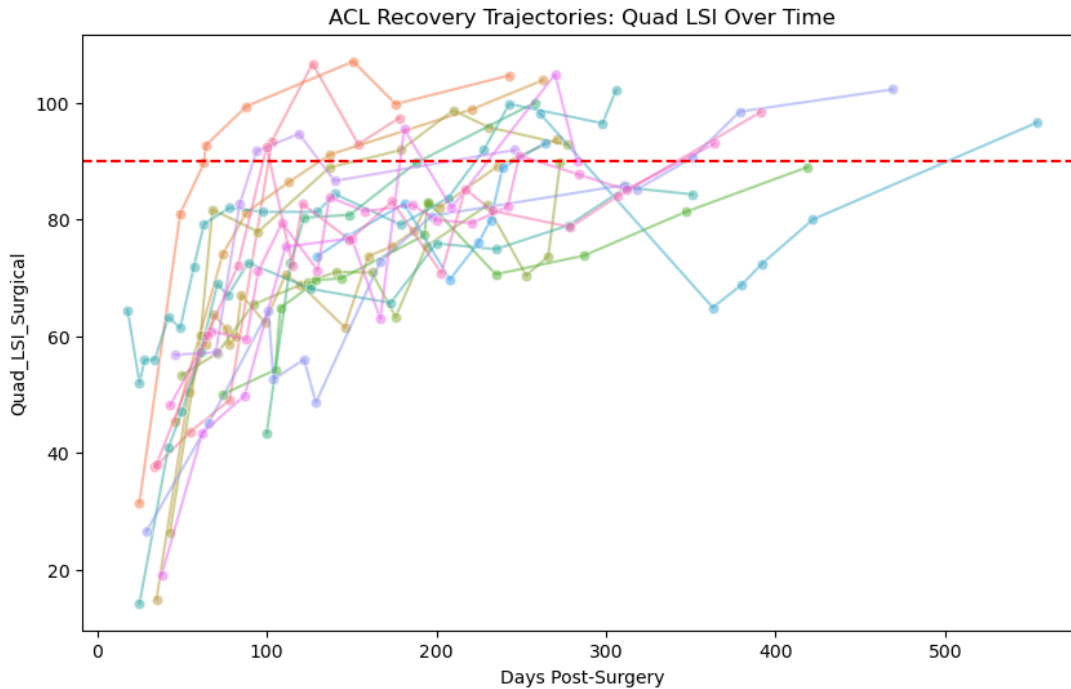


Fig. 4. Longitudinal Recovery Trajectories. The high variability in individual Quad LSI progression necessitates the use of adjusted torque thresholds (6.79 Nm/kg) for early-stage plateau detection.

The model coefficients revealed that the torque (-0.212) has a strong negative correlation with the risk of plateau. LSI (0.018) had a small impact when adjusted torque was taken into account. This confirms that the adjusted torque is the primary feature of recovery, while symmetry is merely a relative indicator that can be influenced by detraining of the opposing limb.

VII. TIMELINE

The project was executed over an eight-week period, moving from a collaborative project practicality phase to a technical data engineering and modeling cycle. The following table summarizes the milestones and the iterative process of working with the physical therapist who shared their data.

Phase	Week	Key Activities & Milestones
Feasibility	Weeks 1-2	Initial brainstorming and multiple meetings with the Physical Therapist (PT) to determine if a data-driven analysis of rehabilitation workbooks was possible.
Architecture	Week 3	Began creating the master file; reorganized varying column structures and researched the clinical testing protocols to define feature significance.
Ingestion	Week 4	Finalized the Google Sheets master template and consolidated the first major batch of athlete workbooks; normalized tabs to minimize Python cleaning.
Cleaning	Week 5	Handled inconsistent formatting from new batches; established phase based groupings (Early, Mid, Late Rehab) and noted structural missingness patterns.
EDA & Pivot	Week 6	Integrated final athlete workbooks ($n = 18$). Visualized recovery trajectories and refined the scope from broad prediction to targeted plateau analysis.
Modeling	Week 7	Resolved minor code issues. Developed the Multivariate Logistic Regression and identified the 6.79 Nm/kg threshold.
Final Delivery	Week 8	Validation of results and completion of the final report/presentation.

TABLE I

PRACTICUM PROJECT TIMELINE: FROM FEASIBILITY TO PREDICTIVE ANALYSIS.

VIII. CONCLUSION

This study successfully followed the data science lifecycle to analyze the rehabilitation trajectories of ACL. It identified key clinical signs that differentiate between successful recovery and plateauing. By applying a Multivariate Logistic Regression model, the research defined a critical functional threshold of 6.79 Nm/kg for the adjusted quadriceps torque. In this exploratory group, models that incorporate adjusted torque demonstrated improved discrimination (ROC-AUC: 0.71) compared to symmetry-only approaches (ROC-AUC: 0.59). This result shows a clear quantitative signal that outperforms traditional and often misleading LSI-only benchmarks. This project demonstrates the value of applying data science methods to real-world clinical datasets, successfully extracting actionable insights from complex, non-linear recovery patterns to support better rehabilitation outcomes.

Future work should involve expanding the dataset to move from this exploratory phase to a validated predictive model. Incorporating a larger athlete cohort would allow for more stable classification and the use of Leave-One-Out (LOO) validation to ensure stability. These findings lay the groundwork for real-time clinician tools that can identify at-risk athletes early in the 0–60 day intervention zone.

REFERENCES

- Alqahtani, T. A., Kardm, S. M., Alnakhli, H. H. et al. (2025). Impact of early vs. delayed physical therapy on functional recovery, proprioception, and return to sport after anterior cruciate ligament (ACL) reconstruction: a cross-sectional study. *Journal of Orthopaedic Surgery and Research*, 20(644).
- Avery, S. (2024, October 9). *The hidden science of ACL injuries*. Duke University School of Medicine. <https://medicine.duke.edu/news/hidden-science-acl-injuries>
- Bruce, A. S., Thompson, X. D., Queen, R. M., et al. (2025). Analysis of limb loading and lower extremity strength recovery across time after anterior cruciate ligament reconstruction. *Sports Health*, 17(5), 1063–1071.
- Compagnin, A., Della Villa, F., La Rosa, G. et al. (2025). Force and power testing during anterior cruciate ligament reconstruction rehabilitation: A world-wide survey of current practices. *Sports Medicine*.
- Pietrosimone, B., Lepley, A. S., Harkey, M. S., et al. (2016). Quadriceps strength predicts self-reported function post-ACL reconstruction. *Medicine & Science in Sports & Exercise*, 48(9), 1671–1677.
- Sheenam, N., Gaur, R., Gonnade, N. M. et al. (2025). Knee functional outcomes and quadriceps hypotrophy after ACL reconstruction: a prospective observational study. *BMC Sports Science, Medicine and Rehabilitation*, 17(120).
- Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLoS One*, 14(11), e0224365.
- Wiggins, A. J., Grandhi, R. K., Schneider, D. K., Stanfield, D., Webster, K. E., & Myer, G. D. (2016). Risk of secondary injury in younger athletes after anterior cruciate ligament reconstruction: A systematic review and meta-analysis. *The American Journal of Sports Medicine*, 44(7), 1861–1876.
- Zhu, X., Kelly, D. K., Kim, G., Hart, J. M., & Gong, J. (2025). Clinically interpretable modeling of ACL reconstruction outcomes using confidence-aware gait analysis. *Biomechanics*, 5(4), 94.