Machine learning-based detection of pitching patterns in Major League Baseball: An analysis of pitch metrics prior to ulnar collateral ligament reconstruction

Ryotaro Ozaki

A thesis submitted to Auckland University of Technology in partial fulfilment of the requirements for the degree of Master of Sport, Exercise, and Health

Sports Research Institute New Zealand
AUT School of Sport, Exercise and Health
April 2026

Abstract

Purpose: Ulnar collateral ligament reconstruction (UCLR) is the most prevalent surgically treated injury among Major League Baseball (MLB) pitchers, yet early detection of pre-surgical biomechanical deterioration remains limited. The purpose of this study was to develop and evaluate an unsupervised machine learning-based anomaly detection framework capable of identifying multivariate pitching metric changes in MLB pitchers in the period preceding UCLR.

Methods: Pitch-tracking data from Statcast were collected for 46 MLB pitchers who underwent primary UCLR between 2016 and 2024. Pitcher-specific vanilla autoencoders were trained on game-level aggregated pitching metrics spanning a 400-day baseline window (200–600 days before last appearance) and applied across a 200-day detection window (0–200 days before last appearance). Reconstruction error served as the anomaly detection metric. Six pitch types were analysed: four-seam fastball, sinker, slider, cutter, changeup, and curveball. Five percentile-based reconstruction error thresholds (90th to 99th) were evaluated. As an exploratory validation analysis, a propensity score-matched control group of 45 non-UCLR pitcher pairs was employed to contextualise the specificity of the pre-surgical signal.

Results: Mean per-game reconstruction error escalated from 0.877 (151–200 days before last appearance) to 2.326 (0–50 days), representing a 2.7-fold increase with a broadly monotonic trajectory. At the 95th percentile threshold, anomaly rates were 67.2%, 76.4%, 71.0%, and 73.6% across the 151–200, 101–150, 51–100, and 0–50 day bins respectively. Escalation ratios increased monotonically with threshold stringency (p90: 0.98; p99: 1.18), with the most extreme deviations concentrated in the final 50 days. Feature-level analysis identified a two-phase deterioration structure: an early phase characterised by elevated slider movement and cumulative workload errors, followed by a late phase dominated by a 24-fold escalation in changeup usage error and sharp increases in rest interval deviation. In the exploratory matched control comparison, UCLR pitchers showed a statistically significant difference in median reconstruction error relative to matched controls in the proximate pre-surgical window ($W = 1245$, $p = 0.036$), with no equivalent directional escalation observed in controls. A Bonferroni-corrected comparison of multivariate and univariate detection identified only marginal gains from multivariate encoding in detection sensitivity, though the autoencoder provided structural interpretability and coherent feature-pattern identification not available through univariate monitoring.

Conclusion: Pitcher-specific autoencoder models applied to Statcast pitch-tracking data can identify progressive multivariate deterioration in the period preceding UCLR, with a signal that appears at least partially specific to the pre-surgical period relative to matched controls. These findings suggest that routine monitoring of individualised pitching profiles may provide a clinically actionable detection window prior to UCL failure.

Table of Contents

List of Figures

List of Tables

Attestation of Authorship

“I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor used artificial intelligence tools or generative artificial intelligence tools (unless it is clearly stated, and referenced, along with the purpose of use), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.”

Signature: 

Co-authorship Contributions within this Thesis

STUDENT AND SUPERVISOR APPROVALS					
By signing you are confirming that the co-author contributions stated in the table(s) below are accurate.					
Student Name	Ryotaro Ozaki	**Signature**	[redacted]	**Date**	16/4/26
Supervisor Name	Mike McGuigan	**Signature**		**Date**	16/4/26

Chapter Number:	**Chapter 3**
Manuscript Title:	**Machine learning-based detection of pitching patterns in Major League Baseball: An analysis of pitch metrics prior to ulnar collateral ligament reconstruction**
Publication Status:	**Unpublished/Ready for submission for Publication**

AUTHOR SURNAME:	CONTRIBUTION:
Ryotaro Ozaki	Conception, design, and conduct of the research.
Mike McGuigan	Drafting significant parts of the research output or critically revising it so as to contribute to its quality and interpretation.
Chris Whatman	Drafting significant parts of the research output or critically revising it so as to contribute to its quality and interpretation.
Tanuj Wadhi	Contributed expertise in machine learning methodology and model development.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my co-supervisors, Professor Mike McGuigan and Associate Professor Chris Whatman. From the very beginning of this project, their knowledge, guidance, and unwavering belief in my abilities set a strong foundation for this work. They were instrumental in keeping me on the right path while generously affording me the creative freedom to pursue my own ideas. Their remarkably prompt correspondence and consistent availability made every stage of this journey feel supported.

I would also like to extend my heartfelt thanks to PhD candidate Tanuj Wadhi, who gave so generously of his time and energy throughout this project. He went above and beyond in introducing me to an entirely new field of study in machine learning, offering fresh perspectives and new ways of thinking that proved invaluable. His openness and willingness to engage at any stage of the process is something I am deeply grateful for.

Finally, and most personally, I would like to thank my fiancée, my whānau, and my friends, whose love and support have carried me through every stage of this journey. Her presence is felt in everything I do, and I am profoundly grateful to have her by my side. My grandparents and mother have been a constant bedrock throughout my life, raising me to trust my own judgement and supporting every path I chose to pursue. My closest friends kept my life balanced and grounded in the best possible way. The shared dinners, the late evenings working side by side, and simply spending time together are memories that will stay with me long after this chapter closes.

I also acknowledge my use of Claude (Anthropic, 2026) as an AI writing assistance tool during the preparation of this thesis. I used Claude to support my writing development, including improvements to grammar, sentence structure, and chapter organisation, to assist with R script debugging and code structure, and to receive editorial feedback on draft text. All AI-assisted content was critically reviewed, substantially revised, and approved by me. The substantive intellectual content, research design, data analysis, and academic arguments presented in this thesis are my own work.

Artificial Intelligence Declaration

Specify the chapter number(s) in your research proposal/research component where AI was used.	
Chapter number(s):	Chapters 2, 3 and 4
Briefly explain why the AI tool was used (e.g., idea generation, summarization, data analysis, image creation, etc.).	
Purpose of AI Use:	The AI tool was used to support writing development, including improving grammar, sentence structure, and prose clarity, as well as to assist with chapter organisation and the iterative refinement of academic writing. The tool was also used to provide critical feedback on existing draft text and to assist with literature search support and synthesis of relevant sources. The AI tool was additionally used to assist with R script debugging, error resolution, and code structure, specifically in relation to data extraction, pre-processing, propensity score matching, and autoencoder model implementation pipelines used in the empirical analysis.
Clearly state the AI tools used.	
AI Tool(s) Used:	Claude (Anthropic), accessed via claude.ai
Prompts or queries entered into the AI tool (only the instructive portion of the prompt is required). Include all relevant variations.	
Prompts Used:	Critique and provide feedback on this chapter draft; Suggest edits to improve sentence clarity and academic tone; Help structure and organise this section; Provide feedback on the framing of this argument; Review this paragraph for consistency with the preceding argument; Debug this R script error; Help restructure this R code for [specific task]; Identify why this function is not producing the expected output; Suggest improvements to the structure of this data processing pipeline.
A summary or description of the AI-generated output. (Optional: attach output in an appendix.)	
Output Received:	The AI tool provided editorial suggestions, structural feedback, and draft text for consideration. Suggested revisions to sentence structure, paragraph organisation, and academic framing were generated in response to existing student-authored text. Literature summaries and citation suggestions were also provided for review. In relation to R scripting, the AI tool provided corrected code, debugging suggestions, and restructured scripts in response to student-submitted code and error messages. Outputs included corrections to data extraction pipelines using the baseballr package, propensity score matching scripts, velocity calculation procedures, and autoencoder model implementation code. In all cases, outputs were provided in direct response to student-submitted code and specific error descriptions rather than generated independently.
Describe how the AI output was used, edited, or transformed.	
Post-AI Processing Methods:	All AI-generated suggestions and draft text were reviewed, critically evaluated, and substantially revised by the student prior to inclusion in the thesis. Final decisions regarding content, argument, and framing were made by the student in consultation with supervisors. AI-generated text was not incorporated verbatim; all output was treated as a drafting aid and edited to reflect the student's own voice, argument, and academic judgement. The substantive intellectual content, research design, data analysis, and findings presented in the thesis are entirely the student's own work. R code suggestions were tested, evaluated, and modified by the student to ensure correct implementation. All analytical decisions regarding methodology and code logic remained the student's own.

STUDENT DECLARATION					
By signing you are confirming that the AI use stated in the table(s) above are accurate and follow AUT's recommended guidelines detailed in the AI Hub and Postgraduate Handbook.					
Student Name:	Ryotaro Ozaki	**Signature:**	[redacted]	**Date:**	**29/03/2026**

Chapter 1. Introduction

Injury surveillance and performance monitoring are increasingly important components of player management in professional baseball, with approaches ranging from wearable sensors to biomechanical assessments aimed at identifying factors that predispose pitchers to injury. Identifying analytical methods that improve the sensitivity and individualisation of these frameworks remains an important direction for the field.

Ulnar collateral ligament reconstruction (UCLR) is the most prevalent surgically treated injury in professional baseball, affecting approximately one quarter of MLB pitchers across their careers (Conte et al., 2015; Leland et al., 2019). UCLR incidence has risen substantially over the past five decades, with a particularly sharp surge between 2010 and 2015 (Almonroeder et al., 2024). The injury typically presents in pitchers during the early career period, with a mean age at surgery of 24.4 years, and carries substantial personal and economic consequences (Camp et al., 2018). Recovery requires 16.8 to 20.5 months of rehabilitation, and the average financial cost to MLB clubs has been estimated at approximately USD 1.9 million per affected athlete (Erickson et al., 2014; Meldau et al., 2020). These figures underscore the significance of developing reliable early detection frameworks capable of identifying at-risk pitchers before clinical injury presentation.

The biomechanical determinants of UCL injury are well established in the laboratory literature. Elbow valgus torque (EVT) represents the primary mechanical pathway to ligament compromise, and its magnitude is shaped by the integrated contributions of lower body mechanics, trunk rotation timing, shoulder kinematics, and arm path dynamics across the kinetic chain (Aguinaldo & Chambers, 2009; Crotin et al., 2022). Crucially, meaningful injury risk patterns are highly individualised; velocity-torque and arm path relationships are consistently stronger within pitchers than between them, which limits the utility of population-level thresholds for injury surveillance (Sakurai et al., 2024; Slowik et al., 2019). Pre-injury changes in pitch velocity, release point, and spin characteristics have been observed in the period preceding UCLR, though the mechanisms underlying these changes remain incompletely understood (Mayo et al., 2021; Portney et al., 2019). Their multivariate and individually variable nature means that population-level supervised models have yielded only moderate predictive performance, in part because meaningful within-pitcher patterns are obscured when individual biomechanical profiles are aggregated (Kang et al., 2025; Karnuta et al., 2020; Oeding et al., 2024).

These limitations support an individualised, unsupervised approach to pitcher monitoring, in which models are trained without labelled injury events and derive each pitcher's typical performance profile directly from unlabelled baseline data. Anomaly detection methods based on autoencoders can learn a pitcher's characteristic performance profile from routine game data and flag deviations from that individual baseline, without requiring labelled injury events for training. By applying this approach to publicly available Statcast pitch-tracking data, which captures the downstream effects of biomechanical efficiency through metrics such as velocity, movement, spin rate, and release point, it becomes possible to monitor pitchers continuously

during competition. This thesis develops and evaluates such a framework, contributing to the development of more actionable and individualised monitoring tools in professional baseball.

Taken together, these observations identify a clear gap. Existing supervised, population-level models do not capture the individualised, multivariate nature of pre-injury change, and no published approach has applied unsupervised anomaly detection to this problem. This thesis therefore asks two research questions. Can pitcher-specific anomaly detection identify changes in pitching metrics in the games preceding UCLR? And how might such an approach enhance understanding of the risk factors contributing to UCL injury? It is anticipated that individualised autoencoder models will reveal progressive multivariate deviations in the pre-surgical period that population-level methods do not detect.

In Chapter 2, the prevalence and consequences of UCL injury in professional pitching are reviewed, alongside the biomechanical evidence base motivating an individualised, data-driven detection approach. Due to this thesis being presented in manuscript format, there is some repetition between the literature review in Chapter 2 and the introduction of the study in Chapter 3. An outline of the thesis is presented in Table 1.1. Chapter 2 is a narrative literature review examining the kinematic and kinetic determinants of UCL injury in high-level adult baseball pitchers. Chapter 3 is an empirical study addressing the two research questions posed above. Chapter 4 discusses the findings of Chapter 3 in an applied context and offers directions for future research.

Table 1.1 Thesis outline

Chapter	Title	Purpose
1	Introduction	General introduction to the thesis
2	Biomechanical and kinetic determinants of UCL injury in baseball pitchers	Narrative literature review on kinematic and kinetic risk factors for UCL injury in high-level adult baseball pitchers
3	Machine learning-based detection of pitching patterns in Major League Baseball: an analysis of pitch metrics prior to UCLR	An empirical study applying pitcher-specific autoencoder models to Statcast pitch-tracking data to detect multivariate performance changes preceding UCLR
4	Summary and future direction	Overall discussion situating the findings of Chapter 3 in an applied context, with limitations and future research directions

Chapter 2. Kinematic and Kinetic Determinants of Ulnar Collateral Ligament Injury in Pitchers

2.1. Introduction

This chapter aims to review and summarise the current academic literature on kinematic and kinetic risk factors for ulnar collateral (UCL) injuries in high-level adult baseball pitchers. It will focus on how joint-level and whole-body pitching mechanics, including elbow valgus torque (EVT), shoulder external rotation (SER) range of motion and torque, arm slot (AS), trunk rotation (TR) velocity and timing, hip-shoulder separation, stride length, and release location, contribute to elbow joint loading and injury risk. The review will also summarise studies that analyse segmental sequencing, mechanical efficiency, and compensation strategies linked to UCL loading, as well as temporal and spatial coordination patterns that may increase injury risk. By synthesising this evidence, the chapter will highlight the diverse methodological approaches used in this field and provide context for applying ML-based techniques to detect multivariate pitching patterns that may precede ulnar collateral ligament reconstruction (UCLR).

In March 2026, a comprehensive literature search was conducted using MEDLINE, Scopus, SPORTDiscus (EBSCO), PubMed, and Google Scholar to identify relevant studies. The search strategy included the terms "UCL reconstruction," "UCL repair," "pitching biomechanics/mechanics," "arm slot," "release point," "shoulder external rotation," "elbow valgus load/stress," "pitch velocity/speed," "elbow torque," "kinematics," and "kinetics" combined with "baseball pitchers" and the names of high-level competition levels, including "Major League Baseball," "Minor League Baseball," "Nippon Professional Baseball," and "Korea Baseball Organisation." The search was restricted to peer-reviewed journal articles published in English within the last 20 years. Studies were included only if they focused on high-level adult baseball pitchers competing at collegiate (National Collegiate Athletic Association Division I), professional (MLB, Minor League Baseball, Nippon Professional Baseball, Korea Baseball Organisation), or national and international team levels. Table 2.1 provides a summary of the studies reviewed.

Table 2.1 *Overview of studies examining kinematic and kinetic factors associated with UCL injuries in high-level pitchers*

Study	Study Design	n	Level	Kinematic/Kinetic Variables	Method	Key Findings	Association with UCL Injury
Aguinaldo and Chambers et al. (2009)	Descriptive laboratory	69	NCAA, MiLB, MLB	EVT, SER angle, Elbow flexion, TR kinematics, AS	3D motion capture analysis	6 parameters associated with peak EVT. Three parameters accounted for 68% of variance in EVT: Peak SER: 169° ± 15° ($r = .60$, $p < .01$), elbow flexion at peak valgus torque: 41° ± 24° ($r = -.36$, $p < .01$), EV loading rate: 29 ± 14 N·m ($r = .74$, $p < .01$).	Increased peak EVT, which has been implicated in UCL loading, was influenced by early TR, reduced elbow flexion, and a sidearm slot position.
Aguinaldo and Escamilla et al. (2019)	Descriptive laboratory	16	MiLB, MLB	PV, EVT, Pelvis/trunk kinetics & kinematics, SER angle, SIR velocity	3D motion capture analysis	PRO had higher absolute peak EVT than HS (71.3 ± 20.0 vs. 50.7 ± 14.6 N·m, $p < .01$), but similar when normalised for body size (0.04 ± 0.01 bw-h both groups). Trunk power and peak SER angle explained 78.4% of peak EVT variance ($r = 0.886$, $p < .001$).	Earlier TR and reduced trunk power efficiency produced less efficient movement patterns with reduced PV, potentially increasing UCL injury risk.
Anz et al. (2010)	Cohort	23	MLB	EVT, SER torque	Kinematic video analysis	Injured pitchers showed significantly higher EVT (91.62 vs 74.70 N·m, $p = .01$) and SER torque (89.83 vs 70.96 N·m, $p < .01$) at peak SER during late cocking phase compared to non-injured pitchers.	Late cocking phase critical for elbow injury risk. Modifying mechanics or using torque measures for risk identification may reduce injury rates.
Barrack et al. (2024)	Descriptive laboratory	87	NCAA	EVarT, PV	IMU kinematics, physical capacity assessments	Higher PV (1.85 N·m per m/s), bodyweight (0.87 N·m per kg), hip abduction, grip strength symmetry, and hip bridge hold time increased EVarT (all $p < .01$). Greater dominant-shoulder flexion ROM, internal rotation strength, and hip external rotation ROM decreased EVarT ($p < .05$ to $p < .01$).	Grip strength imbalance, lead-leg stability, and higher bodyweight increase EVarT; stronger dominant SIR, better shoulder flexion ROM, and balanced scaption strength reduce it. Targeting these factors may reduce UCL injury risk.
Beaudry et al. (2022)	Cohort	223	MLB	Knee angle, Pitching style (TF/DD), Pelvis kinematics, PV	Kinematic video analysis	Among 223 MLB pitchers who underwent UCLR, 72.6% were TF pitchers vs 27.4% DD pitchers (162 vs 61). No significant associations found between pitching style and year of UCLR, BMI, age, handedness, or average PV (all $p > .05$).	TF style overrepresented in MLB pitchers undergoing UCLR. Modifying TF-associated mechanics may reduce EVT and UCL injury incidence.

Beaudry et al. (2023)	Cohort	660	MLB	Knee angle, Pitching style (TF/DD), Pelvis kinematics, PV	Kinematic video analysis	Of 660 pitchers (2019), 62.4% used TF style, 37.6% DD style. Upper extremity injuries occurred in 27.2% of TF vs 15.3% of DD pitchers ($p < .001$). UCLR occurred in 10 TF vs 2 DD pitchers (1.8% overall). UCLR history: 32.8% TF vs 22.6% DD pitchers ($p = .005$).	TF style associated with higher UCL injury rates vs DD style. Targeting TF-specific features such as extended drive-leg knee angles may reduce injury risk.
Camp et al. (2017)	Descriptive laboratory	81	MLB, MiLB	EVarT, AS, Arm speed, Arm external rotation	IMU kinematics	Analysis of 82,000 throws showed each 1 N·m increase in peak EVarT raised arm speed by 116 °/s, increased arm rotation by 8°, and reduced AS by 13° (all $p < .001$). Mean EVarT was 60 ± 15 N·m. Height and weight correlated positively with EVarT ($r = .26$-$.54$, $p \leq .02$).	Increased arm speed, greater arm rotation, and lower AS elevated EVarT, indicating altered mechanics that may increase UCL injury risk despite potential PV gains.
Chalmers et al. (2016)	Case-control	1327	MLB	PV	Archival data analysis	Pitchers with UCLR had higher peak PV (93.3 vs 92.1 mph) and mean PV (87.8 vs 86.9 mph, both $p \leq .001$). Dose-response relationship: 20% with peak PV ≥95.7 mph vs 7.8% with peak PV ≤86.9 mph needed UCLR. Higher BMI, weight, and younger age were also linked to UCLR ($p \leq .01$).	Higher peak PV strongest predictor of UCL tears. Identifying and addressing excessive peak torque patterns may reduce injury incidence.
Cohen et al. (2022)	Case-control	327	MLB	PV, Release location, Ball movement	Archival data analysis	UCL-injured pitchers had more lateral release points (+5.8 cm, $p = .028$) and shifted 2.4 cm laterally over 3 years vs 1.5 cm medially in controls ($p = .016$). Each 2.5 cm lateral shift increased UCLR odds by 3.7%. Injured pitchers had higher BMI (27.6 vs 27.0, both $p \leq .04$).	Lateral release location predicts UCL tears. Monitoring and correcting lateral shifts, particularly in relief pitchers, may reduce injury risk.
Crotin et al. (2022)	Descriptive laboratory	545	PRO, NCAA	EVarT, AS, PV, Joint angles, Release location, Ball movement	3D motion capture analysis	PRO showed greater biomechanical efficiency than NCAA (711.0 vs 657.0; $p < .001$) with higher PV (37.9 vs 35.8 m/s). High-efficiency pitchers had lower normalised torque (0.047 vs 0.063; $p < .001$).	Lower normalised EVarT, greater peak SER angle, and reduced elbow flexion/shoulder abduction at SFC associated with higher efficiency, indicating patterns that may reduce UCL injury risk.

DeZee et al. (2025)	Cross-sectional	44	NCAA	EVarT, PV, Pelvic/trunk kinetics & kinematics	IMU kinematics, Lumbopelvic stability assessments	Two clusters identified: low torque–high velocity (48.3 N·m, 84.8 mph) and high torque–low velocity (61.6 N·m, 81.3 mph; both $p < .001$). Greater transverse trunk (OR = 2.9) and pelvis motion during SLSD predicted high torque–low velocity.	Excessive transverse trunk and pelvis motion during SLSD predicted high EVarT with low PV, indicating lumbopelvic instability increases elbow loading. Improving transverse-plane control may lower UCL injury risk.
Dillon et al. (2025)	Retrospective case-control	7	MLB	PV, Spin rate, Spin Axis, Ball movement, Release point, Release extension	Archival data analysis	Every injury pitch exceeded the 95th percentile MD threshold. PV suppressed by a mean of 2.1 SD and AS angle reduced by a mean of 1.5 SD at the injury pitch. 86% of pitchers showed elevated cumulative MD across the 5 FF preceding injury vs. 7% of matched controls ($p < .001$).	Acute UCL failure appears to reflect a distinct biomechanical pattern characterised by short-term within-game mechanical volatility and abrupt decompensation rather than prolonged cross-outing deterioration.
Dong et al. (2025)	Descriptive laboratory	66	NCAA	AS angle, EVarT, Segment and joint kinetics & kinematics	Archival 3D motion capture data analysis	Pitchers with minimal AS (43°) had higher EVarT (6.7 vs 6.0; $p = .006$) but lower efficiency (5.8 vs 6.7; $p = .003$) than maximal AS (67°). Each 10° AS increase reduced torque and raised efficiency. Minimal AS pitchers showed greater contralateral trunk tilt at MER (–27° vs –16°; $p < .001$).	Higher AS angle, reduced trunk tilt, and improved torque efficiency associated with lower EVarT, indicating optimising AS and trunk posture may reduce medial elbow stress and enhance efficiency.
Dowling et al. (2023)	Descriptive laboratory	182	MLB, MiLB	EVarT, PV, Arm path	3D motion capture analysis	Shorter early arm paths reduced EVarT but also PV. Within pitchers, longer arm paths strongly correlated with higher PV ($R^2 = 0.788$) and EVarT ($R^2 = 0.962$); each 30 cm increase raised PV by 0.79 mph and torque by 1.29 N·m.	Shorter early arm path associated with lower EVarT but reduced PV, indicating trade-off where optimising early arm mechanics may lower UCL injury risk with modest performance cost.
Dowling et al. (2024)	Cross-sectional	322	MLB, MiLB	EVarT, PV, Lead knee kinetics & kinematics	3D motion capture analysis	High-velocity pitchers had greater lead knee extension (17° vs 5°), higher extension velocity (419°/s vs 297°/s), and higher EVarT (91.1 vs 84.0 N·m; all $p < .001$). Lead knee extension and velocity predicted PV ($R^2 = 0.35, 0.33$) but together explained only 13%.	Greater lead knee extension associated with higher PV but not increased EVarT, suggesting enhancing lead leg mechanics may improve velocity without elevating UCL injury risk.

Escamilla et al. (2018)	Cross-sectional	322	MLB, MiLB	EVarT, AS angle, Shoulder/Trunk/ Pelvis kinetics & kinematics	Archival 3D motion capture data analysis	Sidearm showed higher elbow flexion torque (70 vs 61 N·m; p=.01) and SER (169° vs 163°; p=.01) than overhand but trended towards lower EVarT (88 vs 97 N·m; p=.14). Key differences: trunk tilt (overhand: 32°, sidearm: 2°) and shoulder abduction (overhand: 94°, sidearm: 81°; both p<.001).	Lower AS angle, reduced shoulder abduction, and decreased trunk contralateral tilt associated with lower EVarT, suggesting sidearm delivery may reduce medial elbow loading and UCL stress without affecting PV.
Escamilla et al. (2023)	Controlled laboratory	215	PRO	EVarT, Segment and joint kinetics & kinematics	Archival 3D motion capture data analysis	Moderate CTT (15°–25°) produced greater shoulder and elbow forces than minimal (0°–10°) or maximal (30°–40°) tilt, with no PV difference. This group had highest elbow flexion torque (69 N·m), shoulder anterior force (403 N), and elbow proximal force (1145 N; all p<.01). EVarT did not differ (p=.096).	Moderate CTT was associated with greater shoulder and elbow forces, suggesting a trade-off where adopting a three-quarter AS may elevate joint loading and injury risk, despite no clear performance benefit in PV.
Fleisig et al. (2015)	Controlled laboratory	80	MLB	EVarT, PV, SIR torque, Joint angles/velocities, Stride length,	3D motion capture analysis	Pitching biomechanics and passive ROM did not differ between UCLR and non-injured pitchers. EVT (99 vs 99 N·m; p=.86), PV (38.3 vs 38.3 m/s), and other key variables showed no significant differences (all p>.16). No compensatory mechanics evident post-UCLR.	Kinematic patterns such as late shoulder rotation and extreme shoulder abduction have been theorised to increase elbow torque; however, post-UCLR pitchers showed no evidence of these mechanics.
Hodakowski et al. (2025)	Descriptive laboratory	31	MLB, MiLB	EVarT, PV, Cumulative torque, Torque loading rate, Spin rate, Spin efficiency	3D motion capture analysis	FF generated highest PV (40.4 vs 33.4–36.4 m/s) and EVarT (90.1 vs 81.3–87.7 N·m) versus other pitches (all p<.001). They also showed highest cumulative torque (3015 N·m·s) and loading rate (780.7 N·m/s), exceeding others by 3–15%. Spin rate did not correlate with elbow torque.	FF showed higher PV, EVarT, cumulative torque, and loading rate, suggesting that frequent FF use may elevate medial elbow stress and UCL injury risk, independent of spin rate.
Keller et al. (2016)	Case-control	83	MLB	PV	Archival data analysis	No significant differences in pre-injury PV between UCLR and controls for any pitch type. UCLR pitchers threw more FF (46.8% vs 39.7%; p=.035), with each 1% increase linked to 2% higher UCL injury risk. Exceeding 48% FF significantly predicted UCL injury (p=.006).	A high percentage of FF was a main UCL risk factor, with repeated high-velocity pitching leading to overuse injuries. Absolute PV did not increase risk, as UCLR pitchers had similar PV to controls for all pitch types.

Lipa et al. (2025)	Case-control	498	MLB	Elbow angle at peak valgus stress, Release point angle	Kinematic video analysis	Elbow angles at peak valgus stress and release showed no significant link to UCL reconstruction risk (Angle A: OR=1.02, p=.14; Angle B: OR=0.99, p=.20). No pitch type showed significant interaction between elbow angle and injury odds.	Elbow angles at peak stress and release did not predict UCL surgery risk, and no pitch type showed a significant link with elbow angle. This suggests arm angle alone does not meaningfully affect injury risk.
Lizzio et al. (2020)	Descriptive laboratory	12	MiLB	EVT, Arm speed, AS, Shoulder rotation, PV	Kinematic analysis using a wearable IMU	FF produced highest EVT (54.3±7.0 N·m), with CU 4% lower (52.1 N·m, p=.001). Higher BMI linked to lower EVT (β=–1.83, p=.035). PV differed by pitch (FF: 85.2 mph; CH: 77.9 mph; CU: 73.1 mph; all p<.0001) but did not correlate with torque (r=0.30, p=.69).	FF generated highest medial elbow torque, identifying them as primary UCL stressor. PV not correlated with medial elbow torque, indicating velocity alone does not explain UCL injury risk.
Mastroianni et al. (2025)	Case-control	117	MLB	PV, Spin rate, Spin Axis, Ball movement, Release point, Release extension	Archival data analysis (Statcast database)	Pitchers who underwent UCL surgery threw harder (89.9 vs 89.1 mph; p=.01), with each 1 mph increase raising surgery odds by 20% (aOR=1.20; p=.03). Lower FF usage linked to higher surgery risk (aOR=0.07; p=.05).	Higher PV increased UCLR risk regardless of workload. Lower FF use raised odds, while off-speed reliance added strain. Higher Pitching+/Location+ in surgical cases suggest command may elevate mechanical load.
Mastroianni et al. (2026)	Case-control	78	MLB	PV, Spin rate, Spin Axis, Release point, Release extension	Archival data analysis (Statcast database)	UCLR pitchers showed greater baseline intra-outing variability in PV (p = .012) and horizontal release position(p = .005), across sequential time windows, rising spin rate during ΔS2 (p = .019), and a progressive PV decline across the final five outings (slope = −0.082 mph/outing; p = .019); controls remained stable.	Greater intra-outing mechanical variability, particularly in PV and horizontal release position, may represent a baseline risk factor for UCL injury independent of average performance values. PV decline in the final five outings represent the most proximal detectable signal.
Mayberry et al. (2020)	Case-control	274	MLB, MiLB	CMJ metrics	Archival data analysis, Physical capacity assessments	CMJ variables did not predict shoulder injury (p=.694) but improved elbow injury prediction (p=.003). Low ERFD strongly predicted higher risk (β=–0.604; p=.002). Two high-risk profiles: low AVCF with high CVI (45%) and high AVCF with low CVI (19%).	CMJ variables ERFD and AVCF: lower ERFD and imbalanced concentric profiles linked to elbow injury, suggesting poor force transfer elevates UCL stress.

Mayo et al. (2021)	Retrospective case series	223	MLB	PV, Spin rate	Archival data analysis (PITCHf/x technology data)	Pitch characteristics changed in 15 games before UCLR. PV declined for FF (–0.657; p<.001), SI (–0.429; p=.029), and SL (–0.524; p=.008), starting 8–9 games pre-injury. Spin rates dropped for FF (–0.581; p=.003) but rose for FC. CU usage increased (0.486; p=.013).	Decreased PV and altered spin rates in 15 games before UCLR indicated injury, with progressive velocity declines suggesting cumulative fatigue and increased UCL stress.
McCutcheon et al. (2025)	Descriptive laboratory	523	PRO, NCAA	EVarT, PV, Lower body/trunk kinematics, Shoulder & elbow kinematics	Archival 3D motion capture data analysis.	Higher normalised EVarT pitchers: greater PV (38.0 vs 37.1 m/s), shoulder abduction (91.3° vs 86.1°), elbow flexion at foot contact (99.3° vs 93.6°), knee/elbow extension speeds (all p<.001). Velocity explained most torque variance (R^2=0.161).	Higher EVarT linked to increased PV, greater shoulder abduction/elbow flexion at SFC and reduced efficiency. Inefficient force transfer and delayed arm positioning may heighten medial elbow stress.
Oeding et al. (2024)	Case-control	3808 pitcher-years	MLB	PV, Spin rate, Ball movement, Release extension, Spin axis	Archival data analysis	Higher mean PV increased injury risk, driven by faster FF (93.63 vs 93.08 mph) and SL (84.57 vs 83.99 mph). Greater SL usage, FF spin (2265 vs 2241 rpm), more horizontal FF movement (–4.33 vs –3.44 in), and younger age (30.94 vs 31.65 years) were key predictors. Workload had no significant impact.	Higher PV, FF spin, and SL usage predicted future injuries. Combined pitch traits, not workload alone, better reflected biomechanical stress, suggesting pitch-tracking data can identify injury risk.
Oi et al. (2019)	Descriptive laboratory	38	NPB, MiLB, PRO	EVarT, SIR torque, Lower body/trunk kinematics, Shoulder & elbow kinematics	Archival 3D motion capture analysis	Japanese pitchers had longer normalised strides (86% vs 82%; p=.023), greater pelvic rotation (47° vs 31°; p<.001), and higher SIR velocity (8476 vs 6842°/s; p=.002). American pitchers showed higher EVarT (99 vs 86 N·m; p=.018) and greater PV (38.1 vs 34.7 m/s; p<.001).	Greater EVarT was the main UCL injury predictor. American pitchers showed higher EVarT, PV, and elbow flexion torque meaning higher PV increased medial elbow stress and joint loading risk.
Portney et al. (2019)	Case-control	71	MLB	PV, Release location	Archival data analysis	UCLR pitchers showed wider lateral release point (12.2 cm; p=.001) and progressive lateral shift over two years (3.4 cm; p=.036). Vertical release was lower before surgery (–4.8 cm; p=.045). UCLR pitchers used more CU (p=.020), and injured relievers threw fewer pitches per game (p=.046).	More lateral and lower vertical release points predicted UCL injury. Injured pitchers used more CU, and injured relievers threw fewer pitches per game, suggesting pitch mix and recovery may increase risk.

Post et al. (2015)	Cross-sectional	67	NCAA	EVT, SER torque, Shoulder-distraction force, PV, Lower body/trunk kinematics, Shoulder & elbow kinematics	3D motion capture analysis	Peak PV showed weak positive correlation with shoulder-distraction force ($r=0.257$; $r^2=0.066$; $p=.018$) but no significant correlations with EVT ($r=0.199$; $p=.053$) or SER torque ($r=0.097$; $p=.217$). Joint kinetics were largely independent of PV, with very low variance explained ($r^2<0.07$).	EVT and SER torque remain key UCL risk factors but showed no strong link to PV. Inefficient mechanics such as poor timing, excessive trunk tilt, and altered arm positions likely contribute more to joint stress than PV alone.
Prodromo et al. (2016)	Case-control	114	MLB	PV, EVT, Shoulder & elbow kinetics	Case-control study	Pitch type distribution did not differ between UCLR cases and controls ($p>.05$). PV was higher in UCLR group. FF velocity was strongest predictor (92.08 vs 91.33 mph; OR=1.15; $p=.001$). SL, CU, and CH velocities were also significant predictors, but removed from the model due to multicollinearity.	Higher PV increased UCL injury risk due to greater elbow stress at higher speeds. Pitch type did not affect risk, showing throwing intensity matters more than pitch choice. Managing PV important for UCL protection.
Qiao et al. (2025)	Descriptive laboratory	19	NCAA	EVarT, PV, Joint angles, CMJ metrics, GRF measurements	Archival data analysis, Physical capacity assessments	High-EVarT pitchers showed greater EVarTRTD than low-EVarT peers (605±74 vs 353±103 $N \cdot m \cdot s^{-1}$; $p<.001$), indicating faster medial elbow loading. CMJ metrics were mostly similar except higher peak GRF in high-EVarT group (2470 vs 2060 N; $p=.014$).	Higher EVarTRTD was a key UCL risk factor, showing faster elbow loading in high-EVarT pitchers. CMJ performance was similar, but moderate links with CMJ force measures suggest lower-body force reflects elbow stress.
Sakurai et al. (2024)	Descriptive laboratory	81	NCAA	EVarT, PV	Kinematic analysis using a wearable IMU	Higher PV (M = 37.9 m/s) related to greater EVT but explained little variance across pitchers ($R^2 = 0.05$; $\beta = 1.82$; $p = .046$). Within pitchers, PV explained more ($R^2 = 0.29$; $\beta = 1.97$; $p < .001$). Only 48% showed a significant T-V link, and effort had no effect ($p = .30$).	Higher EVarT increased UCL injury risk, but the link between PV and torque varied widely. PV alone did not reliably predict torque, and pitch effort had no effect.
Slowik et al. (2019)	Cross-sectional	64	MLB, MiLB	PV, EVT, Shoulder & elbow kinematics	Archival 3D motion capture data analysis	Between pitchers, PV explained only 7.6% of EVarT variance ($R^2 = 0.076$; $p = .03$), suggesting other factors influence torque. Within pitchers, the link was strong ($R^2 = 0.957$; $p < .001$), with EVarT rising 1.62 Nm for each 1.0 m/s increase in PV.	Higher PV increased EVarT within pitchers, showing that throwing harder raises elbow load and UCL stress. Between pitchers, this link was weak due to mechanical differences.

Tanaka et al. (2024)	Cross-sectional	18	COL, REG	PV, EVT, SER torque, TR velocity, Lower body kinematics, GRF measurements'	3D motion capture analysis, Force plate assessment	More suppressed TR velocity at peak pelvic rotation related to faster PV (R = -0.504; p = .033), higher EVT (R = -0.508; p = .032). Longer pelvic–TR velocity delay also related to greater PV (R = 0.473; p = .047) but not joint torques (p > .10).	Suppressed TR velocity at peak pelvic rotation predicted higher EVT and greater joint loading, showing a trade-off between performance and injury risk.
Whiteside et al. (2016)	Case-control	104	MLB	PV, Release location, Spin rate	Archival data analysis (PITCHf/x database)	A less lateral release point (OR = 0.028; p = .025) and shorter stature (OR = 0.941; p = .013) increased surgery odds. Higher mean PV (OR = 1.381; p = .005) and pitches per game (OR = 1.020; p = .003) also raised risk.	Higher mean PV increased UCLR risk, with each 1 m/s rise raising odds by 38%. A less lateral release point also raised risk, indicating velocity and mechanics jointly shape UCL injury in elite pitchers.

Note. AS = arm slot; AVCF = average concentric force; BMI = body mass index; bw-h = body weight to height; CH = changeup; CMJ = countermovement jump; COL = collegiate pitchers; CTT = contralateral trunk tilt; CU = curveball; CVI = concentric velocity index; ERFD = eccentric rate of force development; EVT = elbow valgus torque; EVarTRTD = elbow valgus torque rate of torque development; FC = cutter; FF = four-seam fastball; GRF = ground reaction force; HS = high school pitchers; IMU = inertial measurement unit; MD = Mahalanobis distance; MER = maximum external rotation; MiLB = Minor League Baseball pitchers; MLB = Major League Baseball pitchers; NCAA = National Collegiate Athletic Association pitchers; NPB = Nippon Professional Baseball pitchers; OR = odds ratio; aOR = adjusted odds ratio; PRO = professional pitchers; PV = pitch velocity; REG = regional baseball team pitchers; ROM = range of motion; SER = shoulder external rotation; SFC = stride foot contact; SI = sinker; SIR = shoulder internal rotation; SL = slider; SLSD = single-leg squat down; TF/DD = three-quarter/drop down pitching styles; TR = trunk rotation; UCL = ulnar collateral ligament; UCLR = ulnar collateral ligament reconstruction; ΔS2 = immediately preinjury versus early injury season.

2.2. Summary of the Studies Reviewed

Of the 37 studies included in this review, the majority were descriptive laboratory studies (13, 35%), with several retrospective case-control (12, 32%) and cross-sectional (6, 16%) designs also represented. All studies were published between 2009 and 2026. The aggregate sample comprised 11,096 high-level adult baseball pitchers, with the majority drawn from professional leagues (n = 10,466, 94%), followed by collegiate programs (n = 635, 6%), and regional competitions (n = 5, <1%). One study utilised pitcher-years as the primary analytical unit (n = 3,808), wherein each pitcher-year constituted a discrete data point representing the comprehensive performance metrics and injury status of an individual pitcher throughout a single competitive season.

The studies collectively analysed a wide range of kinematic and kinetic variables, including EVT, SER, AS angle, release point, PV, and segmental sequencing. While findings varied across studies, most reported significant associations between mechanical patterns and either increased UCL loading or reconstruction risk, reflecting the complex interaction of pitching mechanics in injury development. The primary methods employed across the literature included three-dimensional motion capture analysis, wearable inertial measurement units (IMUs), archival video and pitch-tracking data, and physical capacity assessments. These methodological approaches yielded both comprehensive laboratory-derived biomechanical data and extensive retrospective epidemiological insights obtained from competitive environments, thereby providing complementary analytical perspectives on UCL injury risk. For laboratory-based studies, some limitations must be acknowledged. While pitchers were instructed to throw with game-like effort and delivered from a mound, the controlled environment was atypical compared with practice and competition. Laboratory-based assessments appear to capture general patterns of pitching mechanics, but differences in intensity, fatigue, and context may result in deviations in absolute biomechanical values and effort compared with in-game performance (Lerch et al., 2025; Whiteley, 2007). As such, observed mechanics may not fully represent competitive pitching.

Taken together, the reviewed studies highlight the diversity of methodological approaches used to investigate UCL injury mechanisms and emphasise the importance of integrating both controlled biomechanical data and competition-derived performance metrics. This synthesis provides the necessary context for subsequent sections, which focus on key variables such as EVT, pitch velocity, AS, and biomechanical efficiency.

2.3. Elbow Valgus Torque: Mechanisms and Risk Factors, and Biomechanical Efficiency

2.3.1. Elbow Valgus Torque

EVT is the primary external load placed on the medial elbow during the cocking phase of pitching, driving the forearm into a valgus position that stresses the UCL (Matsuo et al., 2006). The cocking phase culminates at maximum external rotation (MER) of the throwing shoulder, the moment at which EVT peaks and the mechanical demands on the medial elbow are greatest (Fleisig et al., 1995). In adult pitchers, peak EVT at MER approaches 70 newton-metres (N·m)(Aguinaldo & Escamilla, 2019; Fleisig et al., 1995), generating a valgus load that the UCL and surrounding musculature counteract through an internal stabilising force known as elbow varus torque (EVarT) (Werner et al., 1993). Although muscular varus moments can offset much of this demand, some degree of valgus stress on the UCL remains unavoidable (Yanai et al., 2025).

Biomechanical analyses indicate that the UCL, in conjunction with surrounding musculature and bony articulation, contributes approximately one-third of the peak EVarT, meaning the UCL provides around 33 N·m of torque (Buffi et al., 2015; Slowik & Fleisig, 2019). This torque approaches the UCL's physiological failure threshold, estimated at 30 to 35 N·m in cadaveric models (McGraw et al., 2013; Morrey & An, 1983; Udall et al., 2009). This narrow margin between physiological failure limit and in-game loading emphasises EVT as a primary mechanism by which repetitive pitching can compromise ligament integrity. Recent evidence further identifies elevated EVT, EVarT and specific kinematic patterns as significant indicators of UCL injury risk (McCutcheon et al., 2025; Oi et al., 2019; Post et al., 2015; Sakurai et al., 2024).

Building on this biomechanical foundation, several studies have quantified EVT under varying mechanical conditions to identify patterns associated with elevated UCL stress (Aguinaldo & Chambers, 2009; Dong et al., 2025; Escamilla et al., 2018). The overall evidence indicates that EVT magnitude varies substantially between pitchers, primarily due to differences in joint kinematics, kinetic chain dynamics, and individual physical characteristics. These differences mean that EVT cannot be explained solely by PV or throwing intensity but rather emerges from the combined influence of multiple interacting mechanical determinants. As many of these determinants are modifiable, EVT represents both a key causal pathway for UCL injury and a viable target for prevention strategies (Fleisig et al., 2025).

2.3.2. Velocity–Torque Relationship

High PV has been consistently identified as a risk factor for UCL injury, with four-seam fastballs (FF) implicated as the greatest source of medial elbow stress among common pitch types (Makhni et al., 2014; Prodromo et al., 2016). Against this backdrop, the parallel rise in PV and UCLR incidence across professional baseball has prompted growing concern among clinicians and researchers (Almonroeder et al., 2024; Erickson et al., 2016). PV has long been considered a key performance metric in baseball, but its role as a predictor of EVT and, by extension, UCL injury risk is more complex than a simple linear association.

While throwing at higher speeds can elevate joint loading, evidence suggests that the magnitude of EVT at a given PV is highly dependent on mechanical efficiency and individual kinematic patterns.

High PV has been consistently identified as a major risk factor for UCL injury, with strong evidence linking increased PV to a higher risk of elbow injury in professional baseball pitchers (Chalmers et al., 2016; McCutcheon et al., 2025; Prodromo et al., 2016; Whiteside et al., 2016). Peak PV may be particularly influential, with higher peak PV identified as the strongest independent predictor of UCL tears, suggesting that acute spikes in load may represent the primary biomechanical determinant of injury risk rather than cumulative repetitive microtrauma as measured by mean PV (Chalmers et al., 2016). It has been proposed that the effect of velocity on EVT may represent the underlying link between PV and injury (Anz et al., 2010). Research on the velocity–torque (V–T) relationship has shown that PV is associated with the development of EVT in both high school and professional pitchers (Hurd et al., 2012; Post et al., 2015; Slowik et al., 2019). The findings from these studies demonstrate that elevated PV is associated with increased injury risk through increased EVT. However, this relationship is confounded by inter-individual variability, with PV accounting for only a limited proportion of the between-subject variance in injury outcomes (Post et al., 2015; Slowik et al., 2019).

Injured pitchers have demonstrated similar absolute PV to healthy controls across all pitch types, indicating that PV is not a universal predictor of UCL injury (Keller et al., 2016). Moreover, Post et al. (2015) did not assess injury risk; however, their findings indicated little to no association between PV and the kinetic variables considered to contribute to elbow and shoulder injuries. The same pattern was observed in a large-scale cross-sectional study of 81 Division I collegiate pitchers, which examined both across- and within-pitcher relationships between PV and EVarT (Sakurai et al., 2024). This study revealed a weak association between EVT and PV, with half of the pitchers demonstrating no significant V-T relationship. This heterogeneity in individual pitcher V-T relationships explains why PV accounted for merely 5% of the variance in EVT between pitchers in this sample. This variability may also account for the non-significant V-T associations reported in previous across-pitcher analysis (Post et al., 2015; Slowik et al., 2019; Stodden et al., 2005). Collectively, these findings indicate that PV is an important yet insufficient predictor of medial elbow loading and, consequently, UCL injury risk. The critical determinant is not simply the magnitude of throwing PV, but rather the biomechanical mechanisms through which that PV is generated.

This concept that factors beyond simply high PV contribute to elbow kinetics contrasts with previously published findings. Fleisig et al. (1999) found that as pitchers progressed from partial- to full-effort throwing, kinetic variables including EVT increased concurrently with PV. However, kinematic variables (maximum shoulder external rotation [SER] and elbow-flexion angle at stride-foot contact [SFC]) were also observed to have changed simultaneously. SFC, also referred to as foot contact or lead foot contact, is a key temporal event in the pitching cycle defined as the instant the lead foot first contacts the ground, commonly identified either visually or when vertical ground reaction force exceeds 10 N (Dowling et al., 2023; Oyama et al., 2014; Tanaka et al., 2024). It is typically used as the reference point for normalising the pitching cycle from 0% (SFC) to 100% (ball release) and represents the initial moment of energy transfer from the ground to the

body, making it critical for subsequent kinetic chain sequencing (Aguinaldo & Chambers, 2009). These kinematic changes raise the question of whether elevated EVT results from increased PV, changes in pitching mechanics, or a combination of both factors. Previous studies have also demonstrated that increased PV in high school pitchers are closely and positively correlated with increases in EVT (Aguinaldo & Escamilla, 2019; Hurd et al., 2012; Luera et al., 2018). This difference in V-T relationship significance in younger cohorts may be due to lower biomechanical efficiency in high school pitchers, who often have less developed strength, coordination, and segmental sequencing (Luera et al., 2018; Nicholson et al., 2020). It has been shown that more mature pitchers, at higher competition levels generate greater PV and joint forces, while demonstrating greater mechanical efficiency, resulting in higher PV at comparable joint torques when normalised to bodyweight (Aguinaldo & Escamilla, 2019; Fleisig et al., 2009). These competition level–dependent differences in the V-T relationship may diminish the expected linear association between PV and upper extremity kinetics. More mature pitchers appear capable of generating forces in the proximal segments and transferring these forces more effectively through the kinetic chain, thereby reducing shoulder forces and elbow torques compared to younger pitchers (Aguinaldo & Escamilla, 2019).

Similar findings have been reported in other youth pitching studies, where kinetic chain coordination, lead leg stability, and trunk rotation (TR) timing are less consistent compared with collegiate or professional pitchers, resulting in greater reliance on the elbow and shoulder to produce PV (Luera et al., 2018; Nicholson et al., 2020; Oyama et al., 2013). This relative inefficiency means that for high school pitchers, PV functions as a more direct proxy for medial elbow loading, with fewer moderating effects from lower-body force transfer or trunk–arm segmental timing. As pitchers progress to higher competitive levels, improvements in mechanical efficiency can weaken the V-T relationship, allowing for greater PV generation without a proportional increase in torque. This contrast between samples emphasises the central role of biomechanical efficiency in mediating the relationship between PV and EVT. Understanding the mechanical and physical qualities that underpin efficiency provides a framework for identifying modifiable risk factors and developing targeted injury-prevention strategies.

In summary, while EVT represents the direct mechanical loading on the medial elbow that threatens UCL integrity, the V-T relationship demonstrates that PV alone inadequately explains this loading. Rather, biomechanical efficiency and kinematic patterns ultimately determine how PV translates into EVT and subsequent injury risk.

2.3.3. Biomechanical Efficiency

While both EVT and PV contribute to UCL loading, their relationship is moderated by the way PV is generated. Biomechanical efficiency in this context refers to the ratio of PV to normalised EVarT, expressed as a percentage of body weight in Newtons multiplied by height in metres (Crotin et al., 2022). It reflects the ability to maximise PV while minimising elbow loading, allowing meaningful comparisons between athletes

of different ages and body sizes. High biomechanical efficiency reflects optimal force transfer through the kinetic chain and reduced reliance on the elbow joint for PV production.

Pitchers demonstrating optimal kinetic chain sequencing have been shown to achieve significantly higher PV while experiencing no corresponding elevation in EVarT compared to those with less efficient movement patterns (Manzi, Dowling, Wang, et al., 2022). This biomechanical efficiency directly translates to reduced mechanical stress on the UCL, as EVarT represents a primary factor in injury risk (Crotin et al., 2022; McCutcheon et al., 2025). The relationship is particularly relevant given that conventional PV enhancement training may inadvertently increase EVT and associated orthopaedic risks despite improving performance outcomes (Bushnell et al., 2010; Slowik et al., 2019). Efficient mechanics therefore offer a modifiable pathway to reducing medial elbow stress without diminishing performance potential.

Empirical evidence indicates that numerous factors contribute to the efficiency of pitching mechanics and modulate the magnitude of forces at the elbow (Crotin et al., 2022; Davis et al., 2009; Howenstein et al., 2020; Werner et al., 2002). The pitching motion represents a complex kinetic sequence requiring substantial contributions from the lower extremities, pelvis, and trunk, where mechanical alterations throughout this chain can significantly influence EVT. Key determinants of this efficiency include the timing of TR, temporal coordination between pitching phases, and the magnitudes of SER, elbow flexion, and shoulder abduction at SFC (Aguinaldo & Escamilla, 2019; Dowling et al., 2024; Manzi, Dowling, Trauger, et al., 2022; Tanaka et al., 2024). These segmental contributions to biomechanical efficiency demonstrate that optimal pitching mechanics require coordinated integration across multiple body segments, with each component influencing the overall performance-to-stress ratio that determines both PV output and injury risk.

2.3.4. Segmental Contributions to Efficiency

2.3.4.1. Lower Body Mechanics

Force generation from the lower body plays a pivotal role in initiating and sustaining the kinetic chain in pitching. The contribution of the lower body to maximising PV through this sequential energy transfer is well established in baseball pitching research (Dowling et al., 2024; Guido & Werner, 2012; Kageyama et al., 2014). The lower body and trunk are primary sources of force generation in pitching, with lead leg ground reaction ground reaction force (GRF) positively associated with PV in adult, collegiate, and professional pitchers (Kageyama et al., 2014; Matsuo et al., 2001; McNally et al., 2015).

Lead knee extension, defined as the reduction in lead leg knee flexion from foot contact to ball release, represents a key mechanical component of this process (Dowling et al., 2024). This extension facilitates pelvic rotation and energy transfer through the kinetic chain, with both greater extension magnitude and extension velocity associated with increased PV. Dowling et al. (2020) reported that American college pitchers exhibited significantly greater lead knee extension (19°) and higher PV compared with Japanese college pitchers (2°), suggesting that increased lead knee extension enhances kinetic energy contribution and

ultimately PV. In a subsequent study, Dowling et al. (2024) found that greater lead knee extension was significantly associated with PV but not increased EVT, indicating that optimising lead leg mechanics may improve PV without elevating medial elbow stress or the risk of UCL injury.

The laboratory-based findings regarding lower body mechanics and their influence on elbow loading are validated by epidemiological evidence from professional baseball. Analysis of all 660 MLB pitchers during the 2019 season demonstrated that real-world injury patterns directly corresponded to the biomechanical principles identified in controlled studies (Beaudry et al., 2023). Pitchers utilising the "tall and fall" (TF) style, characterised by extended drive-leg knee positioning (140°-180°) at SFC, experienced significantly higher rates of upper extremity injuries (27.2%) and prior UCLR (32.8%) compared to those employing the "drop and drive" (DD) style with greater knee flexion (90°-130°), who sustained injury rates of only 15.3% and 22.6% respectively. The drive leg refers to the back leg that initiates the pitching motion by pushing off the rubber. This epidemiological pattern aligns with controlled biomechanical research showing that every 10° increase in drive leg knee flexion at foot contact produces measurable reductions in peak elbow valgus torque, translating the laboratory-observed protective effects of increased drive leg knee flexion into tangible injury prevention at the professional level (Solomito et al., 2022). Critically, these injury risk differences occurred without performance penalties, as no significant FF PV differences were observed between pitching styles, confirming that the biomechanical efficiency principles demonstrated in laboratory settings (optimising kinetic chain mechanics to maximise performance while minimising joint stress) are both achievable and consequential in elite competition. Thus, both lead leg and drive leg mechanics play distinct but complementary roles in optimising pitching performance while minimising injury risk.

In addition to lower-body kinematics, neuromuscular force development capacity represents a significant predictor of injury risk. Specifically, reduced eccentric rate of force development (ERFD) during countermovement jumps (CMJ) and imbalanced concentric CMJ profiles have been identified as significant predictors of elevated elbow injury risk in professional pitchers (Mayberry et al., 2020). The predictive model demonstrated that each one z-score increase in ERFD corresponded to approximately a 45% reduction in annual elbow injuries. The proposed underlying mechanism suggests that pitchers with deficient force development capabilities during the windup and stride phases may compensate by increasing force generation during the acceleration and follow-through phases, thereby placing excessive demands on the elbow joint (Pappas et al., 1985; Whiteley, 2007).

Building on these associations, concentric force–time characteristics derived from CMJ testing provide further insight into how imbalances in neuromuscular performance may influence elbow loading and injury risk. Regarding concentric performance, an interaction between average vertical concentric force (AVCF) and concentric vertical impulse (CVI) was identified, with imbalanced profiles demonstrating greater injury susceptibility (Mayberry et al., 2020). High AVCF/low CVI profiles reflect high force production over brief durations, which may generate sharp peak forces at the elbow increasing the risk of acute trauma, chronic soreness, or ligament tears (Fleisig et al., 1995). Conversely, low AVCF/high CVI profiles indicated low force production over prolonged durations, potentially overloading joint stabilisers and leading to cumulative

damage and injury (Mayberry et al., 2020). An earlier study found differing results with no significant CMJ kinetic differences between collegiate pitchers with high and low EVarT (Qiao et al., 2025). However, peak CMJ GRF, concentric impulse, and absolute peak power were moderately to strongly correlated with elbow varus torque rate of torque development (EVarTRTD), suggesting CMJ measures better reflect dynamic elbow loading rate than peak torque. Differences from Mayberry et al. (2020) may reflect population and methodological factors, with Qiao et al. (2025) 's smaller, more homogeneous collegiate cohort likely exhibited fewer performance imbalances and reduced cumulative throwing exposure compared to the professional pitcher sample.

Collectively, these findings emphasise that optimal lower body mechanics involve not only force magnitude but also the temporal characteristics and coordination of force production throughout the pitching motion, ultimately contributing to the biomechanical efficiency that maximises performance while minimising elbow stress.

2.3.4.2. Lower Body Mechanics Trunk Mechanics

Owing to its substantial segmental mass, the trunk is proposed to be the predominant contributor to total angular momentum during the pitching motion, with optimal timing of trunk rotation proving essential for maximising energy transfer whilst minimising the mechanical demands placed upon the throwing arm (Aguinaldo et al., 2007; Naito et al., 2014; Oyama et al., 2014). Consequently, alterations in trunk mechanics during the pitching sequence directly impact energy transfer efficiency and subsequently modify throwing arm kinetics and overall pitching performance (Oyama et al., 2014; Roach et al., 2013). Evaluation and optimisation of TR mechanics represent fundamental components of both performance enhancement and injury prevention strategies.

TR velocity has been demonstrated to correlate with elevated elbow and shoulder torques and enhanced PV (Lizzio et al., 2020; Tanaka et al., 2024). Furthermore, the temporal sequence of peak pelvic and TR velocities critically influences pitching mechanics (Putnam, 1993; Stodden et al., 2005). This segmental timing further determines the trunks contribution to performance and injury risk, as pitchers who initiated TR before SFC generated significantly greater EVT and shoulder internal rotation (SIR) torques and faced an increased risk of elbow and shoulder injuries (Aguinaldo & Chambers, 2009; Davis et al., 2009; Douoguih et al., 2015). During the pitching motion, optimal sequencing of peak pelvic and TR velocities maximises the efficiency of momentum transfer from the lower body through the pelvis, trunk, and upper body to the baseball, thereby increasing PV (Putnam, 1993; Stodden et al., 2005). Greater suppression of trunk rotation during peak pelvic rotation appears to augment PV through improved kinetic chain momentum transfer; however, this mechanical pattern concurrently elevates EVT (Tanaka et al., 2024). This finding indicates that whilst specific trunk mechanics may maximise performance outcomes, they may simultaneously intensify joint loading, highlighting the critical need to balance velocity enhancement with elbow health preservation strategies.

Lumbopelvic stability is another important determinant of torque efficiency. Throughout the pitching motion, core and gluteal musculature provides pelvic stabilisation and regulates trunk axial rotation, influencing PV and upper extremity moments including EVT (Oliver & Keeley, 2010). Excessive transverse plane trunk and pelvic motion during single-leg step-down (SLSD) tasks predicted elevated EVT despite reduced PV, indicating that transverse plane instability compromises energy transfer efficiency and increases medial elbow loading (DeZee et al., 2025). Analysis of SLSD kinematics revealed that concurrent transverse-plane pelvic and TR in the same direction reflect trunk kinematics being governed by proximal pelvic stability (Lewis et al., 2015). Contralateral trunk tilt (CTT), also referred to as trunk side tilt or lateral trunk lean, is a kinematic measure in baseball pitching that quantifies the angle of the trunk in the global y–z plane, representing tilt toward the glove side at ball release (Escamilla et al., 2018). Moderate CTT has similarly been associated with increased shoulder and elbow forces without corresponding velocity improvements, suggesting that modified trunk positioning may impose additional upper extremity stress without enhancing performance (Escamilla et al., 2023). Therefore, optimal lumbopelvic stability emerges as a critical component of biomechanical efficiency in pitching, as proximal control deficits compromise the kinetic chain's capacity for effective energy transfer, resulting in both diminished PV output and elevated medial elbow loading that increases UCL injury risk.

The complex interplay between trunk rotation velocity, timing, and stability highlights how this segment integrates lower body forces with upper extremity demands, making trunk mechanics a critical determinant of both pitching effectiveness and injury prevention.

2.3.4.3. Shoulder Mechanics

As the subsequent link in the kinetic chain, the shoulder functions as a key transfer point where generated momentum is channelled into the throwing arm, with its strength, kinematics, and ROM directly influencing elbow loading patterns and overall biomechanical efficiency. In particular, shoulder external rotation (SER) governs how effectively kinetic energy is transmitted, playing a central role in generating and directing forces that contribute to medial elbow loading. SER refers to the backward rotation of the upper arm segment about its long axis relative to the torso, where the humerus rotates away from the midline of the body (Diffendaffer et al., 2022). This movement is critical for optimal kinetic energy transfer during the pitching motion, particularly in the late cocking phase when the shoulder reaches MER (Post et al., 2015). Multiple studies have demonstrated that greater shoulder MER angle is significantly associated with higher EVT (Aguinaldo & Chambers, 2009; Aguinaldo & Escamilla, 2019; Anz et al., 2010; Crotin et al., 2022). This association between greater shoulder MER and increased EVT appears to be explained by mechanical interactions within the kinetic chain. Specifically, increased external rotation frequently occurs concurrently with elbow extension, which serves to lengthen the moment arm relative to the trunk's axis of rotation while simultaneously elevating segmental moments of inertia, thereby amplifying the torque demands placed upon the elbow joint (Aguinaldo et al., 2007; Albright et al., 1978; Matsuo et al., 2006). This mechanical disadvantage is further compounded when the throwing arm demonstrates temporal lag during the

acceleration phase, creating a bending moment that contributes additional valgus loading to the elbow (Aguinaldo & Escamilla, 2019). Contrarily, reduced external rotation paired with greater elbow flexion effectively shortens the lever arm relative to the trunk and has been associated with lower valgus torque production; however, insufficient external rotation, particularly when accompanied by inadequate shoulder abduction, can disrupt the optimal timing sequence of the kinetic chain and paradoxically increase tissue loading despite the theoretically improved mechanical positioning (Diffendaffer et al., 2022; Wilk et al., 2012).

However, the relationship between SER and EVT is further modulated by overall biomechanical efficiency and methodological considerations in torque measurement. Crotin et al. (2022) found that high-efficiency pitchers exhibited greater shoulder MER at SFC but lower normalised EVarT, suggesting that efficient trunk and shoulder mechanics can harness high SER without disproportionately increasing elbow stress. This finding is supported by McCutcheon et al. (2025) who reported that pitchers in the high-torque group exhibited a later arm position, characterised by reduced SER at SFC, compared to the low-torque group. The apparent inconsistency regarding whether increased shoulder MER correlates with elevated elbow joint loading may be attributed to methodological variations in torque quantification approaches. Previous investigations have analysed absolute EVarT values (expressed in N·m), whereas Crotin et al. (2022) employed unitless normalised EVarT expressed as a percentage of a pitcher's weight in newtons multiplied by height in metres. The potential combination of a heavier and taller athlete with greater shoulder MER could give rise to greater biomechanical efficiency, as the relativity of loading is reduced while PV is increased. This normalisation approach suggests that the combination of greater anthropometric dimensions and increased SER may contribute to enhanced biomechanical efficiency, thus optimising the performance-to-stress ratio in larger athletes.

Similarly, shoulder abduction demonstrates a complex, phase-dependent relationship with EVT that further illustrates the importance of kinematic timing and positioning. At SFC, greater shoulder abduction is associated with increased normalised EVarT, whereas pitchers with higher biomechanical efficiency typically demonstrate slightly less than 90 degrees of abduction, a position that may reduce torque and injury risk (Crotin et al., 2022). In contrast, at ball release, greater shoulder abduction is linked to lower normalised EVT, with values near 90 degrees appearing optimal (Fleisig et al., 2015). Importantly, EVT can rise sharply when abduction deviates markedly from 90 degrees, particularly when combined with trunk side tilt, highlighting the interactive role of shoulder abduction in modulating elbow stress and the critical importance of maintaining appropriate arm positioning throughout the throwing motion.

In summary, shoulder kinematics, particularly SER and abduction, are central to the regulation of EVT, with their impact shaped by both timing and coordination within the kinetic chain. While these movements establish the mechanical foundation for energy transfer and torque modulation, their influence cannot be fully understood in isolation. The trajectory of the throwing arm itself, encompassing how the shoulder integrates with the elbow, forearm, and wrist across the pitching sequence, further determines both the

magnitude and timing of loading. This highlights the need to consider arm path dynamics as an extension of shoulder mechanics in evaluating performance efficiency and injury risk.

2.3.5. Arm Path

The influence of shoulder kinematics on elbow loading extends beyond static positioning to encompass the dynamic trajectory of arm movement throughout the pitching sequence. While existing research has identified specific kinematic parameters associated with EVT from SFC onward, including elbow flexion, shoulder horizontal abduction, shoulder rotation, shoulder abduction, and forearm pronation, the contribution of arm path mechanics remains poorly understood. Arm path during the early phases of the pitching motion influences both the magnitude and timing of torque generation at the elbow. Arm path quantifies the cumulative distance traversed by the throwing arm during its trajectory from initial glove position to ball release, incorporating the integrated kinematic contributions of wrist, forearm, elbow, and shoulder across the complete pitching sequence (Dowling et al., 2023) (Figure 2.1).

Early arm path is defined as the total distance travelled by the hand marker during the pitching motion from maximum knee height to foot contact, a critical phase that contributes significantly to overall pitching mechanics (Dowling et al., 2023; Escamilla et al., 2017; Werner et al., 2002) (Figure 2.1). Dowling et al. (2023) found that a shorter early arm path has been associated with reduced EVT but also a modest decrease in PV, indicating a potential performance-injury trade-off that warrants consideration in mechanical optimisation strategies. Conversely, they observed that increases in early, late, and total arm paths were significantly and positively correlated with increased PV within individual pitchers, with every 30 cm increase in early arm path corresponding to a 0.354 m/s (0.79 mph) increase in PV and higher EVT. However, these strong correlations were evident only within individual pitchers; between-pitcher comparisons revealed weak or non-significant relationships, likely attributable to confounding factors including lower body mechanics, kinetic chain timing, and anthropometric variations.

The majority of biomechanical studies have examined the pitching sequence commencing at SFC, operating under the assumption that pre-contact movement phases exert minimal influence due to the ground-body energy transfer interface being established at the initial moment of foot-ground interaction (Escamilla et al., 2017; Fleisig et al., 1999; Manzi, Dowling, Trauger, et al., 2022; Werner et al., 2002). Dowling et al. (2023) conducted the first comprehensive examination of the early arm path phase. Contrary to prevailing assumptions, Dowling et al. (2023) identified considerable throwing arm kinematics during the early arm path interval, with their analysis revealing that 51% of total arm path displacement occurs within this previously overlooked phase, thereby questioning established paradigms regarding the relative importance of pre-stride foot contact movements. Combined with this and a shorter early arm path being associated with reduced EVT, it is arguably paramount that any analysis of the throwing arm includes motion of the early arm path, as this is over half of the total movement during the pitching motion. Given these findings and the association between shorter early arm path duration and reduced EVT, it is imperative that comprehensive

analyses of throwing arm mechanics incorporate the early arm path phase. Previous investigations have demonstrated that pitchers possess greater capacity for kinematic modifications during the initial phases of pitch delivery (i.e., early arm path) compared to subsequent phases where compensatory adjustments are considerably more constrained (Fleisig et al., 2018). Consequently, the strong correlation between arm path dynamics and EVT within individual performances highlights the strategic importance of targeting early-phase mechanical modifications to optimise loading patterns and reduce injury risk.

Arm path integrates with lower body, trunk, and shoulder mechanics to shape the timing and magnitude of elbow loading, making it a key determinant of biomechanical efficiency. Its early-phase influence offers a critical window for kinematic adjustment, reinforcing that efficiency depends on coordinated sequencing across the entire kinetic chain rather than isolated joint actions.

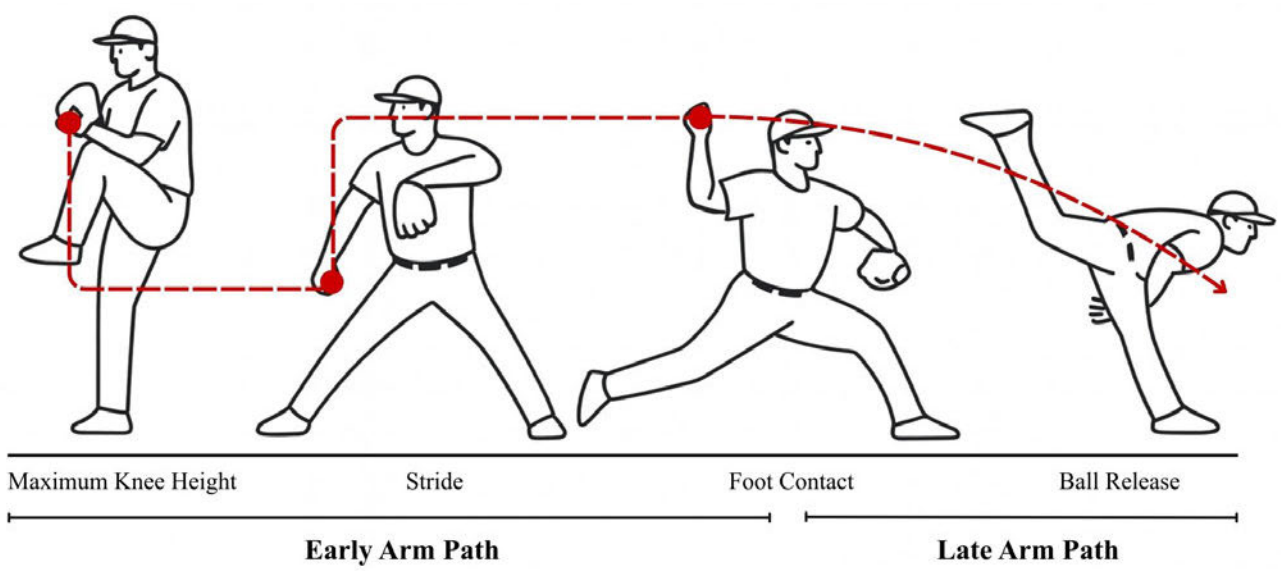

Figure 2.1 Phases of the pitching motion used to define arm path lengths

Note. Early arm path = maximal knee height to foot contact; late arm path = foot contact to ball release. The red dotted line represents total arm path, defined as the cumulative distance traversed by the throwing hand from initial glove position to ball release.

2.3.6. Arm Slot and Release Point

Building upon the established influence of arm path and shoulder kinematics, AS at ball release provides additional refinement to momentum transfer through the upper extremity, consequently modulating torque demands at the elbow joint. Although sometimes considered a stylistic feature, AS influences both trunk mechanics and joint loading patterns. AS describes the angle of the throwing arm relative to the ground at ball release, as seen from the catcher's perspective. It results from the combined effects of trunk lateral flexion, shoulder abduction, and elbow flexion (Camp et al., 2017). It is often classified into overhand (arm near vertical), three-quarter (halfway between), and sidearm (arm near horizontal) (Figure 2.2). Although EVT occurs around the point of maximum shoulder rotation, research has shown that the AS at ball release

plays a significant role in influencing valgus forces on the elbow (Aguinaldo & Chambers, 2009; Camp et al., 2017; Matsuo et al., 2006).

Current literature presents conflicting evidence regarding the relationship between AS and EVT. Several studies suggest that a lower AS results in increased EVT and potentially elevates injury risk. Aguinaldo et al. (2009) found that pitchers employing a more overhand delivery experienced significantly lower EVT compared to those with a sidearm motion, suggesting a biomechanical advantage in vertical arm positions. Camp et al. (2017) similarly demonstrated that a decrease in AS angle (closer to 0°, indicative of a sidearm delivery) was positively correlated with increased EVarT; specifically, a 13° reduction in AS angle led to a 1 N·m increase in EVarT. In contrast, other studies have reported opposing trends. Dong et al. (2025) observed that pitchers with a higher AS angle (~67°), representing a more sidearm motion by their definition, demonstrated significantly lower EVarT and improved torque efficiency compared to those with a lower angle (~43°). Furthermore, they estimated that for every 10° decrease in AS angle increased EVarmT by approximately 4.23 N·m. Escamilla et al. (2018) supported this trend, noting that sidearm pitchers exhibited lower EVT (88 N·m vs 97 N·m) and greater SER (169° vs 163°), although the difference in torque was not statistically significant.

These discrepancies can largely be attributed to inconsistencies in the definition, measurement methodology, and study populations across the literature. The term AS has been variably defined, with some studies referring to it as the vertical-horizontal position of the throwing extremity influenced by trunk tilt, shoulder abduction, and elbow flexion (Aguinaldo & Chambers, 2009) while others conceptualise it as an angular measure from a vertical reference vector extending from the shoulder joint centre to the pitching hand (Dong et al., 2025, p. 2018; Escamilla et al., 2018). Methodological differences further compound these inconsistencies. Some studies relied on indirect surrogates such as contralateral trunk tilt (CTT) or qualitative video assessments (Aguinaldo & Chambers, 2009; Matsuo et al., 2006; Oyama et al., 2013), whereas others employed direct kinematic measurements (Camp et al., 2017; Dong et al., 2025; Escamilla et al., 2018), yielding more accurate and reliable results. Differences in the populations studied also influence outcomes, as mechanics and joint loading vary across levels of play. For example, Dong et al. (2025) focused on elite college pitchers, while Camp et al. (2017) and Escamilla et al. (2018) examined professional athletes, and Aguinaldo et al. (2009) included a broader mix of adult players. Such variation in experience and physical development likely affects the generalisability of findings, as highlighted by Manzi et al. (2023), who observed level-specific correlations between AS angle and elbow flexion torque. Finally, discrepancies may stem from differing throw intensities across studies; Camp et al. (2017) collected data from professional pitchers (MLB and MiLB) across 82,000 throws of all types, including warm-up, long-toss, bullpen, and live game activity. This broad inclusion likely diluted observed mechanical associations and resulted in lower overall mean EVT values compared to studies that focused exclusively on maximal effort pitches. In contrast, Dong et al. (2025) analysed the fastest trial of maximal effort FF from elite college pitchers (velocity > 80 mph), while Escamilla et al. (2018) examined a minimum of five full-effort FF from a windup

in professional pitchers. These methodological and demographic inconsistencies may help explain the divergent conclusions regarding AS angle and elbow injury risk.

In addition to methodological differences, recent evidence suggests that the biomechanical effects of AS are closely linked to trunk mechanics. AS interacts with trunk lateral flexion and rotational timing, both of which influence joint loading. Contralateral trunk tilt (CTT) refers to the lateral inclination of the trunk toward the non-throwing side during baseball pitching mechanics (Escamilla et al., 2023). Escamilla et al. (2023) found that moderate CTT, common in three-quarter slots, was associated with increased shoulder and elbow loading, even though EVarT itself did not vary significantly across trunk tilt categories. Similarly, Dong et al. (2025) reported that pitchers with lower AS exhibited greater CTT and forward trunk flexion at release. These findings underscore the interdependence of AS and trunk posture. Notably, Aguinaldo et al. (2009) used trunk lean to define AS categories, making it a built-in feature of their classification. By contrast, studies using angular AS definitions (Dong et al., 2025; Escamilla et al., 2018) still observed consistent trunk differences between AS groups. Thus, AS should not be viewed as an isolated contributor to medial elbow torque, but rather as part of a coordinated kinetic sequence involving trunk mechanics. In summary, the relationship between AS, EVT and EVarT remains inconclusive due to inconsistent definitions, methodological variability, and the confounding influence of trunk mechanics. While AS may influence medial elbow loading, its role should be interpreted within the broader context of whole-body pitching mechanics and individual variation.

These observations collectively illustrate that AS functions within a complex network of interactions involving trunk posture, shoulder positioning, and arm path kinematics, highlighting the inadequacy of isolated parameter analysis in evaluating pitching performance and injury susceptibility. Consequently, AS emerges as a crucial modifiable variable within the biomechanical optimisation framework, where minor adjustments to arm angle at release can substantially alter the equilibrium between PV production capacity and medial elbow loading. This interdependent relationship among kinematic variables underscores the importance of holistic evaluation strategies that recognise the integrated nature of pitching mechanics rather than focusing on discrete biomechanical elements.

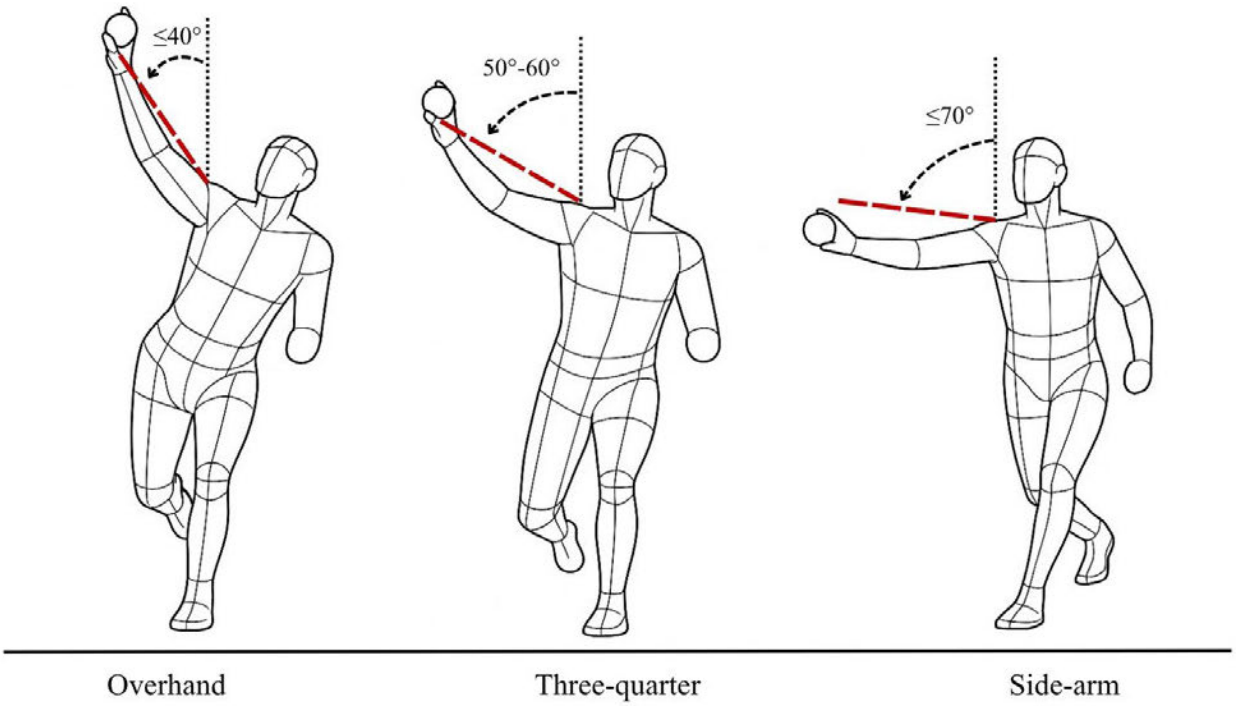

Figure 2.2 Representative examples of overhand (≤40°), three-quarter (50°–60°), and sidearm (≥70°) arm slot angles

Note. Arm slot angle is measured relative to vertical (dotted line) using a vector from the shoulder joint centre to the hand marker. Boundary angles for each delivery type are indicated by dash-dot lines.

2.3.7. Synthesis and Implications

The literature reviewed revealed that UCL injury risk emerges from a complex interplay of biomechanical, kinematic, and temporal factors rather than any single causal pathway. While EVT represents the direct mechanical threat to UCL integrity, its magnitude cannot be predicted reliably from conventional risk factors when considered in isolation. This complexity explains the inconsistent relationships observed between individual variables and injury outcomes across different populations and competitive levels.

The V-T relationship exemplifies this multifactorial reality. Although elevated PV correlates with increased UCL injury risk in some populations, particularly at youth levels, this association weakens considerably among mature, elite-level pitchers who demonstrate superior biomechanical efficiency (Aguinaldo & Escamilla, 2019; Hurd et al., 2012; Luera et al., 2018). Similarly, AS demonstrates conflicting associations with elbow loading across studies, largely due to its interdependence with trunk mechanics, measurement inconsistencies, and population differences (Dong et al., 2025; Escamilla et al., 2018, 2023). These findings collectively indicate that isolated kinematic parameters fail to capture the integrated nature of pitching mechanics.

The concept of biomechanical efficiency provides a unifying framework for understanding these apparently contradictory relationships. Efficient pitchers demonstrate the capacity to generate high PV through optimal kinetic chain sequencing while minimising compensatory loading at the elbow joint (Crotin et al., 2022;

McCutcheon et al., 2025). This efficiency manifests through coordinated contributions from multiple body segments: effective lower body force generation and lead knee extension, precisely timed TR that maximises momentum transfer, appropriate SER and abduction positioning, and streamlined arm path kinematics that reduce unnecessary joint stress (Aguinaldo & Escamilla, 2019; Crotin et al., 2022; Dowling et al., 2023, 2024; McCutcheon et al., 2025; Tanaka et al., 2024). Critically, these components must function as an integrated system rather than independent elements, as evidenced by the strong within-pitcher correlations between kinematic variables and loading patterns compared to weak between-pitcher associations.

2.3.8. Integration with Machine Learning Research

The biomechanical evidence demonstrating that UCL injury risk emerges from complex interactions between multiple kinematic and kinetic variables, rather than isolated factors, provides strong theoretical justification for the ML approach proposed in this thesis. Traditional epidemiological approaches to UCL injury prediction have focused on linear relationships between single variables (such as PV or workload) and injury outcomes. However, the biomechanical literature reveals why these approaches have yielded inconsistent results: the V-T relationship accounts for only 5% of EVarT variance between pitchers, AS effects vary dramatically based on trunk mechanics and measurement methodology, and individual kinematic parameters demonstrate strong within-pitcher correlations that disappear in between-pitcher analyses (Dong et al., 2025; Dowling et al., 2023; Escamilla et al., 2018; Sakurai et al., 2024). These findings suggest that injury risk arises from complex, nonlinear interactions between multiple biomechanical variables, which are the types of relationships that ML methods are particularly effective at detecting (Butani et al., 2025; Nicholson et al., 2022). ML algorithms, particularly advanced deep learning and ensemble methods, are highly effective in identifying intricate patterns and complex interactions that simpler linear models are unable to capture (Butani et al., 2025). Multiple studies have demonstrated that advanced ML models consistently outperform traditional statistical approaches such as logistic regression, Naïve Bayes, and Support Vector Machines in predicting injuries among MLB players (Butani et al., 2025; Karnuta et al., 2020; Nicholson et al., 2022; Oeding et al., 2024; Whiteside et al., 2016).

The concept of biomechanical efficiency as an integrated measure of performance-to-stress ratio provides theoretical validation for using anomaly detection to identify pre-injury changes (Crotin et al., 2022). If UCL injury risk reflects declining biomechanical efficiency rather than absolute changes in individual metrics, then injury-prone pitchers should exhibit systematic deviations from their typical integrated performance patterns. The autoencoder (AE) approach is specifically designed to detect such multivariate anomalies by learning the complex relationships between pitch characteristics and identifying instances where these relationships deviate from normal patterns (Badhoutiya et al., 2023). In this pitching context, "normal" data can be defined as a pitcher’s typical biomechanical measurements, physiological markers, and performance metrics when healthy and performing at an optimal level. Meaningful deviations from these established patterns, such as unexpected changes in PV, release point, spin axis, or increases in joint torque, may be identified as anomalies that signal elevated injury risk.

The biomechanical evidence indicating that pitchers have a greater capacity for kinematic modifications during the initial phases of pitch delivery suggests that compensatory adjustments may be detectable in pitch-tracking data before clinical symptoms become apparent (Fleisig et al., 2018). When biomechanical efficiency declines due to fatigue, subclinical tissue damage, or altered movement patterns, pitchers may unconsciously modify their pitch selection, release point, spin characteristics, or movement patterns to maintain performance while protecting the injured structure (Kang et al., 2025; Mayo et al., 2021; Whiteside et al., 2016). These compensatory adaptations, while often too subtle to be captured by traditional statistical analyses focused on isolated variables, represent the type of complex multivariate pattern changes that deep learning anomaly detection methods are well positioned to identify (Badhoutiya et al., 2023; Gong et al., 2019). Additionally, this study employed Shapley Additive Explanations (SHAP) as an explainable artificial intelligence (XAI) technique to address the black-box nature of deep learning models, enabling the quantification of each feature's contribution to predictions and thereby enhancing interpretability for medical professionals, supporting clinical validation, and guiding the development of personalised prevention strategies (Oeding et al., 2024; Weng et al., 2025).

The ML approach directly addresses several limitations identified in previous biomechanical research. First, the complex interdependencies between TR timing, SER, arm path dynamics, and elbow loading cannot be captured by conventional statistical methods that examine variables in isolation. Although Statcast data does not provide direct kinematic or kinetic joint-level measurements, critical biomechanical features can be inferred through pitch-tracking metrics such as PV, release point, and spin axis. The AE framework leverages these data by learning the underlying interdependencies and detecting when their normal relationships become disrupted, thereby offering a powerful method to approximate and monitor biomechanical efficiency in the absence of direct motion-capture data.

Second, the substantial between-pitcher variability in how kinematic changes translate to injury risk necessitates individualised analysis rather than population-based models. By training on individual pitcher data, the proposed approach accounts for these individual differences and detects personalised patterns of performance degradation.

Finally, biomechanical evidence indicating that efficiency relies on coordinated sequencing across the entire kinetic chain supports the analysis of multiple pitch characteristics simultaneously, with reconstruction error serving as an integrated measure of overall biomechanical integrity.

2.3.9. Implications for Performance and Prevention

The practical implications of this biomechanical efficiency paradigm extend beyond injury prevention to performance optimisation. Traditional approaches that target PV enhancement without considering mechanical efficiency may inadvertently increase EVarT and associated orthopaedic risks despite improving performance outcomes (Crotin et al., 2022; McCutcheon et al., 2025). Conversely, interventions that enhance kinetic chain coordination, including optimising lower body mechanics, improving lumbopelvic stability, and

refining temporal sequencing, provide effective pathways to reduce elbow loading while maintaining or improving PV production (Aguinaldo & Escamilla, 2019; Beaudry et al., 2023; DeZee et al., 2025; Dowling et al., 2024). The epidemiological evidence from professional baseball supports this framework, demonstrating that pitchers utilising more efficient movement patterns experience significantly lower injury rates without PV penalties (Beaudry et al., 2023; Crotin et al., 2022; Dowling et al., 2024). Given that each additional mile per hour of PV reduces runs allowed by approximately 0.3 per nine innings (Nicholson et al., 2022), training approaches that simultaneously optimise performance and prevent injury are critical for modern baseball. By establishing biomechanical efficiency as the foundation for player development, such methods sustain both competitive effectiveness and long-term health (Gray, 2010; Nicholson et al., 2022).

The integration of biomechanical theory with advanced analytics offers potential to identify previously unrecognised risk patterns. While laboratory studies have established that biomechanical efficiency integrates lower body mechanics, TR timing, shoulder kinematics, and arm path dynamics, translating these insights to field-based injury prevention has been challenging due to the difficulty of measuring complex kinematic patterns during actual competition. Statcast data provides an opportunity to detect the downstream effects of declining biomechanical efficiency through changes in the measurable outcomes of the pitching process: PV, movement, spin characteristics, and release consistency. If injury risk reflects integrated mechanical dysfunction rather than isolated parameter changes, then the multivariate anomaly patterns detected by ML should capture this integrated dysfunction in ways that individual variable analyses cannot.

Current evidence indicates that EVT is the principal biomechanical pathway to UCL injury (Buffi et al., 2015; Fleisig et al., 2025). Elbow injury risk is best explained by biomechanical efficiency, rather than by individual mechanical or workload variables. This conceptual shift from focusing on individual risk factors to optimising whole-body mechanical efficiency represents a fundamental advancement in both injury prevention strategies and performance enhancement approaches. The ML methodology proposed in this thesis represents a bridge between controlled biomechanical research and practical injury prevention applications, leveraging theoretical understanding of efficiency-based risk mechanisms to develop data-driven tools for identifying at-risk athletes in real-world competitive environments.

Recent pitch-tracking studies provide direct empirical support for this individualised approach. Mastroianni et al. (2026) demonstrated that injured pitchers exhibited greater intra-outing mechanical variability in FF velocity and horizontal release position relative to matched controls, alongside a cascade of sequential warning signs across the preinjury season, offseason, and final outings before injured list placement. Interestingly, these differences were evident in variability rather than mean values, underscoring that mechanical inconsistency, rather than performance decline alone, may be an early indicator of UCL injury risk. Complementing this, Dillon et al. (2025) found that acute UCL failure was characterised by short-term within-game mechanical volatility, with 86% of injured pitchers exceeding control-derived cumulative deviation thresholds in the five FF preceding injury. Collectively, these findings suggest that individualised baseline deviation, captured across multiple pitch-tracking dimensions simultaneously, may offer a more sensitive injury signal than population-level or univariate approaches.

2.4. Machine Learning in Baseball

2.4.1. Related Work on Machine Learning Approaches to Pitcher's Injuries

Most prior research has excelled in examining the underlying mechanisms of UCL injury but remains somewhat limited, as studies have predominantly focused on analysing and predicting pitcher elbow injuries using non-game-related data, which fails to capture the dynamic, real-time mechanics and stressors encountered during actual competition. However, recent research has increasingly utilised real-game data to predict pitcher injuries through ML and deep learning (DL) techniques (Karnuta et al., 2020; Oeding et al., 2024; Whiteside et al., 2016). These studies have demonstrated varying levels of predictive success, with accuracies ranging from moderate to high depending on the specificity of the injury being predicted and the modelling approach employed. Notably, pitch-tracking metrics derived from game data have consistently proven more predictive than traditional workload measures, suggesting that the biomechanical characteristics captured during actual pitching are critical indicators of injury risk. The specific methodologies, performance metrics, and key findings from these studies are summarised in Table 2.2.

Table 2.2 Overview of machine learning studies investigating pitcher injuries

Study	Models	n	Injury	Performance	Key Predictors/Findings
Kang et al. (2025)	ViT, LSTM, CNN-LSTM, TE, ResNet	620 players	UCLR surgery	F1-score: ViT = 0.73, ResNet = 0.64, TE = 0.46, CNN-LSTM = 0.44, LSTM = 0.35	Lowered release point, horizontal spin axis shift, and reduced PV, indicating increased UCL valgus stress
Karnuta et al. (2020)	RF, kNN, NB, XGB, T3E	3176 players	Next season DL placement risk	F1-score (Mean ± SD): T3E = 0.55 ± 0.02, RF = 0.54 ± 0.02, XGB = 0.54 ± 0.03, LR = 0.54 ± 0.04, kNN = 0.42 ± 0.02 NB = 0.38 ± 0.08	Ensemble ML models showed poor reliability for pitcher injury prediction (AUC=0.65, elbow AUC=0.61).

Oeding et al. (2024)	XGB	3808 pitcher-years	Shoulder/elbow injuries	Accuracy = 0.84 (95% CI, 0.83-0.85), AUC = 0.66 (95% CI, 0.60-0.71)	XGB found pitch-tracking metrics (PV, SL usage, spin rate) substantially more predictive than workload metrics.
Piergiovanni and Ryoo (2019)	3D-CNN (I3D)	5479 pitches	DL placement	Per-pitcher model F1-score = 0.75	3D CNN using optical flow achieved F1-scores of 0.74 (LHP) and 0.69 (RHP) for UCL tear prediction.
Whiteside et al. (2016)	NB, SVM	208 pitchers	UCLR surgery	NB performance: Accuracy = 72%, Sensitivity =70%, SVM performance: Accuracy = 75%, Sensitivity =74%,	SVM and NB models identified reduced recovery time, increased PV, higher pitch counts, and limited pitch repertoire as key predictors of UCLR.

Note. ViT = Vision Transformer; LSTM = Long Short-Term Memory; CNN-LSTM = Convolutional Neural Networks and Long Short-Term Memory; TE = Transformer-Encoder; ResNet = Deep Residual Learning Network; XGB = Extreme Gradient Boosting; LR = Logistic Regression; RF = Random Forest; kNN = k-Nearest Neighbours; NB = Naïve Bayes; T3E = Top 3 Ensemble; DL = disabled list; PV = pitch velocity; UCLR = ulnar collateral ligament reconstruction; UCL = ulnar collateral ligament; AUC = area under the curve, 3D-CNN = 3D Convolutional Neural Network; I3D = Inflated 3D Convolutional Neural Network, SVM = Support Vector Machine.

While these studies demonstrate the potential of ML and DL methods, their predictive accuracy has often been constrained by limited injury specificity, short observation windows, or a lack of clinical utility. Kang et al. (2025) advanced the field by leveraging pitch-by-pitch Statcast data within a real-time framework, enabling both early classification of UCL injury risk and regression-based estimates of days until surgery. Their models demonstrated robust performance (F1 = 0.73; R^2 = 0.79) whilst incorporating explainable artificial intelligence (i.e., SHAP values) to identify key biomechanical indicators, including lowered FF release point, increased horizontal spin axis, and reduced velocity. This real-time, interpretable approach utilising granular pitch data distinguishes Kang (2025)'s methodology from earlier research and renders it particularly relevant to the present study.

Kang et al. (2025) therefore provide a critical foundation for data-driven UCL injury prediction. However, important questions remain regarding how well population-level models capture the subtle, within-pitcher deviations that often precede injury. The biomechanical literature reveals substantial individual variability in

how mechanical patterns translate to injury risk, with surprisingly low correlations between PV and EVT. Specifically, half of the pitchers demonstrated no significant velocity-torque relationship, and PV accounted for merely 5% of the variance in EVT across pitchers (Post et al., 2015; Sakurai et al., 2024). More critically, strong correlations between arm path dynamics and EVT were evident only within individual pitchers; between-pitcher comparisons revealed weak or non-significant relationships, likely attributable to confounding factors including lower body mechanics, kinetic chain timing, and anthropometric variations. This pattern suggests that meaningful injury risk patterns exist at the individual level but are obscured in population-based analyses. Furthermore, pitchers possess greater capacity for kinematic modifications during the initial phases of pitch delivery compared to subsequent phases where compensatory adjustments are considerably more constrained, indicating that subtle compensatory adaptations may precede clinical symptoms (Fleisig et al., 2018). When biomechanical efficiency declines due to fatigue, subclinical tissue damage, or altered movement patterns, pitchers may unconsciously modify their pitch selection, release point, spin characteristics, or movement patterns to maintain performance while protecting the injured structure (Kang et al., 2025; Mayo et al., 2021; Whiteside et al., 2016). These within-pitcher variations in biomechanical efficiency and compensatory behaviour represent precisely the type of complex, individualised patterns that population-level models may fail to detect, highlighting the need for pitcher-specific analytical approaches that can identify personalised deviations from typical performance patterns.

2.4.2. Anomaly Detection in Machine Learning

Anomaly detection involves identifying data points or observations that exhibit substantial deviations from established baseline patterns or expected behaviours (Nawaz et al., 2024). For time series data, anomalies might represent sudden spikes, shifts, or subtle changes in temporal dynamics. In the broader healthcare context, it plays a critical role in the early detection of abnormalities that could indicate possible health problems (Badhoutiya et al., 2023). Healthcare settings benefit from anomaly detection algorithms that provide crucial capabilities for identifying irregular signals that may indicate developing pathological conditions, supporting proactive medical intervention and better patient outcomes (Badhoutiya et al., 2023). The complex and heterogeneous nature of healthcare data poses significant challenges for conventional analytical methods.

Autoencoders (AE), a subclass of artificial neural network architectures, offer a practical solution by leveraging neural networks to automatically learn meaningful data representations (Pereira & Silveira, 2019). While traditional dimensionality reduction techniques such as Principal Component Analysis (PCA) and linear methods assume linear relationships between variables, AE can capture the nonlinear interactions that characterise complex biomechanical systems (He et al., 2021). Kneifl et al. (2023) demonstrated AE superiority over traditional dimensionality reduction techniques in capturing musculoskeletal system nonlinearities. While further research in sport-specific biomechanics is warranted, current evidence supports AE as effective tools for modelling complex biomechanical system interactions. This distinction is critical in the context of UCL injury prediction, where the biomechanical literature demonstrates that injury risk

emerges from complex, nonlinear interactions between multiple kinematic and kinetic variables rather than linear combinations of individual factors (Butani et al., 2025). PCA identifies orthogonal linear combinations that maximise variance across the dataset, making it effective for data compression but fundamentally limited in its ability to model the intricate interdependencies between TR timing, SER, arm path dynamics, and elbow loading that govern injury risk. Moreover, the substantial between-pitcher variability documented in biomechanical research, where V–T relationships account for only 5 percent of variance between pitchers yet show strong within-pitcher correlations, suggests that meaningful injury patterns exist in nonlinear and pitcher-specific manifolds that linear dimensionality reduction cannot adequately represent (Post et al., 2015; Sakurai et al., 2024).

Unsupervised learning methods such as AE overcome the limitations of traditional threshold-based approaches by automatically extracting meaningful patterns from raw data. These unsupervised learning models function by encoding input data into a compressed latent space and then decoding it back to the original dimensionality. Unlike PCA's variance-maximising objective, AE are trained to minimise reconstruction error, forcing the network to learn the essential features and relationships necessary to accurately reproduce normal data (Alain & Bengio, 2014) In anomaly detection contexts, AE are trained exclusively on normal data, allowing them to develop an internal representation of typical patterns; anomalies are subsequently identified as data points that cannot be accurately reconstructed due to their unfamiliarity to the model (Fernando et al., 2021). This approach offers value in applications like pitching biomechanics, where anomaly identification typically requires extensive manual analysis by domain experts. Importantly, unsupervised models are well suited to this setting because labelled injury events are relatively scarce and often diagnosed retrospectively, making supervised training impractical. By establishing pitcher-specific baselines without relying on subjective labelling, they can accommodate the wide variability in mechanics and capture complex, multivariate patterns within high-dimensional Statcast data, enabling earlier and more individualised detection of meaningful deviations. The capacity of AE to model nonlinear manifolds and complex interdependencies makes them particularly suitable for detecting the integrated biomechanical dysfunction that precedes UCL injury, representing patterns that would be obscured or entirely missed by linear methods such as PCA.

2.4.3. Autoencoder Approaches for Anomaly Detection

AE have become a fundamental part of unsupervised and self-supervised learning in recent years. They are essential for tasks such as dimensionality reduction, denoising, anomaly detection, and generative modelling (Pinaya et al., 2020). Initially conceived as straightforward neural network architectures that learn to reconstruct input data through compressed latent representations, AE have since evolved into a diverse family of models capable of addressing various data types, constraints, and learning objectives (Mienye & Swart, 2025). The primary function of AE is to compress data into a lower-dimensional latent space and then use this representation to recreate the original input (decoding) (Pinaya et al., 2020). As illustrated in Figure

2.3, the encoder function $f\theta\ (x) = z$ transfers an input x to a latent code z, while the decoder $g\emptyset\ (z) = \hat{x}$ attempts to reconstruct the input. Depending on the type of input data, the training goal is to minimise the reconstruction loss $L(x, \hat{x})$, which is usually assessed using mean squared error (MSE) and cross-entropy loss (Mienye & Swart, 2025). AE are useful tools for feature learning because the reconstruction target compels the model to capture the most prominent aspects of the data. To overcome limitations including poor generalisation, a lack of structure in latent spaces, and susceptibility to noise, several AE variations have been developed over time. With distinct designs and applications, these variations include sparse, denoising, variational, contractive, adversarial, convolutional, and sequence-to-sequence AE.

Among AE variants, vanilla AE (feedforward) architectures potentially offer distinct advantages for detecting gradual deterioration by maintaining fixed baselines rather than adapting to progressive changes. Long Short-Term Memory (LSTM) AE are a recurrent neural network architecture designed to retain information across sequential time steps, making them theoretically better suited to capturing temporal dependencies in longitudinal data (Hochreiter & Schmidhuber, 1997). However, recent research found that simpler feed-forward architectures consistently outperformed them in practice. Specifically, the standard AE statistically outperformed LSTM-AE with 90% confidence (Wilcoxon's test), achieving superior results across key metrics: a higher average F1-score (0.27 vs. 0.25) and better average ranking (2.75 vs. 3.46) using optimal threshold selection and a window size of 100 (Markovic et al., 2024). This superiority held across nearly all experimental configurations and threshold methods, with LSTM-AE only performing slightly better at the largest window size tested (200), suggesting that the feed-forward AE's non-recurrent structure maintains more stable baselines for measuring deviations. It has been proposed that LSTM-based variants may inadvertently adapt to slow, progressive changes, incorporating deterioration into their learned notion of 'normal' and reducing reconstruction error for genuinely anomalous behaviour (Markovic et al., 2024). This fixed baseline advantage is reinforced by work on turbofan engine degradation, where a standard AE trained only on normal operating conditions (RUL > 130 cycles) achieved 93.8% F1 score by measuring deviations from an explicitly non-adaptive baseline of healthy operation (Jakubowski et al., 2022). However, both studies indicate context matters. Sequence-aware models became competitive at larger window sizes, while success in degradation detection depended on carefully isolating stable operational periods for training, suggesting optimal architecture choice depends on the specific degradation patterns and detection requirements (Jakubowski et al., 2022; Markovic et al., 2024).

For the specific application of injury detection in MLB pitchers, the vanilla AE's independence property offers particularly relevant advantages. Unlike recurrent architectures such as LSTM-AE, which model temporal dependencies by maintaining hidden states across sequential inputs, vanilla AE treat each data point as independent (Sak et al., 2014). In this context, each game's performance metrics are encoded, compressed into a lower-dimensional representation, and reconstructed without reference to previous games' patterns. The reconstruction error, which is calculated as the mean squared difference between input and output, serves as the anomaly score for that specific observation, allowing the model to detect deviations from a pitcher's established baseline without adapting to gradual performance decline that may signal emerging

injury. For injury detection specifically, the critical distinction between vanilla AE and temporal architectures lies in their handling of gradual changes. When trained exclusively on a fixed baseline period representing healthy performance (e.g., the first 60% of a season or pre-season games), vanilla AE maintain this static reference throughout the monitoring period. Because the model never updates or adapts its learned representation of "normal" performance, it can detect gradual deterioration that adaptive models might normalise. This characteristic is particularly relevant given Mayo et al.'s (2021) findings that UCL reconstruction is preceded by small but statistically significant game-to-game changes in pitch metrics over the 15 games leading up to surgery. Specifically, they identified monotonic trends including decreased FF velocity (approximately 0.1 mph per game decline starting around 9 games before injury), decreased SL velocity (0.5 mph total drop over 8 games), declining FF spin rate (around 5 rpm per game), and increased curveball (CU) usage. These patterns are theorised to reflect subclinical UCL injury and compensatory adjustments. A vanilla AE trained on a pitcher's healthy baseline could potentially flag these subtle deviations because it compares each game to the original healthy standard rather than a continuously updating baseline. For instance, even a cumulative velocity decrease of 0.7 mph over several games would trigger elevated reconstruction error as each observation deviates from the fixed healthy baseline. In contrast, temporal models that learn sequential patterns may adapt to this gradual decline, incorporating the deteriorating trajectory into their learned representation of "normal" and thereby failing to flag the progression as anomalous until more severe changes occur.

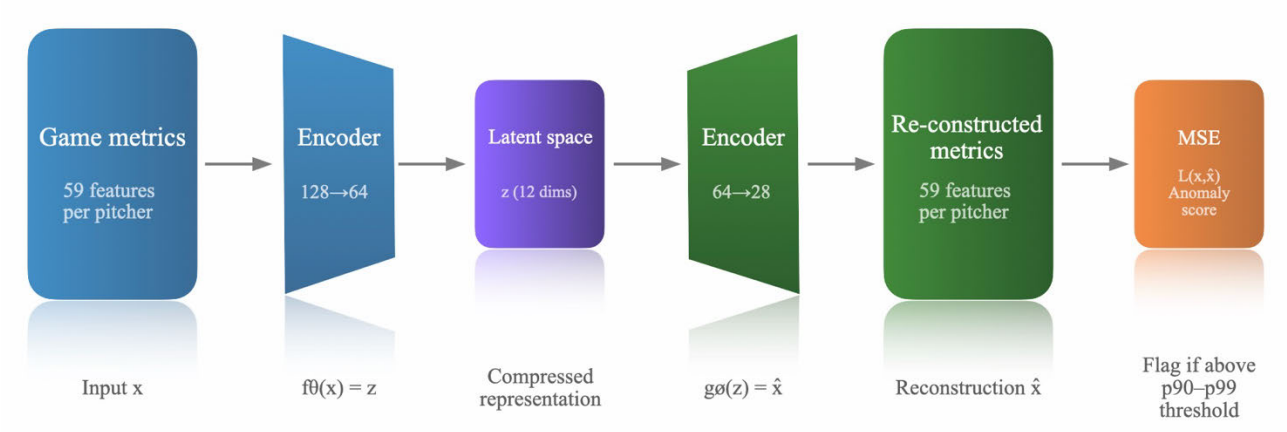

Figure 2.3 Vanilla autoencoder architecture for pitcher anomaly detection

Note. The encoder compresses per-game pitching metrics (x) into a 12-dimensional latent representation (z) via two fully connected layers (128→64 neurons), and the decoder reconstructs the original input (x̂) via a mirrored architecture (64→128 neurons). Reconstruction error, calculated as mean squared error between x and x̂, serves as the anomaly score. Games exceeding a pitcher-specific percentile threshold derived from the baseline reconstruction error distribution (p90–p99) are flagged as anomalous.

2.4.4. Implementation Considerations

The effectiveness of vanilla AE for pitcher injury monitoring depends critically on several methodological design choices that influence both detection sensitivity and practical applicability. First, the baseline period

definition requires careful consideration to ensure training data represents genuinely healthy performance. The temporal window selected for establishing each pitcher's baseline, whether spring training, early-season games, or a predetermined number of initial appearances, must balance the need for sufficient data to learn robust patterns against the risk of including subclinical injury states. This challenge is particularly acute in professional baseball, where pitchers may enter seasons with varying degrees of residual fatigue or minor tissue damage that does not preclude competition but may already represent subtle deviations from optimal biomechanical function (Escamilla et al., 2007; Grantham et al., 2014; Stone & Schilling, 2020)

Second, encoding dimensionality presents a fundamental trade-off between pattern extraction and overfitting (Mounayer et al., 2025). The latent space must be sufficiently compressed to force the network to learn meaningful, generalisable representations rather than simply memorising individual games (Pham et al., 2022). However, excessive compression may eliminate the network's capacity to represent legitimate inter-game variability inherent in healthy pitching performance (Lapenda et al., 2020). Such variations are attributable to factors such as opponent quality, weather conditions, or natural day-to-day fluctuations in neuromuscular coordination. Determining the optimal compression ratio requires systematic evaluation across multiple pitchers with diverse mechanical profiles, as the ideal latent dimensionality may itself vary according to individual complexity (Way et al., 2020).

Third, threshold calibration determines the balance between sensitivity (detecting true deterioration) and specificity (avoiding false alarms) (Ko et al., 2022). Anomaly thresholds are typically established using the distribution of reconstruction errors observed on the healthy baseline data itself, with common approaches setting the detection boundary at the 95th or 99th percentile of this distribution (Ndubuaku et al., 2019). This statistical approach provides a principled framework for defining "normal" variation, yet the choice of percentile reflects an implicit value judgment regarding the acceptable trade-off between missed injuries and unnecessary interventions. In practical deployment, threshold selection would require input from medical staff and coaching personnel regarding their tolerance for false positives relative to the consequences of failing to detect developing injuries.

2.4.5. Comparative Framework and Research Rationale

The primary theoretical advantage of vanilla AE over univariate statistical methods lies in their capacity for multivariate pattern recognition (Chen et al., 2016). While Mayo et al. (2021) demonstrated that trends in individual pitch metrics (velocity, spin rate) provide injury signals, AE architectures may detect injury signatures that emerge from specific combination of changes across multiple variables. For instance, injury risk may be elevated specifically when decreased velocity occurs concurrently with altered spin axis and shifted release point. This kind of multivariate signature, where no single variable exceeds alarm thresholds but their joint configuration indicates biomechanical compromise, is precisely what AE holistic encoding of all metrics simultaneously may detect earlier or more reliably than monitoring variables independently.

This hypothesis aligns with biomechanical evidence that injury risk emerges from complex interactions between kinematic and kinetic factors rather than univariate thresholds (Butani et al., 2025). However, the practical significance of this theoretical advantage remains an empirical question. If the univariate signals identified by Mayo et al. (2021) capture the dominant injury patterns, then AE additional complexity may provide minimal incremental value. Alternatively, multivariate detection may prove advantageous specifically for a subset of pitchers whose injury signatures manifest through subtle coordinated changes that escape univariate monitoring. The present study's comparative framework addresses this question directly by evaluating both approaches on identical data.

2.5. Conclusion

The evidence reviewed in this chapter consistently demonstrates that UCL injury risk in elite adult baseball pitchers emerges from complex, multivariate interactions across the kinetic chain rather than from any single mechanical or workload variable. The review revealed that UCL injury risk cannot be attributed to any single causal variable. EVT represents the primary mechanical pathway to ligament compromise, yet its magnitude is shaped by the integrated contributions of lower body force generation, lumbopelvic stability, TR timing, shoulder kinematics, arm path dynamics, and arm slot positioning (Aguinaldo & Chambers, 2009; Crotin et al., 2022; Dowling et al., 2023; McCutcheon et al., 2025). The concept of biomechanical efficiency, reflecting the ratio of PV to normalised EVT, provides a unifying framework for understanding these interactions. Pitchers who transfer force effectively through the kinetic chain can generate high PV without proportional increases in elbow loading, whereas those with inefficient mechanics may sustain elevated joint stress even at moderate velocities (Crotin et al., 2022). Individual differences in mechanical efficiency confound population-level comparisons, which accounts for the inconsistent associations between PV and AS and injury risk across studies, and explains why the velocity-torque and arm path relationships are consistently stronger within individual pitchers than between them (Dowling et al., 2023; Sakurai et al., 2024; Slowik et al., 2019). Collectively, these findings indicate that meaningful injury risk patterns are inherently individualised, underscoring the need for analytical approaches that can detect personalised deviations from established performance baselines rather than relying on population-level thresholds.

The review of ML-based approaches to pitcher injury prediction reinforces the same conclusion from an analytical perspective. Population-level supervised models have yielded moderate and variable predictive performance, in part because meaningful injury risk patterns are obscured when individual biomechanical profiles are aggregated (Karnuta et al., 2020; Oeding et al., 2024). Pre-injury changes in PV, release point, and spin characteristics identified at the within-pitcher level likely reflect compensatory adaptations to subclinical tissue damage rather than absolute performance thresholds (Cohen et al., 2022; Kang et al., 2025; Mayo et al., 2021; Portney et al., 2019). Their multivariate and individually variable nature provides direct theoretical justification for a pitcher-specific anomaly detection approach over both univariate monitoring and population-based supervised classification.

Taken together, the biomechanical and ML literature converges on the need for an individualised, multivariate analytical framework capable of detecting subtle departures from each pitcher's established performance profile. Chapter 3 describes the design and implementation of such a framework, applying pitcher-specific vanilla autoencoders to game-level Statcast data to characterise the temporal pattern of biomechanical deterioration in MLB pitchers in the period preceding UCLR.

Chapter 3. Machine Learning-based Detection of Pitching Pattern Changes Prior to Ulnar Collateral Ligament Reconstruction

3.1. Introduction

3.1.1. Background

The prevalence of overuse conditions and throwing-related injuries remains problematic within professional baseball, with incidence rates continuing to escalate at an unprecedented rate (Conte et al., 2016). Ulnar collateral ligament reconstruction (UCLR) is a significant surgical intervention that affects approximately one-quarter of surveyed Major League Baseball (MLB) pitchers throughout their professional careers (Leland et al., 2019). Recent research has demonstrated a consistent rise in UCLR surgeries in professional baseball over the past 50 years, with a significant surge between 2010 and 2015 (Almonroeder et al., 2024). Among professional pitchers, UCL injuries occur at a mean age of 24.4 years, coinciding with early career stages and potentially influencing long-term trajectories (Camp et al., 2018; Cinque et al., 2022). These injuries frequently require surgical intervention, with 25% of major league pitchers and 15% of minor league pitchers having undergone UCL reconstruction (Conte et al., 2015).

The UCL of the elbow serves as the primary soft tissue structure that provides stability against valgus stress during overhead throwing activities (Erickson et al., 2015). The anterior bundle of the UCL is particularly susceptible to injury in the dominant arm of pitchers, as a result of the high-velocity forces and substantial valgus torque generated during overhead throwing movements (Chen et al., 2001). These injuries typically present as chronic and degenerative conditions; however, acute ruptures and acute-on-chronic presentations are also observed in clinical practice (Erickson et al., 2015). Irrespective of their aetiology, UCL tears consistently result in medial elbow pain, valgus instability, and diminished velocity and accuracy during overhead throwing activities (Erickson et al., 2015; Jensen et al., 2020). Although UCL injuries were historically considered career-terminating for overhead athletes, the introduction of UCLR by Dr Frank Jobe in 1974 has enabled numerous overhead throwers with UCL tears to achieve successful return to elite levels of competition (Camp et al., 2018; Erickson et al., 2015; Jensen et al., 2020). However, UCLR represents a significant, long-term injury requiring an extensive rehabilitation period of approximately 16.8 to 20.5 months, resulting in considerable economic implications (Erickson et al., 2014). Research examining 194 MLB pitchers who underwent UCLR demonstrated an average financial cost of $1.9 million per athlete (Meldau et al., 2020). Given the substantial personal and economic consequences, there is a clear need to address the rising incidence of UCL injuries and prioritise the preservation of elbow joint integrity in pitchers across all levels of competition.

UCL injuries in pitchers emerge from complex interactions between biomechanical factors including elbow valgus torque (EVT), kinematic sequencing, and kinetic chain efficiency (Aguinaldo & Chambers, 2009;

Crotin et al., 2022). Research has established that biomechanical efficiency varies substantially between pitchers and may decline prior to injury (Mayo et al., 2021; Post et al., 2015). Biomechanical efficiency refers to the ability to generate velocity while minimising elbow torque, often expressed as the ratio of peak velocity to normalised EVarT relative to body size (Crotin et al., 2022). High-efficiency pitchers demonstrate superior segmental sequencing through the integration of lower body drive, trunk rotation timing, and shoulder mechanics, which enables them to achieve greater velocities without proportionally elevating medial elbow stress (Aguinaldo & Escamilla, 2019; Crotin et al., 2022; Oyama et al., 2014; Tanaka et al., 2024). However, direct biomechanical assessment remains impractical for routine injury surveillance in professional settings. The relevance of biomechanical efficiency to the present study lies in its capacity to unify disparate risk factors into a performance-to-stress ratio that can be approximated using in-game pitch-tracking data. This study leveraged publicly available Statcast pitch-tracking data, which captures the downstream effects of biomechanical efficiency through metrics such as velocity, spin rate, release point, and movement. By applying machine learning (ML) anomaly detection to these readily accessible performance metrics, this research aims to identify pitcher-specific deviations from established baselines that may signal elevated UCL injury risk before clinical symptoms arise. Recent ML studies applied to pitcher injury prediction have demonstrated the viability of this approach, with pitch-tracking metrics consistently proving more predictive of shoulder and elbow injury than traditional workload or demographic variables (Kang et al., 2025; Oeding et al., 2024). Notably, Kang et al. (2025) demonstrated that deep learning models applied to in-game Statcast data could identify injury risk up to 100 days prior to surgery, providing direct precedent for a data-driven, performance-based detection framework.

3.1.2. Purpose

The purpose of this study was to evaluate whether ML-based anomaly detection techniques can identify patterned changes in MLB pitchers' performance metrics during the games leading up to UCLR surgery, and to compare this approach against simpler statistical methods. Using autoencoder (AE) architecture trained on pitcher-specific baseline data (defined as periods without injury or surgery within 12 months), the model analyses multivariate pitching metrics, including velocity, movement, release point, and spin rate, with reconstruction error serving as the anomaly detection metric. This approach aimed to bridge the gap between population-level injury prediction models and the individualised, biomechanically-driven performance degradation that research suggests precedes UCL injury in professional pitchers.

3.2. Data

3.2.1. Data Source

Pitch-tracking data for this study was sourced from Statcast via the Baseball Savant application programming interface (API) using the baseballr package in R version 4.4.2 (Petti & Gilani, 2024; R Core

Team, 2024). Statcast is a proprietary player tracking and ball tracking system developed and operated by MLB Advanced Media for MLB. The baseballr package provides programmatic access to Statcast data through the Baseball Savant API, enabling efficient queries by date range, pitch type, and player attributes. Statcast uses a combination of high-speed cameras and radar systems to capture and report ball-tracking metrics. The system, introduced by MLB in collaboration with Washington State University, has demonstrated a spatial resolution of 0.03 inches per pixel, with exposure times of 50 μs and motion blur measuring less than 0.080 inches (Platt et al., 2021). The sample of players was identified using a publicly available database, which has tracked UCLR surgeries across both MLB and MiLB players since 2012 (Roegele, 2012). Pitchers with previous UCLR were excluded to ensure only primary UCLR cases were analysed. Injury occurrences were cross-verified using the transaction database available on the official MLB website (Major League Baseball, 2024). Both databases are open-access and have been used in previous research (Cinque et al., 2022; Kang et al., 2025; Portney et al., 2019).

3.2.2. Temporal Window Definition

All data were anchored to a reference date corresponding to each pitcher's last game before UCLR surgery, with the time axis expressed as days before last appearance. This standardisation enabled consistent temporal alignment across pitchers despite differences in individual game schedules and provided a uniform framework for modelling injury progression. Two distinct temporal windows were defined: a baseline training window and a detection window.

The baseline training window spanned 200 to 600 days before last appearance, providing a 400-day period for learning each pitcher's characteristic performance profile. This window was selected to ensure that training data represented pitching performance well before the documented period of pre-surgical deterioration, minimising the likelihood that injury-related biomechanical changes contaminated the baseline. However, it is acknowledged that pitchers in the baseline window cannot be confirmed as entirely injury-free, as subclinical UCL degeneration may be present well in advance of any detectable performance change or clinical presentation, given that UCL injuries typically develop as chronic and degenerative conditions rather than discrete acute events (Erickson et al., 2016). The evidence indicates that measurable performance deterioration occurs across multiple time scales, with season-long declines in performance metrics in the year before surgery, alongside a more pronounced deterioration approximately 50 to 60 days or around 15 games before UCLR, suggesting that biomechanical changes become detectable during this late pre-surgery period (Erickson et al., 2014; Keller et al., 2014; Mayo et al., 2021). By concluding the baseline window at 200 days pre-last appearance, the model training period remained well separated from this documented deterioration phase. The 400-day baseline span was informed by Kang et al. (2025), who demonstrated that extended historical data (up to 550 days pre-injury) enhanced predictive stability and accuracy for UCLR prediction.

The detection window encompassed the final 200 days before last appearance, during which the trained AE model was applied to identify anomalous performance patterns. This 200-day span was justified on multiple grounds. First, it captured the documented deterioration period identified by Mayo et al. (2021) at 50–60 days pre-last appearance while also providing temporal coverage for earlier, more subtle biomechanical changes that may precede overt performance decline. Second, successful injury risk prediction has been demonstrated up to 100 days prior to the last game before UCLR surgery, suggesting that detectable performance patterns exist within this timeframe (Kang et al., 2025). Moreover, in the context of partial UCL injuries, nonoperative rehabilitation protocols typically span approximately 12 weeks, suggesting that early detection of mechanical deterioration at least 200 days in advance could provide a clinically actionable window (Kato et al., 2019; Podesta et al., 2013). Identifying potential breakdown during this period enables coaches, medical staff, and performance practitioners to adjust training loads, monitor recovery trajectories, or initiate targeted interventions, with the goal of mitigating injury progression and potentially avoiding the need for surgical reconstruction.

3.2.3. Preliminary Analysis of Minimum Data Requirements

To establish minimum data requirements for both baseline training and detection windows, preliminary analysis was conducted on 34 pitchers from the 2016–2024 UCLR cohort meeting minimum data availability criteria (≥25 baseline games and ≥7 detection games). AE models were trained across varying sample sizes (baseline: 5–25 games; detection: 3–12 games) with 10 iterations per configuration to assess model stability, yielding 3,400 total model trainings. For the baseline window, validation loss decreased monotonically from 1.120 mean squared error (MSE) at 5 games to 0.779 MSE at 25 games, with the most substantial improvement occurring between 5 and 20 games (20% reduction). The overfitting gap (validation loss minus training loss) similarly decreased from 0.629 to 0.455, indicating improved model generalisation with larger training samples. Based on these findings, a minimum baseline threshold of 20 games was established, representing the point where validation loss stabilisation and acceptable overfitting characteristics converged. Although the coefficient of variation (CV) remained elevated (66.3%), this reflected genuine inter-pitcher variability rather than model instability, as mean validation loss continued to improve. Reliability theory demonstrates that variance-based metrics depend on the ratio of between-subject to total variance, such that increased heterogeneity can elevate dispersion without implying reduced stability (Bland & Altman, 1990; Quan & Shih, 1996). Given the concurrent reduction in overfitting gap and improved bootstrap stability, the elevated dispersion most plausibly represents genuine inter-individual heterogeneity.

Using the same sample of pitchers, a minimum detection window of 5 games was established, as the reconstruction error stabilised at this point. This threshold provided superior temporal coverage compared with three games (20 vs 10 days), aligned with the timeframe reported by Mayo et al. (2021), and yielded an almost twofold improvement in bootstrap stability (10.94 vs 21.73 MSE). Although 7 games showed marginally improved stability (14% improvement), this modest gain did not justify the resulting reduction in available pitchers, which would compromise statistical power and limit the generalisability of findings

to the broader population of MLB pitchers undergoing UCLR. Additionally, reconstruction error stabilised empirically at 5 games rather than continuing to improve at higher thresholds, suggesting this threshold represents a meaningful inflection point in detection reliability. However, persistently high CV (>250%) indicated substantial individual differences in pre-injury performance detectability. This heterogeneity reflects genuine biological variability in injury presentation rather than methodological instability, as bootstrap analysis confirmed measurement consistency at the selected threshold.

3.3. Methods

3.3.1. Inclusion and Exclusion Criteria

Pitchers identified through the UCLR databases were included in the final analytical sample if they met data availability requirements in both temporal windows. Based on preliminary analysis (Section 3.2.3), pitchers were required to have ≥20 games within the baseline training window (200–600 days before last appearance) and ≥5 games within the detection window (0–200 days before last appearance). These thresholds ensured sufficient data for stable AE training whilst maximising sample retention. Pitchers with prior UCLR surgery were excluded to focus analysis exclusively on primary reconstruction cases. Following application of these criteria, the final analytical sample comprised 46 MLB pitchers who underwent primary UCLR between 2016 and 2024 (see Appendix D for full cohort details). The descriptive statistics for the pitchers are summarised in Table 3.1.

Table 3.1 Sample characteristics by handedness (n = 46)

Handedness	Sample (n)	Height (cm)	Weight (kg)	Age (years)	Fastball velocity (kph)	Career appearances (n)
Right	33	188.7 ± 4.8	99.9 ± 9.8	28.7 ± 3.0	151.8 ± 4.1	200.9 ± 172.7
Left	13	189.5 ± 4.0	94.8 ± 9.5	30.3 ± 3.1	148.9 ± 4.8	265.5 ± 183.9

Note. Data present mean ± SD physical, performance, and experience characteristics by throwing hand. Career appearances reflect total MLB pitching appearances recorded through March 2026.

3.3.2. Variables and Measurements

Table 3.2 presents the pitch metrics used in the analysis. The pitching metrics selected for this study were informed by prior research on pitch-tracking variables associated with UCL injury risk (Kang et al., 2025; Mastroianni et al., 2026; Mayo et al., 2021; Whiteside et al., 2016; see also Sections 2.3.6 and 2.4.1). Release speed was included given its consistent identification as a predictor of elevated elbow torque and UCLR risk (Chalmers et al., 2016; Mastroianni, Kunes, Mueller, et al., 2025; Oeding et al., 2024; Prodromo et al., 2016; Whiteside et al., 2016). Horizontal and vertical release position were selected based on evidence that

progressive lateral and vertical shifts in release point precede UCL injury (Cohen et al., 2022; Kang et al., 2025; Portney et al., 2019). Release extension was retained as it captures forward displacement at release, relating to arm path mechanics associated with injury risk (Mastroianni et al., 2025). Spin rate and spin axis were included following findings that spin rate declines and spin axis shifts emerged as significant pre-injury indicators (Kang et al., 2025; Mayo et al., 2021; Oeding et al., 2024). Horizontal and vertical pitch movement (pfx_x, pfx_z) reflect the combined effects of spin and release characteristics on ball trajectory, providing an integrated measure of pitch quality used in prior injury prediction research (Kang et al., 2025; Oeding et al., 2024).

In addition to pitch-specific metrics, workload and fatigue variables exhibit a complex, interconnected relationship with injury risk, as fatigue demonstrates co-dependence between kinematic changes and performance decline (Birfer et al., 2019; Chalmers et al., 2021; Coughlin et al., 2019). Including workload metrics enables the model to capture the interplay between cumulative stress, biomechanical adaptation, and performance changes that may precede UCL injury. Table 3.3 presents the workload metrics used in the analysis. Player age was also included as an additional contextual variable. While these metrics exhibit substantial individual variability, reflecting differences in pitcher roles, physical conditioning, developmental stages, and team usage patterns, this variability is precisely what the individualised AE approach is designed to capture. By training separate models for each pitcher on their baseline performance, the model learns each pitcher's typical workload patterns and the individual relationships between age, workload, and pitch metrics. This enables detection of deviations from a pitcher's own normal performance profile rather than comparing against population averages. The analysis was restricted to six pitch types: four-seam fastball (FF), sinker (SI), slider (SL), cutter (FC), changeup (CH), and curveball (CU). These represent the core repertoire thrown by most MLB pitchers, as reflected in the usage percentages presented in Table 3.4, and align with pitch type classifications used in prior Statcast-based injury research (Kang et al., 2025; Mayo et al., 2021; Oeding et al., 2024). Less common pitch types were excluded due to low usage rates, which would result in excessive missing data at the game level.

Table 3.2 Descriptions of extracted pitching variables from Baseball Savant Statcast data

Pitching Metric	**Description**
pfx_x	“Horizontal movement from the catcher's perspective.”
pfx_z	“Vertical movement from the catcher's perspective.”
release_extension	“Release extension of the pitch in feet, as tracked by Statcast.”
release_pos_x	“Horizontal release position from the catcher's perspective.”
release_pos_z	“Vertical release position from the catcher's perspective.”
release_speed	“Pitch velocity at the release point”
release_spin_rate	“Spin rate of pitch.”
spin_axis	“Spin axis in the 2D X-Z plane in degrees (from 0 to 360).”

Note. All variables were extracted directly from Baseball Savant Statcast. Movement and position metrics are measured in feet from the catcher's perspective. Release speed is reported in kph. Spin rate is in RPM. Spin axis is in degrees (0–360).

Table 3.3 Workload and contextual features

Pitching Metric	**Description**	**Calculation Method**
days_rest	Number of days between consecutive pitching appearances	Calculated as the difference in calendar days between the current game and the previous game in which the pitcher appeared
age	Pitcher's age at the time of each game appearance	Calculated as the difference in days between the pitcher's date of birth and the date of the current game, expressed in years
total_pitch_count	Total number of pitches thrown in a given game appearance	Sum of all pitches recorded for the pitcher in that game across all pitch types
cumulative_7day	Cumulative pitch count over the preceding 7-day period	Rolling sum of total_pitch_count across all games within the previous 7 days, inclusive of the current game
cumulative_30day	Cumulative pitch count over the preceding 30-day period	Rolling sum of total_pitch_count across all games within the previous 30 days, inclusive of the current game

Note. All variables were derived from raw Statcast game-level data post-extraction. days_rest is missing for each pitcher's first recorded appearance. Rolling pitch count windows are inclusive of the current game date.

Table 3.4 League-wide pitch characteristics by pitch type in Major League Baseball (2019-2024)

Pitch Type	Abbreviation	Velocity (kph)	Spin Rate (RPM)	Usage Percentage
Change-up	CH	136.4 (± 1.4)	1775 (± 58)	10.6%
Curve	CU	126.2 (± 1.4)	2523 (± 81)	7.5%
Cutter	FC	143.0 (± 1.4)	2363 (± 54)	6.7%
Four-seam Fastball	FF	150.6 (± 0.7)	2277 (± 28)	34.4%
Sinker	SI	149.2 (± 0.9)	2143 (± 28)	17.3%
Slider	SL	136.7 (± 1.1)	2390 (± 46)	16.0%

Note. Data presents weighted league-wide mean (± SD) from MLB 2019-2024 seasons.

3.3.3. Pre-processing

The data extracted using baseballr were pivoted to a wider format so that each pitch type formed its own set of columns. For example, the feature 'release_speed' became 'FF_release_speed' for fastballs, and 'release extension' became 'FF_release_extension'. Pitch-level metrics were then aggregated to game-level averages, with each game representing a single observation. This aggregation reduces pitch-to-pitch variability while preserving temporal patterns across games, aligning with prior research demonstrating that injury-related changes manifest as progressive game-to-game trends (Mayo et al., 2021). All pitching metrics affected by handedness were normalised to right-handed orientation to account for differences between left- and right-handed pitchers. Most variables were normally distributed, as determined through graphical inspection and evaluation of skewness and kurtosis values. Minor skewness was observed in some variables, but the overall distributions were deemed acceptable for the intended analysis.

Missing values in the pitch-tracking data arise from two distinct sources. First, when pitchers do not utilise certain pitch types across the entire baseline period (e.g., a pitcher who never throws a CH), all metrics for that pitch type exhibit zero variance and are removed from the feature set entirely. This ensures the model only learns patterns for pitch types that are part of each pitcher's established repertoire, rather than attempting to reconstruct metrics for pitches that were never thrown. Second, within a pitcher's active repertoire, individual games may lack data for a particular pitch type if it was not thrown in that specific appearance. In these cases, the raw zero values are replaced with the training mean for each affected column after scaling, rendering them neutral (a scaled value of zero) rather than extreme outliers. This prevents games in which a pitcher chose not to throw a specific pitch from being misinterpreted as anomalous performance. Together, these approaches preserve the individuality emphasised throughout this study while ensuring that pitch repertoire variation does not distort the anomaly detection process.

The vanilla AE model approach accounts for developmental trends and individual variation through its extended baseline training period. By learning each pitcher's characteristic performance patterns over 400 days, including gradual improvements and routine adjustments, the model establishes a representation of "normal" that encompasses typical variation rather than a single static snapshot. When applied to the detection window, deviations flagged by the model represent changes that fall outside this learned range. Because the baseline remains fixed rather than adapting to progressive decline, the model maintains sensitivity to injury-related deterioration that might otherwise be normalised by continuously updating reference points (Jakubowski et al., 2022; Markovic et al., 2024). For consistency of timelines across pitchers, all data were anchored to the reference date corresponding to the last game prior to surgery. Pitching records were then reordered in reverse chronological order, with the time axis expressed as days before the last game. This ensured temporal alignment across pitchers despite differences in individual game schedules and provided a consistent framework for modelling injury progression.

3.3.4. Model Architecture

A Vanilla (feedforward) AE architecture was employed for anomaly detection (Jakubowski et al., 2022). Vanilla AE process each observation independently without incorporating temporal dependencies or sequential information, in contrast to recurrent architectures such as Long Short-Term Memory (LSTM) autoencoders. This independence property offers a critical advantage for detecting gradual performance deterioration. Vanilla autoencoders maintain a fixed representation of normal performance learned during training, enabling them to flag progressive decline that temporal models might inadvertently normalise (Jakubowski et al., 2022; Markovic et al., 2024). The model was implemented in R 4.4.2 using the keras3 and tensorflow packages, with hyperparameters optimised through grid search to ensure accurate anomaly detection (Abadi et al., 2015; Chollet & others, 2015; R Core Team, 2024).

When trained exclusively on a baseline period representing healthy performance, vanilla AE preserve a static reference standard. Because the model never updates its learned representation during the detection phase, it can identify gradual deterioration that adaptive architectures might incorporate into their evolving notion of normality. This characteristic is particularly relevant given Mayo et al.'s (2021) findings that UCLR is preceded by small but statistically significant game-to-game changes in pitch metrics over the 15 games leading up to surgery. A vanilla AE trained on healthy baseline performance can flag these subtle deviations because it compares each detection-window game against the original healthy standard rather than a continuously adapting baseline.

The AE architecture comprised an encoder network that compressed input data into a lower-dimensional latent space and a decoder network that reconstructed the original input from this compressed representation. The encoder consisted of two fully connected layers with 128 and 64 neurons respectively, using ReLU activation. The latent space dimensionality was set to 12 dimensions, selected through hyperparameter optimisation to balance information retention and compression. The decoder mirrored the encoder structure with two layers of 64 and 128 neurons, reconstructing the original input vector from the latent representation. Input dimensionality ranged from 22 to 59 features per pitcher depending on individual pitch repertoire, as pitch types not utilised during the baseline period were excluded from the feature set. The architecture implementation is provided in Appendix A. All analysis code, including the full pipeline scripts for data collection, feature engineering, propensity score matching, and visualisation, is available at https://github.com/ozakiryotaro/uclr-biomechanical-anomaly-detection.

Reconstruction error, calculated as the MSE between original and reconstructed values, served as the anomaly score for each game. Rather than adopting a single fixed anomaly threshold, an empirical threshold optimisation approach was employed. Reconstruction error distributions were computed on each pitcher's baseline data, and a range of candidate percentile thresholds (90th, 92.5th, 95th, 97.5th, and 99th) were evaluated across the full 46-pitcher sample. For each threshold, the proportion of detection window games flagged as anomalous was measured, with particular attention to whether flagging rates increased as games approached the reference date. Rather than nominating a single recommended operating point, results are reported across all five thresholds to demonstrate the robustness of findings to threshold choice and to reflect

the practical reality that optimal sensitivity will vary by deployment context. In applied settings, medical staff would reasonably adjust threshold stringency based on their tolerance for false positives relative to the consequences of missed detections, making a fixed recommended threshold inappropriate as a general prescription (Ali et al., 2013; Hossain et al., 2024). To examine the co-occurrence of feature-level deviations across time bins, Pearson pairwise correlations were computed between the per-game squared reconstruction errors of the top 10 features by overall mean squared error, separately within each 50-day bin.

3.3.5. Hyperparameters Optimisation

Hyperparameters were optimised through a two-stage grid search. An initial pilot across 5 pitchers evaluated 243 configurations, revealing that the 128-64 encoder architecture and 0.1 dropout rate consistently outperformed alternatives. These were fixed, reducing the search to 27 configurations applied across 34 representative pitchers. Each pitcher's baseline data was split chronologically (75/25 train/validation), with models trained using the Adam optimiser, mean squared error loss, and early stopping (patience = 20 epochs). The optimal configuration, selected by mean validation loss across pitchers, comprised 12 latent dimensions, a 128-64 encoder with mirrored decoder, learning rate of 0.001, batch size of 4, and dropout rate of 0.1. These hyperparameters were applied to all 46 pitchers in the subsequent analysis. The grid search implementation is provided in Appendix B.

3.3.6. Comparison with Propensity Score-Matched Controls

As an exploratory validation analysis, a propensity score-matched control group was constructed to evaluate whether the reconstruction error escalation observed in UCLR pitchers was specific to the pre-surgical period or consistent with normal performance variation in non-UCLR pitchers over an equivalent time window. Candidate control pitchers were drawn from all MLB pitchers active across the 2016–2024 seasons with no history of UCLR. Candidates were required to meet the same minimum data availability thresholds as the UCLR cohort, specifically at least 20 games within the baseline window and at least 5 games within the detection window, each anchored to the reference date of their prospective UCLR match.

For each UCLR pitcher, a logistic regression model was fitted to estimate propensity scores from four covariates: age at the reference date, throwing hand, mean pitch velocity, and cumulative pitch count during the baseline window. Nearest-neighbour matching without replacement was then applied, selecting the candidate control with the smallest absolute propensity score distance. Each control pitcher was eligible for selection only once. Covariate balance was assessed using standardised mean differences (SMDs), with values below 0.10 indicating negligible imbalance (Austin, 2011). Full pre- and post-matching balance diagnostics are reported in Appendix C. The reference date assigned to each matched control was set equal to their UCLR counterpart's last game before surgery, providing a pseudo-reference date for temporal alignment across the detection window, consistent with matched control designs used in previous UCL pitching research (Dillon et al., 2025; Mastroianni, Kunes, El-Najjar, et al., 2025; Whiteside

et al., 2016). One UCLR pitcher was excluded due to insufficient active controls meeting the minimum appearance thresholds within his baseline window, yielding a final matched sample of 45 pairs. AE models were trained and applied to all control pitchers using an identical pipeline to the UCLR cohort, and reconstruction error trajectories and anomaly rates across the four 50-day bins were compared between groups. Reconstruction error distributions were assessed for normality using the Shapiro-Wilk test. Given the strongly right-skewed distributions observed in both groups, between-group differences in median reconstruction error were evaluated using the Mann-Whitney U test (test statistic reported as W), a non-parametric alternative appropriate for non-normally distributed data. Comparisons were conducted across the full detection window and separately within each 50-day bin, with statistical significance set at $\alpha = 0.05$.

3.3.7. Univariate Comparison

To evaluate whether the multivariate AE encoding offered detection advantages over simpler single-feature monitoring, a Bonferroni-corrected univariate comparison was conducted. For each pitcher, individual per-feature anomaly thresholds were derived from the same percentile-based baseline distribution used for the AE reconstruction error, applied separately to each feature's raw values. To control for the inflated probability of spurious detections arising from simultaneous comparisons across the full feature set, the significance threshold was Bonferroni-corrected by dividing $\alpha = 0.05$ by the number of features tested for each pitcher (Andrade, 2019). A detection window game was classified as univariately anomalous if it exceeded the Bonferroni-corrected threshold on at least one feature. The proportion of AE-flagged games that were also flagged under this corrected univariate criterion was then calculated separately within each 50-day bin, enabling a direct assessment of the incremental detection contribution of the multivariate AE approach relative to individual feature monitoring.

3.4. Results

Anomaly rates across the five reconstruction error thresholds are presented in Figure 3.1. Across all thresholds, a greater proportion of detection window games were flagged as anomalous in bins closer to the last appearance. At the 95th percentile threshold, anomaly rates were 67.2%, 76.4%, 71.0%, and 73.6% for the 151-200, 101-150, 51-100, and 0-50 day bins respectively (n = 67, 106, 210, and 443 games across 12, 12, 25, and 46 pitchers).

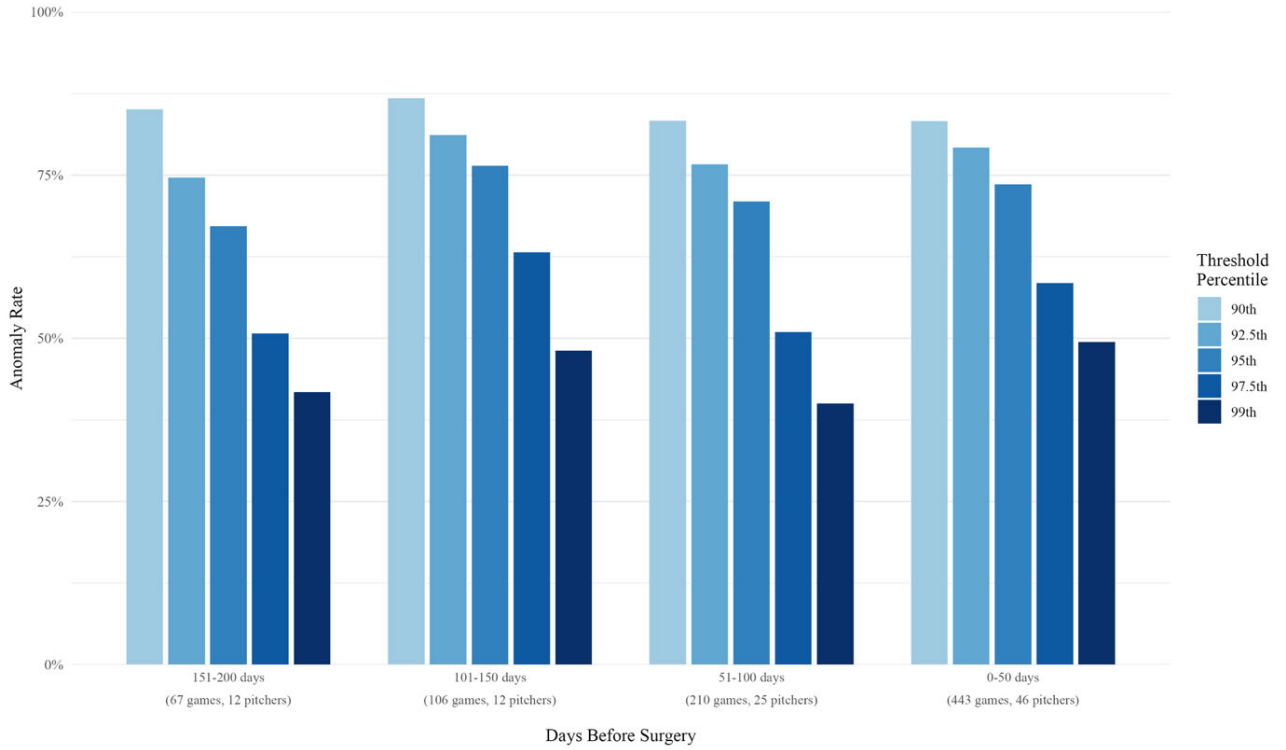

Figure 3.1 Proportion of detection window games flagged as anomalous by days before last appearance and reconstruction error threshold

Note. Each bar represents the mean proportion of games flagged as anomalous within each time bin across 46 pitchers. Five reconstruction error thresholds are displayed, derived from the 90th, 92.5th, 95th, 97.5th, and 99th percentiles of each pitcher's baseline reconstruction error distribution. Sample sizes vary across bins due to differences in individual game schedules and data availability.

The distribution of per-game reconstruction errors shifted progressively across time bins as surgery approached as shown in Figure 3.2. Mean reconstruction error increased from 0.877 (151-200 days) to 2.326 (0-50 days). Median values followed a similar pattern: 0.659 (151-200), 0.885 (101-150), 0.871 (51-100), and 1.059 (0-50) MSE. Distributions widened considerably in the 0-50 day bin (interquartile range [IQR]: 0.527 to 1.918), reflecting increasing inter-game variability in the period proximate to surgery. The y-axis was capped at the 95th percentile of the overall reconstruction error distribution (3.92 MSE) to aid visualisation; extreme values were excluded from display but retained in all analyses.

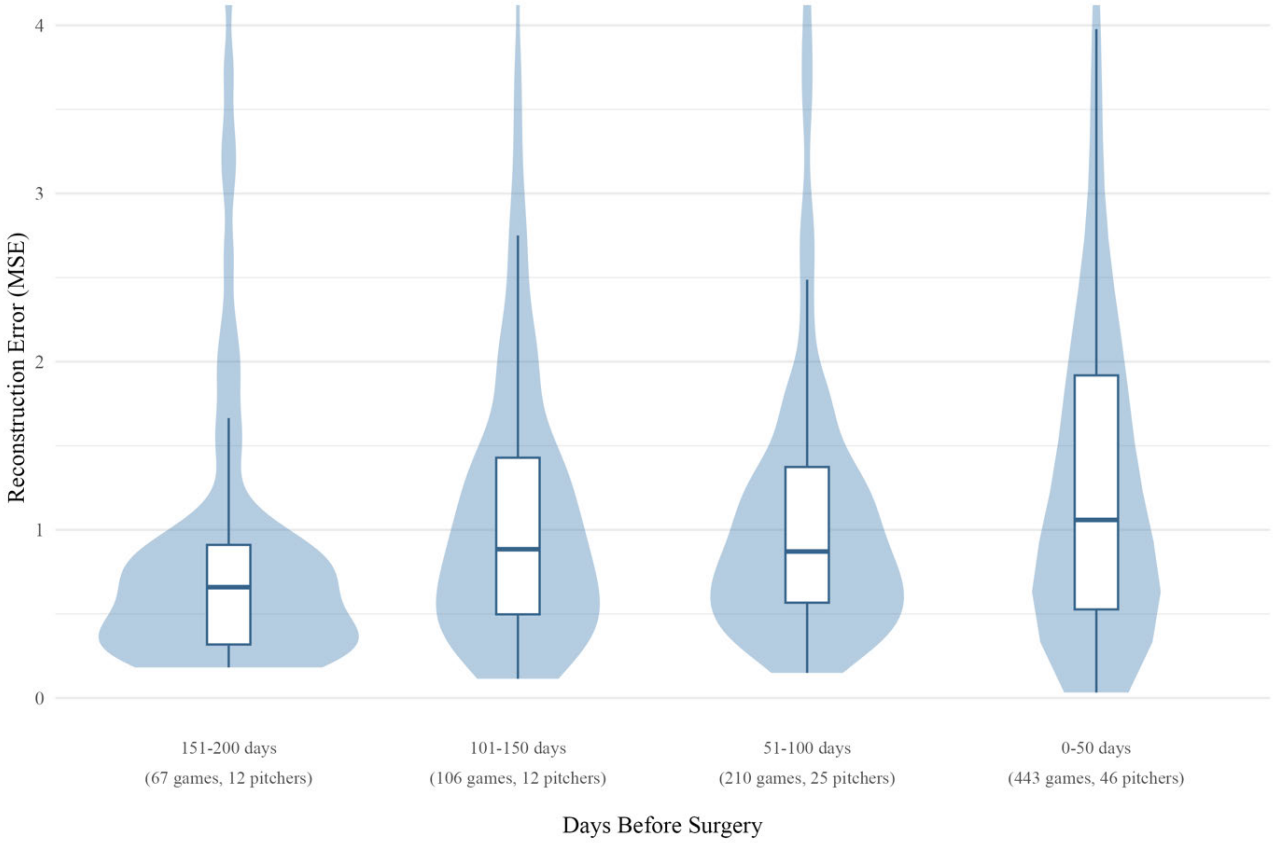

Figure 3.2 Distribution of per-game reconstruction error across time bins in the detection window

Note. Violin plots display the full distribution of mean squared reconstruction error for each time bin, with overlaid boxplots showing the median and interquartile range. The y-axis is capped at the 95th percentile of reconstruction error to aid visualisation; extreme outliers are excluded from display but retained in all analyses. Sample sizes vary across bins.

The top 20 features ranked by overall mean squared reconstruction error are displayed in Figure 3.3. CH_pitch_count produced the highest overall mean squared error (11.020), followed by days_rest (7.740) and SL_pfx_z (5.172). CH-related features (CH_pitch_count, CH_spin_axis, CH_pfx_z) and SL-related features (SL_pfx_z, SL_pitch_count, SL_spin_axis, SL_pfx_x, SL_release_pos_x, SL_release_pos_z) were prominent across the top 20, alongside FF release position and spin axis features (FF_release_pos_z, FF_release_pos_x, FF_spin_axis, FF_pfx_x) and workload metrics (days_rest, cumulative_30day). CH_pitch_count showed the largest temporal variation, with MSE rising from 0.835 to 1.832 across the 151-200 to 51-100 day bins before increasing substantially to 19.987 in the 0-50 day bin.

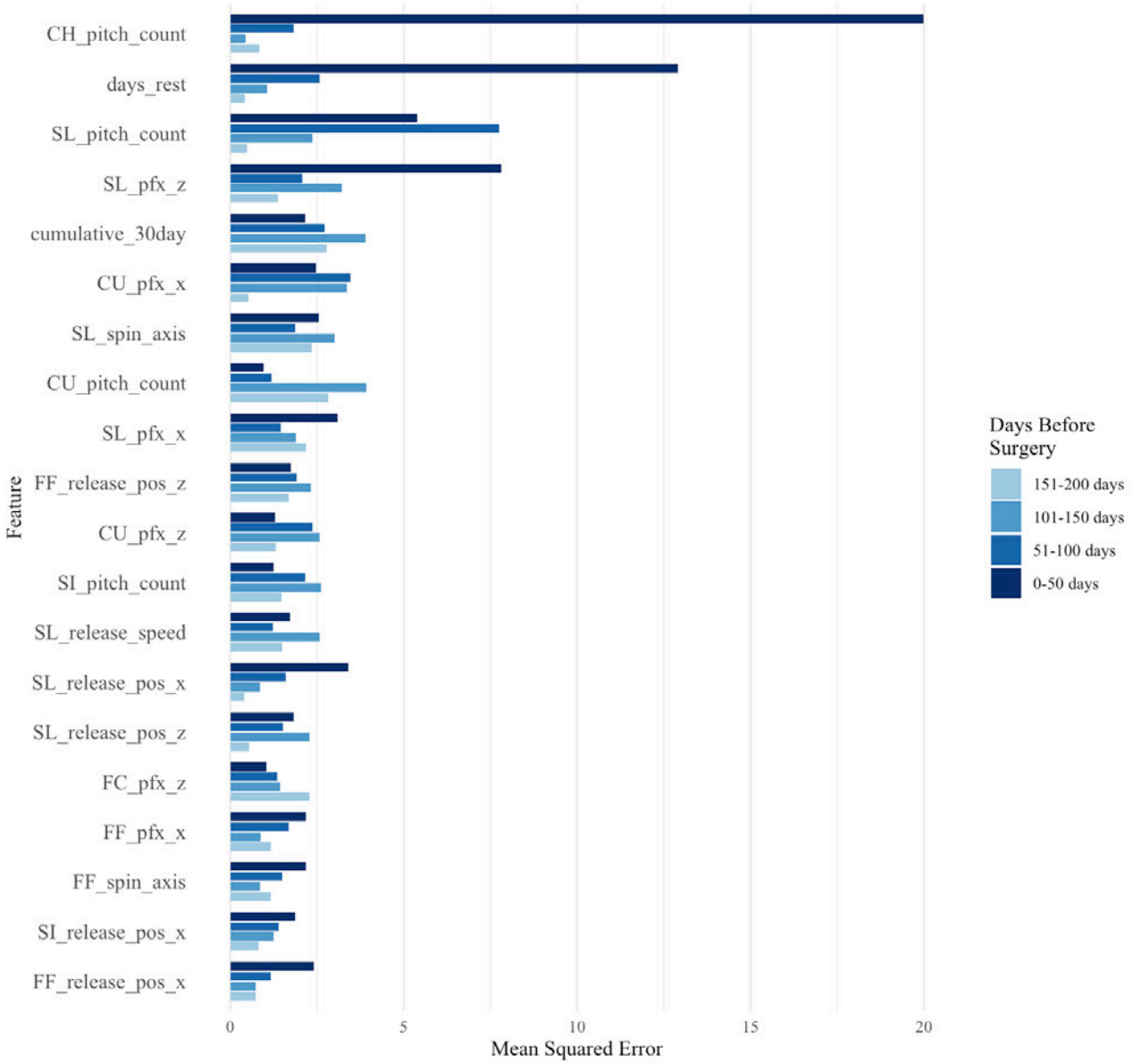

Figure 3.3 Mean feature-level reconstruction error by days before last appearance for the top 20 features

Note. Features are ordered by overall mean squared error across all time bins, with the highest-error features shown at the top. Each bar represents the mean squared error for a given feature within a time bin, averaged across all pitchers contributing games to that bin. Only features present in the baseline training data of at least one pitcher are included. Pitch type prefixes: CH = changeup; CU = curveball; FC = cutter; FF = four-seam fastball; SI = sinker; SL = slider. Feature suffixes: pfx_x = horizontal movement; pfx_z = vertical movement; release_pos_x = horizontal release position; release_pos_z = vertical release position; release_speed = pitch velocity at release; spin_axis = spin axis angle; pitch_count = number of pitches thrown of that type per game. Workload features: days_rest = days between consecutive appearances; cumulative_30day = cumulative pitch count over the preceding 30-day period.

Pearson pairwise correlations between the per-game squared reconstruction errors of the top 10 features are presented across time bins in Figure 3.4. In the 0-50 day bin, 32 of 45 feature pairs showed positive correlations (mean r = 0.167). The 51-100 and 101-150 day bins showed near-zero mean correlations (r = 0.096 and 0.075 respectively), with a greater proportion of negative pairs (22 and 18 of 45 respectively). The 151-200 day bin showed a mean correlation of r = 0.121 (23 of 45 pairs positive). Sample sizes in the

101-150 and 151-200 day bins were limited (n = 12 pitchers each), and correlations within these bins should be interpreted with caution.

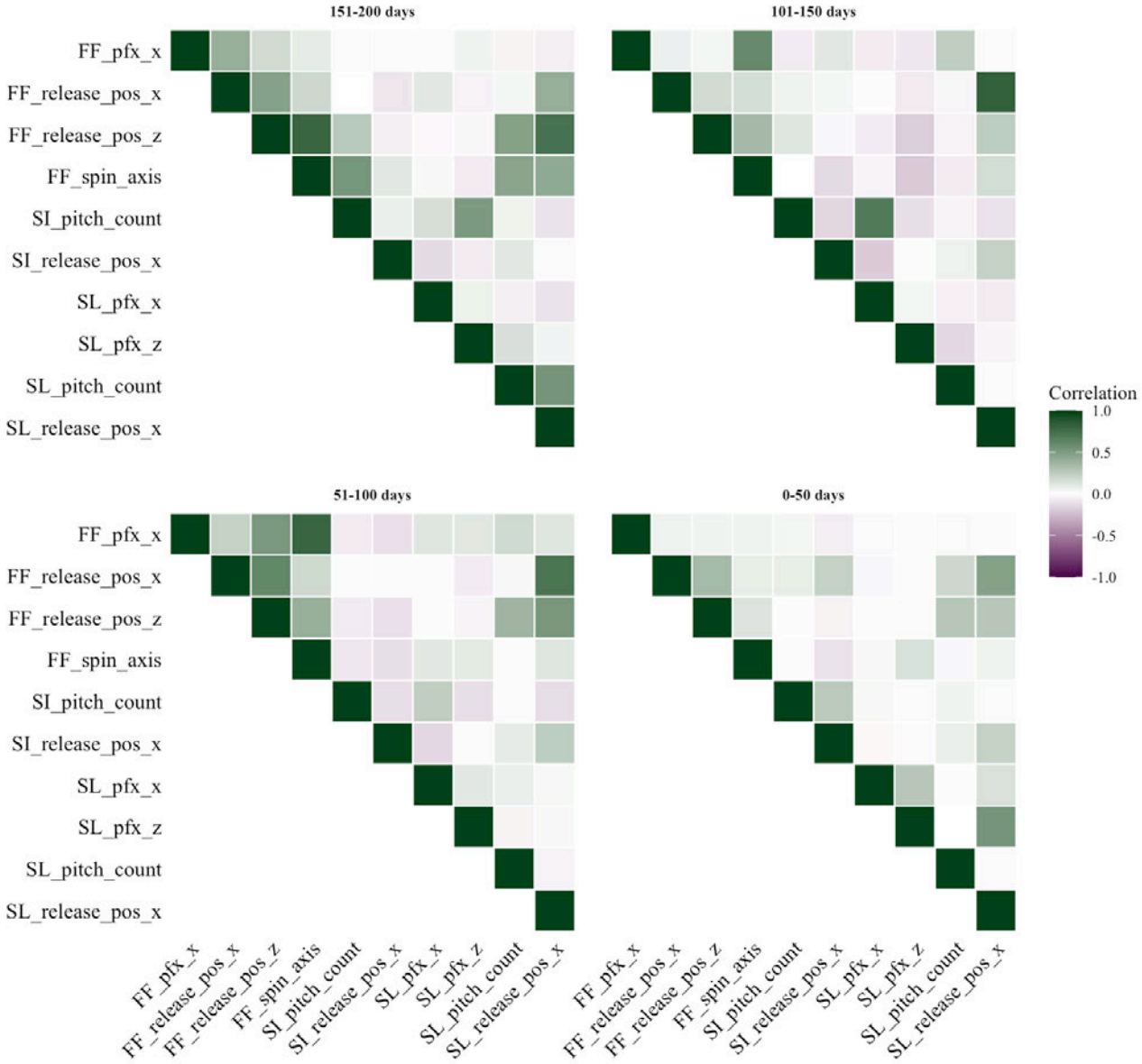

Figure 3.4 Pairwise correlations of feature-level reconstruction errors by days before last appearance

Note. Lower triangle correlation matrices display Pearson correlations between per-game squared reconstruction errors for the top 10 features, grouped by pitch type. Each panel represents one time bin. Positive values (green) indicate features whose reconstruction errors tend to deviate together within a game; negative values (purple) indicate features that deviate in opposing directions. Empty cells represent feature pairs where insufficient data were available to compute a stable correlation estimate. Sample sizes are small in the 151-200 and 101-150 day bins (12 pitchers each) and correlations should be interpreted with caution. Pitch type prefixes: FF = four-seam fastball; SI = sinker; SL = slider. Feature suffixes: pfx_x = horizontal movement; pfx_z = vertical movement; release_pos_x = horizontal release position; release_pos_z = vertical release position; spin_axis = spin axis angle; pitch_count = number of pitches thrown of that type per game.

Matched control and UCLR groups were well balanced across all four matching covariates, with no statistically significant differences in handedness, age, mean pitch velocity, or cumulative pitch count (Table

3.5). Reconstruction error distributions were strongly right-skewed in both groups across all comparisons (Shapiro-Wilk $p < 0.001$; skewness range 2.9 to 4.5), and Mann-Whitney tests were used for all between-group comparisons accordingly.

Table 3.5 Sample characteristics by group (n = 90)

Characteristic	UCLR (n = 45)	Control (n = 45)	p-value
Right-handed, n (%)	33 (73.3%)	32 (71.1%)	1.000
Age at reference (years)	29.0 ± 3.0	29.5 ± 3.3	0.410
Mean pitch velocity (kph)	88.8 ± 3.3	89.0 ± 3.8	0.824
Cumulative pitch count	1,705 ± 910	1,679 ± 1,102	0.907

Note. Data present mean ± SD. p-values derived from independent samples t-tests (continuous variables) and chi-square test (handedness). Cumulative pitch count reflects total pitches thrown during the baseline window (200–600 days before the reference date).

Mean reconstruction error trajectories across the detection window are presented in Figure 3.5. UCLR pitchers showed a broadly progressive escalation from 1.005 MSE (151-200 days) to 2.431 MSE (0-50 days). Control pitchers showed a non-progressive trajectory, with values of 1.455 (151-200 days), 1.620 (101-150 days), 1.216 (51-100 days), and 2.057 (0-50 days). Across the full detection window, UCLR pitchers showed a higher median reconstruction error than controls (1.409, IQR 0.913–2.277 vs 1.015, IQR 0.686–1.642), though this difference did not reach statistical significance (Mann-Whitney $W = 1235$, $p = 0.073$). Restricted to the 0-50 day bin, UCLR pitchers showed a significantly higher median reconstruction error than controls (1.513, IQR 0.713–2.375 vs 0.986, IQR 0.559–1.692; $W = 1245$, $p = 0.036$). A sensitivity analysis excluding one control pitcher with a documented concurrent shoulder injury during his detection window produced a stronger between-group difference in the 0-50 day bin ($W = 1245$, $p = 0.020$). Anomaly rates at the 95th percentile threshold in the 0-50 day bin were 72.9% for UCLR pitchers and 65.7% for controls.

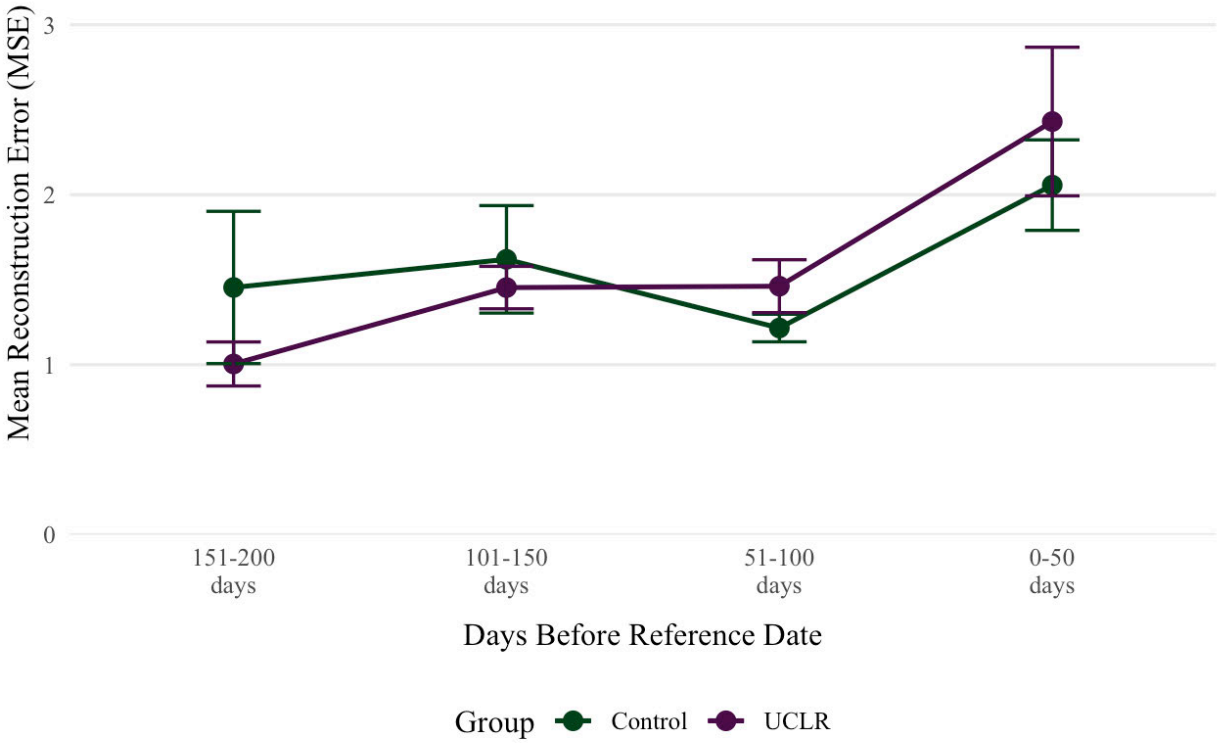

Figure 3.5 *Mean reconstruction error trajectories across the detection window: UCLR pitchers compared to propensity score-matched controls*

Note. Mean reconstruction error (MSE) across four successive 50-day bins preceding the reference date for UCLR pitchers (n = 45) and propensity score-matched controls (n = 45). Error bars represent the standard error of the mean.

3.5. Discussion

This study applied unsupervised ML-based anomaly detection to identify multivariate changes in pitching metrics in the games preceding UCLR in MLB pitchers. A vanilla AE was trained on each pitcher's baseline performance (200-600 days before last appearance) and applied to a detection window spanning the 200 days prior to last appearance. Across the 46-pitcher cohort, mean per-game reconstruction error increased from 0.877 MSE in the 151-200 day bin to 2.326 MSE in the 0-50 day bin, and anomaly rates were elevated in the period closest to surgery across all 5 reconstruction error thresholds evaluated (90th to 99th percentile). Feature-level analysis identified CH and SL metrics alongside workload variables as the primary contributors to reconstruction error, with the largest deviations concentrated in the 0-50 day bin. The following sections interpret these findings in the context of the existing biomechanical literature and discuss the methodological decisions underpinning the analysis.

3.5.1. Evidence of Pre-surgical Biomechanical Deterioration

The findings provide evidence that multivariate pitching performance deviated progressively from established baseline patterns in the period preceding UCL reconstruction. Mean MSE increased progressively across successive 50-day bins preceding the last pitching appearance before UCLR, rising from 0.877 (151-200 days) to 1.232 (101-150 days), 1.351 (51-100 days), and 2.326 (0-50 days), representing a 2.65-fold increase from the most distal to most proximate bin. The magnitude and consistency of this gradient suggests that the AE was detecting genuinely progressive biomechanical change rather than random fluctuation. The sharpest escalation occurred in the final 50 days, a period at which subclinical UCL stress likely begins to meaningfully compromise mechanical output. Median values showed a broadly similar pattern (0.659, 0.885, 0.871, and 1.059, respectively), though a modest dip between the 101-150 and 51-100 day bins indicates that the deterioration trajectory was not strictly monotonic across all pitchers at all time points. The IQR widened substantially in the 0-50 day bin (0.527 to 1.918) compared to 151-200 days (0.318 to 0.911), reflecting increasing inter-game variability in the period proximate to final pre-UCLR appearance.

This pattern corroborates a broader body of evidence suggesting that multivariate pitching performance deteriorates progressively in the period preceding UCL injury. At the seasonal level, baseline variability in velocity and release position alongside progressive velocity decline across the final five outings ($p = 0.019$) have been documented in the preinjury season (Mastroianni et al., 2026). At the game level, progressive reductions in FF velocity ($\tau b = -0.657$; $p < 0.001$), SL velocity, and spin rate across the 15 games preceding UCLR have been interpreted as reflecting subclinical UCL damage and compensatory mechanical adjustment (Mayo et al., 2021). This deterioration signal appears detectable across multiple timescales, extending to within a single outing. In a cohort of 7 acutely UCL injured MLB pitchers, every injury pitch exceeded the 95th percentile multivariate deviation threshold of each pitcher's individual control distribution (Dillon et al., 2025). Velocity was suppressed by a mean of 2.1 SD and arm angle reduced by 1.5 SD at the moment of failure, with 86% of pitchers also showing elevated cumulative mechanical deviation across the five FFs immediately preceding injury compared with only 7% of matched controls ($p < 0.001$).

The progressive reconstruction error gradient is most plausibly attributed to the accumulation of compensatory mechanical adaptations driven by declining biomechanical efficiency in the period preceding UCLR. When biomechanical efficiency deteriorates due to fatigue or subclinical tissue damage, pitchers may unconsciously modify pitch selection, release point, spin characteristics, and movement patterns to maintain competitive output (Birfer et al., 2019; Mayo et al., 2021; Yanagisawa, 2024). For instance, evidence that pitchers can alter stride length to reduce physiologic stress without sacrificing ball velocity illustrates how such compensatory adjustments may be enacted through subtle mechanical reorganisation that leaves no single dominant signature in the data (Crotin et al., 2014). Critically, these adaptations do not manifest in a single metric in isolation. Pitching mechanics function as an integrated kinetic chain in which lower body drive, trunk rotation timing, and arm path kinematics operate as interdependent components, meaning disruption to one element propagates across the full multivariate profile (Aguinaldo & Escamilla, 2019; Crotin et al., 2022). It is potentially this system-level disruption that the AE encodes as elevated

reconstruction error, as the complex interdependencies between features learned during the baseline period may become progressively destabilised in the period preceding UCLR.

Population-level models have identified significant biomechanical risk factors for UCL injury (DeFroda et al., 2016; Mastroianni et al., 2024; Whiteside et al., 2016). However, individualised analytical approaches applied directly to pitch-tracking data have demonstrated an ability to detect deterioration patterns that population-level models may be less sensitive to, with subject-specific baseline modelling successfully identifying mechanical deviation both at the game level (Kang et al., 2025) and at the pitch level immediately preceding acute UCL failure (Dillon et al., 2025). This distinction is particularly relevant when the signal of interest is individual deterioration from an established personal baseline rather than deviation from a population norm. Kinematic parameters demonstrate strong within-pitcher correlations that largely disappear in between-pitcher comparisons, with pitch velocity accounting for only 5% of EVT variance across pitchers (Post et al., 2015; Sakurai et al., 2024). Consistent with this, the persistently high CV across individual autoencoder models in the present study (CV > 250%) indicates that the magnitude of pre-surgical biomechanical disruption varied substantially between pitchers, reflecting genuine biological variability in injury presentation rather than methodological instability. This heterogeneity further supports the case for individualised rather than population-referenced detection, as a single group-level threshold would be insensitive to the differing trajectories through which individual pitchers approach UCLR. However, whether these kinematic changes reflect deliberate protective adaptation or are instead a consequence of fatigue and progressive tissue compromise remains unclear, and the present data cannot distinguish between these mechanisms (Murray et al., 2001). Nevertheless, that the AE captured this signal across 46 pitchers without any labelled injury data or population-level reference points suggests that the progressive multivariate disruption preceding UCLR is a sufficiently consistent phenomenon to be detectable through individualised unsupervised learning, even in the presence of substantial inter-pitcher variability in how and when that deterioration manifests.

3.5.2. A Two-Phase Deterioration Structure

The feature-level reconstruction error analysis revealed a two-phase deterioration structure consistent with the compensatory adaptation framework that has been proposed to precede UCLR in professional pitchers. An earlier phase, centred on the 101-150 day bin, was characterised by elevated errors in SL movement and cumulative workload features, most notably SL_pfx_z (MSE = 3.221), suggesting that mechanical disruption precedes overt performance decline by several months. A later phase, concentrated in the 0-50 day window, was dominated by a 24-fold escalation in CH usage error (CH_pitch_count MSE = 19.987 versus 0.835 at 151-200 days) and sharp increases in rest interval deviation, alongside broad increases across FF release position and spin axis features, consistent with a progressive shift from biomechanical instability toward strategic adaptation as the final appearance approaches. This progression mirrors the theoretical pathway wherein pitchers unconsciously modify pitch selection, release point, and spin characteristics to maintain

performance while protecting a deteriorating UCL (Fleisig et al., 2018; Kang et al., 2025; Mayo et al., 2021), and represents a pattern that univariate or population-level approaches would be unlikely to detect.

Among mechanical pitch features, SL-related metrics dominated the early-phase error profile. SL_pfx_z recorded the fourth-highest overall MSE (5.172) of all features analysed and exhibited a qualitatively distinct temporal pattern from the late-phase features, with its peak error occurring in the 101-150 day bin (MSE = 3.221) before declining modestly in the 51-100 day window (MSE = 2.076) and escalating again in the 0-50 day bin (MSE = 7.821). SL_pfx_x, SL_spin_axis, SL_release_pos_x, SL_release_pos_z, SL_release_spin_rate, and SL_release_extension all appeared among the higher-error features, indicating that SL mechanics were pervasively disrupted across multiple interdependent dimensions well in advance of surgery. The SL generates elbow valgus torque and shoulder internal rotation velocity comparable to the FF and significantly greater than the CH during arm cocking, placing substantial stress on medial elbow structures (Escamilla et al., 2017). Given that subtle arm path disruptions would be expected to manifest earliest in the pitch type most sensitive to internal rotation dynamics, the early prominence of SL features in the reconstruction error profile is biomechanically coherent (Fleisig et al., 2018). Horizontal spin axis shift was identified as one of three key predictors in the best-performing UCL injury model (Kang et al., 2025), and greater SL usage has been found to independently predict future injuries across 3,808 pitcher-years (Oeding et al., 2024), lending further support to the interpretation that SL mechanics constitute a particularly sensitive early indicator of emerging UCL stress.

In the 0-50 day window, the dominant signal shifted from mechanical disruption toward observable strategic and workload adaptation. The feature with the greatest overall MSE across the detection window was CH_pitch_count (MSE = 11.020), with a sharply escalating temporal trajectory: error was modest in the 151-200 day bin (MSE = 0.835), rose incrementally through the 51-100 day period (MSE = 1.832), then increased dramatically in the 0-50 day window (MSE = 19.987), representing a 24-fold escalation relative to the earliest detection bin. The prominence of pitch count as the highest-error feature reflects usage patterns rather than mechanics in isolation, yet warrants careful biomechanical interpretation. This pattern suggests that CH usage undergoes a pronounced and abrupt departure from each pitcher's established baseline in the final weeks before their final appearance, rather than accumulating gradually across the detection window. Pitchers approaching UCL failure may alter their CH usage in opposing directions, either increasing it as a velocity-sparing strategy to reduce elbow valgus torque demands, or abandoning it as elbow discomfort progressively constrains their repertoire. Either pattern constitutes a meaningful departure from the model's learned baseline, producing elevated reconstruction error. The CH’s relatively low velocity and arm-speed deception requirements may make it a pitch of preference when pitchers seek to reduce elbow stress while maintaining effectiveness. The CH produces significantly lower kinetic values than the FF across the majority of elbow and shoulder force and torque variables, with maximum elbow extension and shoulder internal rotation velocities both reduced during CH delivery relative to FF, SL or CU mechanics (Escamilla et al., 2017; Fortenbaugh et al., 2009). A pitcher with subclinical UCL compromise may therefore gravitate

toward increased CH usage as a load-management strategy, even without conscious awareness of doing so, rendering atypical usage patterns a plausible late-stage indicator of UCL integrity deterioration.

A comparable interpretive ambiguity applies to the workload features, particularly days_rest, which recorded the second-highest overall MSE (7.740). Although deviation in rest interval and cumulative pitch count is consistent with the biomechanical fatigue framework described above, it may alternatively reflect intentional load management by medical or coaching staff already aware of subclinical symptoms, in which case workload deviation would represent a downstream consequence of existing clinical knowledge rather than an independent early warning signal. The present data cannot distinguish between these mechanisms, as reconstruction error captures the magnitude of deviation from an established baseline but provides no information regarding its origin. Prospective designs incorporating clinical records and injury reporting timelines would be necessary to resolve this ambiguity, and the contribution of workload features to reconstruction error in the present study should therefore be interpreted with caution.

3.5.3. Repertoire Composition and Compensatory Adaptation

This pattern extends to repertoire composition more broadly. A smaller repertoire of unique pitch types has been identified as a significant independent predictor of UCL reconstruction, with each additional unique pitch type associated with a 33% reduction in the odds of undergoing surgery (Whiteside et al., 2016). This finding has been interpreted as evidence that greater pitch variety distributes mechanical load across differing joint positions and force profiles, potentially attenuating the repetitive microtrauma thought to underpin overuse UCL injury (Escamilla et al., 1998; Fleisig et al., 1995). Pitchers who rely predominantly on a limited number of high-velocity pitch types may therefore subject the UCL to more consistent, high-magnitude stress across outings. Any detectable shift in repertoire composition during the detection window, whether a narrowing of pitch variety or an atypical redistribution of usage, may thus represent an additional multivariate signal that the AE is sensitive to. However, the protective value of pitch variety and strategic pitch selection adjustments should not be overstated, as the evidence also points to important limitations and complications in this compensatory framework.

Two of seven pitchers in the Dillon et al. (2025) study sustained acute UCL injuries on a CH rather than on a higher-stress FF thrown earlier in the same outing, suggesting that once the ligament is sufficiently primed for failure, even a relatively low-stress pitch may trigger rupture. This is further complicated by evidence that higher velocity, superior command, and increased usage of the CH are themselves independently associated with elevated UCL injury risk (Mastroianni et al., 2025), indicating that the CH is not simply a protective pitch type. Furthermore, pitch-mix compensation appears to have inherent limits, as pitchers typically experience monotonic velocity declines across FF and SL as surgery approaches (Dillon et al., 2025; Mastroianni et al., 2026; Mayo et al., 2021), suggesting that subclinical UCL damage eventually compromises the quality of the entire arsenal regardless of selection strategy. Taken together, these findings suggest that atypical CH usage and broader repertoire shifts may represent early compensatory responses to

UCL deterioration, but that such adaptations are ultimately unable to prevent the progressive mechanical decline that manifests across the full multivariate profile as surgery approaches and which the present model appears to detect.

3.5.4. Multivariate vs. Univariate Detection

A Bonferroni-corrected comparison of multivariate and univariate detection was conducted to address the empirical question posed in the study rationale of whether the AE's multivariate encoding offers detection advantages over simpler single-feature monitoring. Individual per-pitcher feature thresholds were Bonferroni-corrected by dividing the significance level by the number of features tested, ensuring that the expected false-positive rate across all 58 features was held at 5% and controlling for the inflated probability of spurious detections inherent in multiple simultaneous comparisons (Andrade, 2019). Under this corrected framework, only one AE-flagged game in the 51-100 day bin and one in the 0-50 day bin were not also flagged by at least one individual feature, representing 0.7% and 0.3% of AE-flagged games in those bins respectively. This finding suggests that the deterioration in individual pitch features before surgery was strong enough to be picked up by both approaches, raising the question of whether the added complexity of the autoencoder provides meaningful additional value over simpler univariate methods when the signal in individual features is already this pronounced (Mayo et al., 2021).

However, the practical advantage of the AE lies not in identifying games that univariate monitoring would miss, but in how detection is achieved. A practitioner implementing univariate monitoring across this feature set would need to maintain 58 individually calibrated, pitcher-specific thresholds, interpret simultaneous threshold crossings across multiple features on any given flagged game, and determine which combinations of crossings are meaningful versus coincidental. The AE collapses this interpretive burden into a single composite reconstruction error score per game, whose baseline distribution is well-characterised for each pitcher individually. This operational simplification is relevant in applied sports medicine settings where clinical staff have limited time and require clear, actionable outputs rather than complex multi-feature analyses.

A further limitation of univariate monitoring is revealed by the pairwise feature correlation structure (Figure 3.4). Within-pitch-type feature errors are strongly correlated throughout all four time bins: FF_release_pos_x and FF_release_pos_z co-deviate consistently, as do SL_release_pos_x and SL_release_pos_z, reflecting the biomechanical coupling of release point coordinates within a given pitch type. Under univariate monitoring, a single mechanical disruption in SL release mechanics could simultaneously trigger alarms across multiple individual features, creating interpretive noise from what is effectively one underlying signal. The AE learns the relationships between features during training and registers coordinated deviations as a single anomaly score rather than triggering multiple independent alarms, a property that has been demonstrated to improve anomaly detection sensitivity in multivariate time series data relative to univariate monitoring approaches (Garg et al., 2022; Li & Jung, 2023; Pota et al., 2023). This property is particularly relevant in the 0-50 day

bin, where feature correlations are strongest across all time bins (Figure 3.4), suggesting that deterioration in the final weeks before last appearance involves coordinated disruption across multiple features simultaneously rather than isolated changes in individual metrics. The AE's individualised architecture accommodates this heterogeneity implicitly through its per-pitcher baseline, whereas a population-level univariate threshold calibrated to the cohort mean cannot.

The most substantive advantage of the AE over univariate approaches in this study is not detection sensitivity but post-hoc interpretability. The feature-level reconstruction error decomposition, which identified the two-phase deterioration structure and the disproportionate contribution of SL movement and CH usage features, is only possible because the AE produces a structured, decomposable error signal across all features simultaneously. A univariate monitor produces a binary flag per feature with no capacity for the kind of cross-feature, cross-temporal synthesis that generated the novel findings of this study. In this sense, the AE may function less as a superior detector and more as a superior analytical lens through which the multivariate signature of pre-surgical deterioration can be characterised.

3.5.5. Detection Anomaly Rate Patterns Across Thresholds

Anomaly rates across all five thresholds were non-monotonic, with rates rising from the 151-200 to the 101-150 day bin, dipping in the 51-100 day bin, then rising again in the 0-50 day bin across every percentile evaluated (Figure 3.1). This pattern may reflect the genuinely non-linear nature of pre-surgical UCL deterioration rather than being a methodological artefact or a product of threshold sensitivity. Dillon et al. (2025) independently observed that pitch mechanics in the period immediately preceding acute UCL injury were characterised by instability and inconsistency rather than a clear progressive decline, with no consistent directional trends identifiable across the five pitches before injury. Such acute presentations account for nearly half of all UCL reconstructions in some surgical cohorts (Cain et al., 2010). Although they appear instantaneous, they are widely considered acute-on-chronic events representing the final rupture of a ligament already structurally compromised by repetitive microtrauma (Erickson et al., 2015), a notion supported by imaging evidence of chronic damage even in first-time tears (Garcia et al., 2019; Kooima et al., 2004). This has been interpreted as a distinct biomechanical profile of acute UCL failure, in which short-term mechanical instability is superimposed on a ligament already weakened by chronic stress, ultimately pushing it past the point of failure rather than through a gradual, predictable decline (Dillon et al., 2025). The dip in anomaly rates in the 51-100 day bin suggests that deterioration does not follow a smooth linear trajectory. Instead, pitchers appear to pass through periods of relative mechanical stability interspersed with phases of elevated deviation, producing the wave-like pattern observed here. This interpretation is supported by the consistency of the pattern across all five thresholds (Figure 3.1), which indicates it may reflect a structural feature of the pre-surgical deterioration process rather than an artefact of any particular threshold choice. The non-monotonic profile is consistent with the two-phase deterioration structure identified earlier, where the 51-100 day window represents a transitional period between early mechanical disruption and the terminal compensatory phase rather than genuine stabilisation.

Threshold stringency had a discernible effect on the discriminative capacity of the anomaly rate signal. At the most lenient threshold (90th percentile; p90), the 0-50 day bin anomaly rate (83.3%) was marginally lower than the 151-200 day bin rate (85.1%), yielding an escalation ratio of 0.98, effectively no discrimination between the earliest and latest detection windows. At the strictest threshold (p99), the 0-50 day bin rate (49.4%) exceeded the 151-200 day bin rate (41.8%) by a factor of 1.18, representing the strongest separation across the threshold range evaluated. This pattern indicates that the pre-surgical signal is concentrated in the tail of the reconstruction error distribution. The most extreme deviations from each pitcher's baseline profile are disproportionately concentrated in the final 50 days, even if moderate deviations are distributed relatively evenly across the detection window. Escalation ratios increased monotonically with threshold stringency across all five percentiles (p90: 0.98x, p92.5: 1.06x, p95: 1.10x, p97.5: 1.15x, p99: 1.18x), suggesting that stricter thresholds progressively isolate a more temporally specific signal. No single threshold is nominated as a recommended operating point, as optimal sensitivity will vary by deployment context. Practitioners prioritising comprehensive surveillance may favour lenient thresholds, whilst those seeking a cleaner, lower-noise signal may prefer stricter settings (Ali et al., 2013; Hossain et al., 2024).

Reconstruction error magnitude provided a cleaner characterisation of the temporal deterioration pattern than anomaly rates alone. Mean reconstruction error escalated from 0.877 in the 151-200 day bin to 2.326 in the 0-50 day bin, a 2.7-fold increase, with a broadly monotonic trajectory across bins (Figure 3.2). The standard deviation of reconstruction error grew substantially over the same period (0.860 to 8.733), indicating that whereas the central tendency escalated progressively, individual pitcher trajectories diverged markedly in the terminal window. This divergence is consistent with the feature-level findings discussed earlier, where a subset of pitchers showed extreme CH_pitch_count and days_rest deviations that pulled the cohort mean upward whereas other pitchers showed more modest, distributed signal. Taken together, these findings suggest that the magnitude and trajectory of reconstruction error over time may carry more diagnostic information than simply flagging games as anomalous above a fixed threshold, a question that future work employing continuous risk scoring approaches could directly evaluate.

3.5.6. Exploratory Analysis: Comparison with Propensity Score-Matched Controls

The matched control comparison was conducted as an exploratory validation analysis to contextualise the UCLR reconstruction error gradient from the main analysis. Groups were well balanced on all matching covariates, confirming that observed between-group differences cannot be attributed to systematic differences in age, handedness, velocity, or workload. The most informative contrast between groups lies in the directionality of their trajectories rather than the magnitude of reconstruction error at any single time point. UCLR pitchers showed a broadly progressive escalation across the detection window consistent with the deterioration signal described in the main analysis, whereas control pitchers showed an irregular, non-directional trajectory, rising in the 101-150 day bin before declining in the 51-100 day bin and escalating again in the 0-50 day bin. The key distinction between groups is therefore not the magnitude of the terminal escalation, which was comparable between groups, but the absence of a consistent directional deterioration

pattern across the full detection window in controls. This divergence in trajectory shape is consistent with matched control evidence from the broader UCL pitching literature. At the seasonal level, pitch-tracking metrics have been shown to distinguish UCLR cases from matched controls (Mastroianni et al., 2026; Whiteside et al., 2016), and data-driven individualised approaches have demonstrated detectable differences between injured and non-injured pitchers up to 100 days prior to surgery (Kang et al., 2025). At a finer temporal resolution, mechanical deviation in the period immediately preceding acute UCL failure has been found to be present in a substantially greater proportion of injured pitchers than matched controls (Dillon et al., 2025).

Between-group differences in median reconstruction error were non-significant across the full detection window ($W = 1235$, $p = 0.073$) but reached statistical significance in the 0-50 day bin ($W = 1245$, $p = 0.036$), indicating that the between-group separation is most pronounced in the period proximate to surgery. A sensitivity analysis excluding one control pitcher with a documented concurrent shoulder injury during his detection window strengthened this finding ($p = 0.020$), suggesting the primary result is robust rather than dependent on a single influential observation. The control group's reconstruction error distribution was strongly right-skewed in the 0-50 day bin (skewness = 4.494, kurtosis = 25.596), driven principally by this pitcher, whose documented concurrent shoulder injury during the detection window provides a plausible explanation for his elevated reconstruction error. This observation illustrates a broader limitation of assuming a truly injury-free control group from publicly available records alone. As noted by Whiteside et al. (2016), some control pitchers may sustain throwing-related injuries not captured in public databases or may subsequently undergo UCLR themselves, and case-control designs of this nature cannot fully account for between-person differences in mechanics, injury history, and physical development through matching alone (Mastroianni et al., 2026; Whiteside et al., 2016), meaning the assumption of a truly comparable control group cannot be confirmed with certainty.

The non-trivial reconstruction error observed in control pitchers throughout the detection window is consistent with evidence that within-individual biomechanical variability is an inherent feature of healthy pitching mechanics (Camp et al., 2017; Fleisig et al., 2009) , and that fatigue and accumulated workload drive inter-session kinematic changes even in non-injured pitchers (Birfer et al., 2019). Furthermore, the limited sample sizes in the 151-200 and 101-150 day bins preclude reliable between-group comparisons outside of the 0-50 day window, and the bin-based trajectory presented in Figure 3.5 should be treated as descriptive rather than inferential. Taken together, these findings provide preliminary exploratory evidence that the reconstruction error escalation observed in UCLR pitchers in the proximate pre-surgical window reflects a signal that is at least partially specific to the pre-surgical period, though prospective validation in larger samples will be necessary to establish the clinical specificity of this signal more firmly.

3.6. Limitations

Several limitations of the present study warrant consideration. The analytical sample comprised 46 MLB pitchers, determined by data availability and inclusion criteria rather than a priori power calculations. While sufficient for the exploratory aims of this study, this limits generalisability and prevents subgroup analyses by handedness, role, or injury severity, a constraint shared by comparable studies in this area (Dillon et al., 2025; Kang et al., 2025; Mayo et al., 2021). The exclusive reliance on publicly available Statcast data meant that medical records, concurrent injury history, and non-UCL musculoskeletal conditions were inaccessible. Pitchers who sustained shoulder, lower body, or other injuries during the baseline or detection windows could not be systematically excluded, introducing a potential source of confounding. However, given that UCL injuries represent the primary cause of extended absence among MLB pitchers, the detected patterns are likely to predominantly reflect UCL-related biomechanical changes rather than incidental injury noise (Kang et al., 2025). Relatedly, anchoring the detection window to each pitcher's last appearance before UCLR surgery means the cohort is restricted to pitchers who proceeded to surgical reconstruction, excluding those managed nonoperatively and potentially biasing findings towards more severe injury presentations that limit applicability to the full spectrum of UCL pathology (Mastroianni et al., 2025). It should also be noted that Statcast metrics represent downstream kinematic outputs rather than direct measures of internal joint loading or ligamentous strain. Reconstruction error therefore reflects deviation in observable pitching performance and cannot be interpreted as a direct measure of UCL integrity, consistent with the observation that mechanical deviation metrics remain indirect proxies for tissue-level stress (Dillon et al., 2025).

The matched control comparison is subject to several additional limitations. The analysis was underpowered for formal between-group inference in the earlier time bins, where control pitcher sample sizes were limited to 10 and 13 pitchers in the 151-200 and 101-150 day bins respectively, precluding reliable trajectory comparisons outside of the 0-50 day window. The pseudo-reference date design assigns each control pitcher a reference date based on their matched UCLR counterpart's last game before surgery, meaning controls are not observed across a genuinely equivalent injury-risk period. One control pitcher sustained a concurrent shoulder injury during his detection window, elevating his reconstruction error and contributing to right skew in the control group distribution, though a sensitivity analysis confirmed the primary 0-50 day result was robust to his exclusion.

At the strictest threshold evaluated (p99), 13.0% of UCLR pitchers produced no flagged games in the 0-50 day bin, indicating that the approach does not capture all injury trajectories even under the most sensitive operating conditions. This likely reflects genuine biological heterogeneity in how UCL deterioration manifests across individuals, but represents a meaningful boundary on clinical applicability. Finally, findings are restricted to MLB pitchers and may not generalise to other competitive levels. Future work incorporating Minor League, collegiate, or international cohorts alongside multimodal data such as wearable sensors and biomechanical measurements would strengthen both the reliability and generalisability of the approach.

3.7. Conclusion

This study demonstrated that ML-based anomaly detection applied to individualised pitcher baselines can identify progressive multivariate deviations in pitching performance preceding UCLR in MLB pitchers. Reconstruction error escalated across the 200-day detection window, with the sharpest increase concentrated in the final 50 days, and anomaly rates were elevated proximate to surgery across all five thresholds evaluated. Feature-level analysis revealed a two-phase deterioration structure: SL mechanics and workload variables showed early disruption, followed by a pronounced late-phase shift in CH usage and rest interval patterns. This progression is consistent with a movement from subclinical biomechanical instability toward compensatory adaptation. An exploratory matched control comparison provided preliminary support for the specificity of this signal, with UCLR pitchers showing significantly higher reconstruction error than propensity score-matched controls in the proximate pre-surgical window, alongside a directional escalation pattern across the full detection window that was not observed in controls. These findings extend the existing literature by demonstrating that individualised unsupervised anomaly detection can characterise the multivariate signature of pre-surgical deterioration without requiring labelled injury data or population-level reference points. Although sample size and reliance on publicly available data limit generalisability, the consistency of the deterioration signal across pitchers, thresholds, and features supports the methodological feasibility of this approach as a basis for future injury surveillance research. Prospective validation in larger cohorts incorporating biomechanical and medical record data will be necessary to establish its clinical utility as a component of pitcher health monitoring in professional baseball.

Chapter 4. Summary and Future Directions

The primary aim of the study presented in this thesis was to determine whether unsupervised ML-based anomaly detection, applied to individualised pitcher baselines derived from Statcast pitch-tracking data, could identify multivariate pitching metric changes in MLB pitchers in the period preceding UCLR. We hypothesised that pitcher-specific vanilla AE models trained on baseline performance data would detect progressive deviations in pitching metrics across the 200-day detection window, with escalating reconstruction error in the period proximate to surgery. The key findings were that: (i) mean per-game reconstruction error escalated progressively across the detection window, rising 2.7-fold from the most distal to most proximate 50-day bin, with the sharpest increase concentrated in the final 50 days before last appearance; (ii) feature-level analysis revealed a two-phase deterioration structure, characterised by early disruption in SL mechanics and cumulative workload variables, followed by a pronounced late-phase shift in CH usage and rest interval deviation; and (iii) an exploratory matched control comparison provided preliminary evidence that the reconstruction error escalation in the proximate pre-surgical window was at least partially specific to UCLR pitchers relative to propensity score-matched non-UCLR controls. This chapter discusses the primary practical applications of these findings and outlines recommendations for future research.

4.1. Application to Practice

The results of this study extend the existing pitch-tracking literature by demonstrating that pitcher-specific AE models, trained without labelled injury data or population-level reference points, can detect a progressive multivariate deterioration signal in the period preceding UCLR. The approach has not been prospectively validated, and the exploratory matched control comparison, though encouraging, was underpowered outside the proximate pre-surgical window. As such, the present findings do not support recommendation of the AE framework as a ready monitoring tool. Rather, they provide justification for further investment in developing and prospectively validating this class of approach, and demonstrate that the data infrastructure and methodological framework necessary to do so already exist within the current MLB tracking environment. In line with Kang et al. (2025), any future application of model outputs of this nature should be viewed as supportive rather than definitive, with clinical decisions regarding pitcher readiness or referral remaining the responsibility of qualified medical staff.

The feature-level findings nonetheless generate specific, testable hypotheses regarding the clinical patterns that may precede UCL failure. The two-phase deterioration structure identified here, characterised by early disruption in SL mechanics and workload accumulation approximately 100 to 150 days before last appearance, followed by a pronounced late-phase shift in CH usage in the final 50 days, provides a basis for targeted prospective observation. It should be noted, however, that workload deviation may reflect

intentional load management by medical or coaching staff already aware of subclinical symptoms, rather than constituting an independent early warning signal. Prospective designs with access to clinical records would be necessary to distinguish between these explanations, and the applied value of workload features as surveillance inputs should therefore be interpreted with caution pending such investigation. Should the SL and CH patterns be replicated in larger cohorts, they would warrant direct clinical investigation, as early-phase SL disruption and late-phase CH deviation may represent observable indicators of emerging UCL stress that could inform the design of future validated surveillance tools. Similarly, the fact that Statcast data are captured in real time during every MLB game raises the possibility that the game-level framework presented here could inform the development of within-game pitch-level surveillance, as proposed by Dillon et al. (2025), though this would require substantially further methodological development before applied use could be considered.

A further limitation of the present framework is its reliance on Statcast pitch-tracking metrics as proxies for underlying biomechanical function. As discussed in Chapter 3, the biomechanical literature identifies EVT, kinetic chain sequencing, and segmental coordination as the primary causal mechanisms of UCL injury, yet none of these variables are directly measurable from public Statcast data. Pitch-tracking outcomes such as release speed, movement, and spin characteristics reflect the downstream products of these mechanical processes rather than the processes themselves, meaning the AE learns deviations in observable outputs rather than in the joint-level dynamics that drive UCL loading. This represents an inherent constraint of any publicly available data-driven approach; however, MLB organisations have access to proprietary in-game biomechanical data that extends well beyond what is publicly released. KinaTrax, a markerless motion-capture system deployed across a number of MLB organisations, enables collection of joint-level kinematic data including arm path dynamics, hip-shoulder separation, and elbow kinematics across every pitch of every outing during live game conditions without requiring any attached sensors (Giordano et al., 2024). Kinetic data derived from such systems, including EVT estimates, have been identified as an important direction for future research (Fleisig et al., 2024; Giordano et al., 2024). Future studies conducted in partnership with MLB organisations could therefore apply the individualised AE framework directly to KinaTrax-derived kinematic and kinetic data, grounding the anomaly detection signal in the primary biomechanical determinants of UCL injury rather than their downstream pitch-tracking correlates.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., … Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems* [Computer software]. https://www.tensorflow.org

Aguinaldo, A., Buttermore, J., & Chambers, H. (2007). Effects of Upper Trunk Rotation on Shoulder Joint Torque among Baseball Pitchers of Various Levels. *Journal of Applied Biomechanics, 23(1), 42-51.* https://doi.org/10.1123/jab.23.1.42

Aguinaldo, A., & Chambers, H. (2009). Correlation of throwing mechanics with elbow valgus load in adult baseball pitchers. *The American Journal of Sports Medicine*, *37*(10), 2043–2048.

Aguinaldo, A., & Escamilla, R. (2019). Segmental power analysis of sequential body motion and elbow valgus loading during baseball pitching: Comparison between professional and high school baseball players. *Orthopaedic Journal of Sports Medicine*, *7*(2), 2325967119827924. https://doi.org/10.1177/2325967119827924

Alain, G., & Bengio, Y. (2014). *What Regularized Auto-Encoders Learn from the Data Generating Distribution* (arXiv:1211.4246). arXiv. https://doi.org/10.48550/arXiv.1211.4246

Albright, J. A., Jokl, P., Shaw, R., & Albright, J. P. (1978). Clinical study of baseball pitchers: Correlation of injury to the throwing arm with method of delivery. *The American Journal of Sports Medicine*, *6*(1), 15–21. https://doi.org/10.1177/036354657800600104

Ali, M. Q., Al-Shaer, E., Khan, H., & Khayam, S. A. (2013). Automated Anomaly Detector Adaptation using Adaptive Threshold Tuning. *ACM Trans. Inf. Syst. Secur.*, *15*, 17:1-17:30.

Almonroeder, T., Jones, M. T., Fields, J. B., Erickson, J. L., Taylor, W. A., Bittner, M. H., & Jagim, A. R. (2024). Examining changes in ulnar collateral ligament reconstruction surgery patterns among professional baseball players. *Clinical Journal of Sport Medicine: Official Journal of the Canadian Academy of Sport Medicine*. https://doi.org/10.1097/JSM.0000000000001319

Andrade, C. (2019). Multiple Testing and Protection Against a Type 1 (False Positive) Error Using the Bonferroni and Hochberg Corrections. *Indian Journal of Psychological Medicine*, *41*(1), 99–100. https://doi.org/10.4103/IJPSYM.IJPSYM_499_18

Anthropic. (2026, April 11). Claude Sonnet 4.6 [Large language model]. https://www.anthropic.com

Anz, A. W., Bushnell, B. D., Griffin, L. P., Noonan, T. J., Torry, M. R., & Hawkins, R. J. (2010). Correlation of torque and elbow injury in professional baseball pitchers. *The American Journal of Sports Medicine*, *38*(7), 1368–1374. https://doi.org/10.1177/0363546510363402

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Badhoutiya, A., Singh, D. P., Raj, J. R. F., Srivastava, A. P., Chari, S. L., & Khan, A. K. (2023). Anomaly Detection in Healthcare: A Deep Learning Approach with Autoencoders. *2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI)*, 1–6. 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI). https://doi.org/10.1109/ICAIIHI57871.2023.10489790

Barrack, A. J., Sakurai, M., Wee, C. P., Diaz, P. R., Stocklin, C., Karduna, A. R., & Michener, L. A. (2024). Investigating the influence of modifiable physical measures on the elbow varus torque – ball velocity relationship in collegiate baseball pitchers. *Orthopaedic Journal of Sports Medicine*, *12*(11), 23259671241296496. https://doi.org/10.1177/23259671241296496

Beaudry, M. F., Beaudry, A. G., Bradley, J. P., Davis, S., Baker, B. A., Holland, G., Jacobson, B. R., & Chetlin, R. D. (2022). Retrospective analysis of ulnar collateral ligament reconstructions in Major League Baseball pitchers: A comparison of the “tall and fall” versus “drop and drive” pitching styles. *Orthopaedic Journal of Sports Medicine*, *10*(10), 23259671221128041. https://doi.org/10.1177/23259671221128041

Beaudry, M. F., Beaudry, A. G., Bradley, J. P., Haynes, D. E., Holland, G., Edwards, A., Baker, B. A., Jacobson, B. R., & Chetlin, R. D. (2023). Comparison of the “tall and fall” versus “drop and drive” pitching styles: Analysis of Major League Baseball pitchers during a single season. *Orthopaedic Journal of Sports Medicine*, *11*(5), 23259671231173691. https://doi.org/10.1177/23259671231173691

Birfer, R., Sonne, M. W., & Holmes, M. W. (2019). Manifestations of muscle fatigue in baseball pitchers: A systematic review. *PeerJ*, *7*, e7390. https://doi.org/10.7717/peerj.7390

Bland, J., & Altman, D. (1990). A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Computers in Biology and Medicine*, *20*(5), 337–340.

Buffi, J. H., Werner, K., Kepple, T., & Murray, W. M. (2015). Computing muscle, ligament, and osseous contributions to the elbow varus moment during baseball pitching. *Annals of Biomedical Engineering*, *43*(2), 404–415. https://doi.org/10.1007/s10439-014-1144-z

Bushnell, B. D., Anz, A. W., Noonan, T. J., Torry, M. R., & Hawkins, R. J. (2010). Association of maximum pitch velocity and elbow injury in professional baseball pitchers. *American Journal of Sports Medicine*, *38*(4), 728–732.

Butani, R. R., Chintalapati, K., Sridharan, A., Oltean, T. I., & Shastri, A. (2025). Comparative Machine Learning Analysis Highlights Novel Predictive Capability of Deep Neural Decision Forest for Ulnar Collateral Ligament Reconstruction in Baseball Athletes. *2025 IEEE 13th International Conference on Healthcare Informatics (ICHI)*, 1–10. https://doi.org/10.1109/ICHI64645.2025.00009

Cain, E. L., Andrews, J. R., Dugas, J. R., Wilk, K. E., McMichael, C. S., WalterII, J. C., Riley, R. S., & Arthur, S. T. (2010). Outcome of Ulnar Collateral Ligament Reconstruction of the Elbow in 1281 Athletes: Results in 743 Athletes with Minimum 2-Year Follow-up. *The American Journal of Sports Medicine*, *38*(12), 2426–2434. https://doi.org/10.1177/0363546510378100

Camp, C. L., Conte, S., D'Angelo, J., & Fealy, S. A. (2018). Epidemiology of ulnar collateral ligament reconstruction in Major and Minor League Baseball pitchers: Comprehensive report of 1429 cases. *Journal of Shoulder and Elbow Surgery*, *27*(5), 871–878. https://doi.org/10.1016/j.jse.2018.01.024

Camp, C. L., Tubbs, T. G., Fleisig, G. S., Dines, J. S., Dines, D. M., Altchek, D. W., & Dowling, B. (2017). The relationship of throwing arm mechanics and elbow varus torque: Within-subject variation for professional baseball pitchers across 82,000 throws. *The American Journal of Sports Medicine*, *45*(13), 3030–3035. https://doi.org/10.1177/0363546517719047

Chalmers, P. N., Erickson, B. J., Ball, B., Romeo, A. A., & Verma, N. N. (2016). Fastball pitch velocity helps predict ulnar collateral ligament reconstruction in Major League Baseball pitchers. *The American Journal of Sports Medicine*, *44*(8), 2130–2135. https://doi.org/10.1177/0363546516634305

Chalmers, P. N., Mcelheny, K., Dangelo, J., Ma, K., Rowe, D., & Erickson, B. (2021). Is Workload Associated with Ulnar Collateral Ligament Tears in Professional Baseball Players? An Analysis of

Days of Rest, Innings Pitched and Batters Faced (126). *Orthopaedic Journal of Sports Medicine*, *9*(10_suppl5), 2325967121S00269. https://doi.org/10.1177/2325967121S00269

Chen, F. S., Rokito, A. S., & Jobe, F. W. (2001). Medial Elbow Problems in the Overhead-Throwing Athlete. *JAAOS - Journal of the American Academy of Orthopaedic Surgeons*, *9*(2), 99.

Chen, J., Xie, B., Zhang, H., & Zhai, J. (2016). Deep Autoencoders in Pattern Recognition: A Survey. In *Bio-Inspired Computing Models and Algorithms* (pp. 229–255). WORLD SCIENTIFIC. https://doi.org/10.1142/9789813143180_0009

Chollet, F. & others. (2015). *Keras* [Computer software]. https://github.com/fchollet/keras

Cinque, M. E., LaPrade, C. M., Abrams, G. D., Sherman, S. L., Safran, M. R., & Freehill, M. T. (2022). Ulnar collateral ligament reconstruction does not decrease spin rate or performance in Major League pitchers. *The American Journal of Sports Medicine*, *50*(8), 2190–2197. https://doi.org/10.1177/03635465221097421

Cohen, S. A., Portney, D. A., Cohen, L. E., Bolia, I. K., Weber, A. E., & Saltzman, M. D. (2022). Using pitch-tracking data to identify risk factors for medial ulnar collateral ligament reconstruction in Major League Baseball pitchers. *Orthopaedic Journal of Sports Medicine*, *10*(3), 23259671211065756. https://doi.org/10.1177/23259671211065756

Conte, S. A., Camp, C. L., & Dines, J. S. (2016). Injury trends in Major League Baseball over 18 seasons: 1998-2015. *American Journal of Orthopedics*, *45*(3), 116–123.

Conte, S. A., Fleisig, G. S., Dines, J. S., Wilk, K. E., Aune, K. T., Patterson-Flynn, N., & ElAttrache, N. (2015). Prevalence of ulnar collateral ligament surgery in professional baseball players. *The American Journal of Sports Medicine*, *43*(7), 1764–1769. https://doi.org/10.1177/0363546515580792

Coughlin, R. P., Lee, Y., Horner, N. S., Simunovic, N., Cadet, E. R., & Ayeni, O. R. (2019). Increased pitch velocity and workload are common risk factors for ulnar collateral ligament injury in baseball players: A systematic review. *Journal of ISAKOS*, *4*(1), 41–47. https://doi.org/10.1136/jisakos-2018-000226

Crotin, R. L., Kozlowski, K., Horvath, P., & Ramsey, D. K. (2014). Altered Stride Length in Response to Increasing Exertion among Baseball Pitchers. *Medicine & Science in Sports & Exercise*, *46*(3), 565–571. https://doi.org/10.1249/MSS.0b013e3182a79cd9

Crotin, R. L., Slowik, J. S., Brewer, G., CainJr, E. L., & Fleisig, G. S. (2022). Determinants of Biomechanical Efficiency in Collegiate and Professional Baseball Pitchers. *The American Journal of Sports Medicine*, *50*(12), 3374–3380. https://doi.org/10.1177/03635465221119194

Davis, J. T., Limpisvasti, O., Fluhme, D., Mohr, K. J., Yocum, L. A., ElAttrache, N. S., & Jobe, F. W. (2009). The Effect of Pitching Biomechanics on the Upper Extremity in Youth and Adolescent Baseball Pitchers. *The American Journal of Sports Medicine*, *37*(8), 1484–1491. https://doi.org/10.1177/0363546509340226

DeFroda, S. F., Kriz, P. K., Hall, A. M., Zurakowski, D., & Fadale, P. D. (2016). Risk Stratification for Ulnar Collateral Ligament Injury in Major League Baseball Players: A Retrospective Study From 2007 to 2014. *Orthopaedic Journal of Sports Medicine*, *4*(2), 2325967115627126. https://doi.org/10.1177/2325967115627126

DeZee, Z. J., Barrack, A. J., Bucci, K., Zerega, R. J., Straub, R. K., Karduna, A. R., & Michener, L. A. (2025). Association between lumbopelvic stability during a single-legged step down and elbow-varus torque during baseball pitching. *Journal of Athletic Training*, *60*(2), 143–153. https://doi.org/10.4085/1062-6050-0697.23

Diffendaffer, A. Z., Bagwell, M. S., Fleisig, G. S., Yanagita, Y., Stewart, M., Cain, E. L., Dugas, J. R., & Wilk, K. E. (2022). The Clinician's Guide to Baseball Pitching Biomechanics. *Sports Health*, *15*(2), 274–281. https://doi.org/10.1177/19417381221078537

Dillon, M. R., Mastroianni, M. A., Frappa, N., Nicholson, K., LeVasseur, M. R., Luzzi, A. J., Alexander, F. J., Ablove, R., & Ahmad, C. S. (2025). Using Pitch-Tracking Metrics to Identify Warning Signs Immediately Prior to Acute Ulnar Collateral Ligament Injuries in Major League Baseball Players. *Orthopaedic Journal of Sports Medicine*, *13*(12), 23259671251389226. https://doi.org/10.1177/23259671251389226

Dong, M., Li, M., Qu, Q., Kim, Y., & Kim, S. (2025). Arm slot angles affect elbow and shoulder joint torque in elite college pitchers. *Sports Biomechanics*, *0*(0), 1–18. https://doi.org/10.1080/14763141.2024.2431927

Douoguih, W. A., Dolce, D. L., & Lincoln, A. E. (2015). Early Cocking Phase Mechanics and Upper Extremity Surgery Risk in Starting Professional Baseball Pitchers. *Orthopaedic Journal of Sports Medicine*, *3*(4), 2325967115581594. https://doi.org/10.1177/2325967115581594

Dowling, B., Hodakowski, A., Olmanson, B. A., Cohn, M. R., Pauley, P. J., Verma, N. N., Nicholson, G. P., & Garrigues, G. E. (2023). Relationship between arm path, ball velocity, and elbow varus torque in professional baseball pitchers. *Orthopaedic Journal of Sports Medicine*, *11*(12), 23259671231202524. https://doi.org/10.1177/23259671231202524

Dowling, B., Laughlin, W. A., Gurchiek, R. D., Owen, C. P., Luera, M. J., Hansen, B. R., & Fleisig, G. S. (2020). Kinematic and kinetic comparison between American and Japanese collegiate pitchers. *Journal of Science and Medicine in Sport*, *23*(12), 1202–1207. https://doi.org/10.1016/j.jsams.2020.04.013

Dowling, B., Manzi, J. E., Raab, G., Coladonato, C., Dines, J. S., & Fleisig, G. S. (2024). The relationship among lead knee extension, fastball velocity and elbow torque in professional baseball pitchers. *Sports Biomechanics*, *23*(12), 2664–2674. https://doi.org/10.1080/14763141.2022.2050801

Erickson, B. J., Chalmers, P. N., Bush-Joseph, C. A., & Romeo, A. A. (2016). Predicting and Preventing Injury in Major League Baseball. *American Journal of Orthopedics (Belle Mead, N.J.)*, *45*(3), 152–156.

Erickson, B. J., Gupta, A. K., Harris, J. D., Bush-Joseph, C., Bach, B. R., Abrams, G. D., San Juan, A. M., Cole, B. J., & Romeo, A. A. (2014). Rate of Return to Pitching and Performance After Tommy John Surgery in Major League Baseball Pitchers. *The American Journal of Sports Medicine*, *42*(3), 536–543. https://doi.org/10.1177/0363546513510890

Erickson, B. J., Harris, J. D., Chalmers, P. N., BachJr, B. R., Verma, N. N., Bush-Joseph, C. A., & Romeo, A. A. (2015). Ulnar Collateral Ligament Reconstruction: Anatomy, Indications, Techniques, and Outcomes. *Sports Health*, *7*(6), 511–517. https://doi.org/10.1177/1941738115607208

Escamilla, R. F., Barrentine, S. W., Fleisig, G. S., Zheng, N., Takada, Y., Kingsley, D., & Andrews, J. R. (2007). Pitching Biomechanics as a Pitcher Approaches Muscular Fatigue during a Simulated Baseball Game. *The American Journal of Sports Medicine*, *35*(1), 23–33. https://doi.org/10.1177/0363546506293025

Escamilla, R. F., Fleisig, G. S., Barrentine, S. W., Zheng, N., & Andrews, J. R. (1998). Kinematic Comparisons of Throwing Different Types of Baseball Pitches. *Journal of Applied Biomechanics*, *14*(1), 1–23. https://doi.org/10.1123/jab.14.1.1

Escamilla, R. F., Fleisig, G. S., Groeschner, D., & Akizuki, K. (2017). Biomechanical comparisons among fastball, slider, curveball, and changeup pitch types and between balls and strikes in professional baseball pitchers. *The American Journal of Sports Medicine*, *45*(14), 3358–3367. https://doi.org/10.1177/0363546517730052

Escamilla, R. F., Slowik, J. S., Diffendaffer, A. Z., & Fleisig, G. S. (2018). Differences among overhand, 3-quarter, and sidearm pitching biomechanics in professional baseball players. *Journal of Applied Biomechanics*. https://journals.humankinetics.com/view/journals/jab/34/5/article-p377.xml

Escamilla, R. F., Slowik, J. S., & Fleisig, G. S. (2023). Effects of contralateral trunk tilt on shoulder and elbow injury risk and pitching biomechanics in professional baseball pitchers. *The American Journal of Sports Medicine*, *51*(4), 935–941.

Fernando, T., Gammulle, H., Denman, S., Sridharan, S., & Fookes, C. (2021). *Deep Learning for Medical Anomaly Detection—A Survey* (arXiv:2012.02364). arXiv. https://doi.org/10.48550/arXiv.2012.02364

Fleisig, G. S., Andrews, J. R., Dillman, C. J., & Escamilla, R. F. (1995). Kinetics of Baseball Pitching with Implications About Injury Mechanisms. *The American Journal of Sports Medicine*, *23*(2), 233–239. https://doi.org/10.1177/036354659502300218

Fleisig, G. S., Barrentine, S. W., Zheng, N., Escamilla, R. F., & Andrews, J. R. (1999). Kinematic and kinetic comparison of baseball pitching among various levels of development. *Journal of Biomechanics*, *32*(12), 1371–1375. https://doi.org/10.1016/S0021-9290(99)00127-X

Fleisig, G. S., Chu, Y., Weber, A., & Andrews, J. (2009). Variability in baseball pitching biomechanics among various levels of competition. *Sports Biomechanics*, *8*(1), 10–21. https://doi.org/10.1080/14763140802629958

Fleisig, G. S., Diffendaffer, A. Z., Ivey, B., & Aune, K. T. (2018). Do baseball pitchers improve mechanics after biomechanical evaluations? *Sports Biomechanics*, *17*(3), 314–321. https://doi.org/10.1080/14763141.2017.1340508

Fleisig, G. S., Leddon, C. E., Laughlin, W. A., Ciccotti, M. G., Mandelbaum, B. R., Aune, K. T., Escamilla, R. F., MacLeod, T. D., & Andrews, J. R. (2015). Biomechanical performance of baseball pitchers with a history of ulnar collateral ligament reconstruction. *The American Journal of Sports Medicine*, *43*(5), 1045–1050.

Fleisig, G. S., Slowik, J. S., Wassom, D., Yanagita, Y., Bishop, J., & Diffendaffer, A. (2024). Comparison of marker-less and marker-based motion capture for baseball pitching kinematics. *Sports Biomechanics*, *23*(12), 2950–2959. https://doi.org/10.1080/14763141.2022.2076608

Fleisig, G. S., Slowik, J. S., Wychgram, C., D'Angelo, J., Chalmers, P. N., Erickson, B. J., Froom, R., Pollack Porter, K. M., & Curriero, F. C. (2025). Risk Factors for an Ulnar Collateral Ligament Injury Resulting in Surgery: A Prospective Longitudinal Study of 305 Professional Baseball Pitchers. *Orthopaedic Journal of Sports Medicine*, *13*(7), 23259671251351339. https://doi.org/10.1177/23259671251351339

Fortenbaugh, D., Fleisig, G. S., & Andrews, J. R. (2009). Baseball Pitching Biomechanics in Relation to Injury Risk and Performance. *Sports Health*, *1*(4), 314–320. https://doi.org/10.1177/1941738109338546

Garcia, G. H., Gowd, A. K., Cabarcas, B. C., Liu, J. N., Meyer, J. R., White, G. M., Romeo, A. A., & Verma, N. N. (2019). Magnetic Resonance Imaging Findings of the Asymptomatic Elbow Predict Injuries and Surgery in Major League Baseball Pitchers. *Orthopaedic Journal of Sports Medicine*, *7*(1), 2325967118818413. https://doi.org/10.1177/2325967118818413

Garg, A., Zhang, W., Samaran, J., Savitha, R., & Foo, C.-S. (2022). An Evaluation of Anomaly Detection and Diagnosis in Multivariate Time Series. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(6), 2508–2517. https://doi.org/10.1109/TNNLS.2021.3105827

Giordano, K. A., Schmitt, A., Nebel, A., Yanagita, Y., & Oliver, G. D. (2024). Normative In-Game Data for Collegiate Baseball Pitchers Using Markerless Tracking Technology. *Orthopaedic Journal of Sports Medicine*, *12*(10), 23259671241274137. https://doi.org/10.1177/23259671241274137

Gong, D., Liu, L., Le, V., Saha, B., Mansour, M. R., Venkatesh, S., & Van Den Hengel, A. (2019). Memorizing Normality to Detect Anomaly: Memory-Augmented Deep Autoencoder for Unsupervised Anomaly Detection. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 1705–1714. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). https://doi.org/10.1109/ICCV.2019.00179

Grantham, W. J., Byram, I. R., Meadows, M. C., & Ahmad, C. S. (2014). The Impact of Fatigue on the Kinematics of Collegiate Baseball Pitchers. *Orthopaedic Journal of Sports Medicine*, *2*(6), 2325967114537032. https://doi.org/10.1177/2325967114537032

Gray, R. (2010). Expert Baseball Batters Have Greater Sensitivity in Making Swing Decisions. *Research Quarterly for Exercise and Sport*, *81*(3), 373–378. https://doi.org/10.1080/02701367.2010.10599685

Guido, J. A. J., & Werner, S. L. (2012). Lower-Extremity Ground Reaction Forces in Collegiate Baseball Pitchers. *The Journal of Strength & Conditioning Research*, *26*(7), 1785. https://doi.org/10.1519/JSC.0b013e31824e1211

He, X., He, Q., & Chen, J.-S. (2021). Deep autoencoders for physics-constrained data-driven nonlinear materials modeling. *Computer Methods in Applied Mechanics and Engineering*, *385*, 114034. https://doi.org/10.1016/j.cma.2021.114034

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hodakowski, A. J., Dowling, B., Streepy, J. T., Olmanson, B., Schmitt, L., Richard, M. J., Verma, N. N., & Garrigues, G. E. (2025). Pitch types and their influence on elbow varus torque and spin rate in professional baseball pitchers. *The American Journal of Sports Medicine*, *53*(3), 543–548. https://doi.org/10.1177/03635465241309316

Hossain, M. S., Hossain, M. S., Klüttermann, S., & Müller, E. (2024). Evaluating Anomaly Detection Algorithms: A Multi-Metric Analysis Across Variable Class Imbalances. *2024 International Joint Conference on Neural Networks (IJCNN)*, 1–7. https://doi.org/10.1109/IJCNN60899.2024.10650351

Howenstein, J., Kipp, K., & Sabick, M. (2020). Peak horizontal ground reaction forces and impulse correlate with segmental energy flow in youth baseball pitchers. *Journal of Biomechanics*, *108*, 109909. https://doi.org/10.1016/j.jbiomech.2020.109909

Hurd, W. J., Jazayeri, R., Mohr, K., Limpisvasti, O., ElAttrache, N. S., & Kaufman, K. R. (2012). Pitch Velocity Is a Predictor of Medial Elbow Distraction Forces in the Uninjured High School–Aged Baseball Pitcher. *Sports Health*, *4*(5), 415–418. https://doi.org/10.1177/1941738112439695

Jakubowski, J., Stanisz, P., Bobek, S., & Nalepa, G. J. (2022). Anomaly Detection in Asset Degradation Process Using Variational Autoencoder and Explanations. *Sensors*, *22*(1), 291. https://doi.org/10.3390/s22010291

Jensen, A. R., LaPrade, M. D., Turner, T. W., Dines, J. S., & Camp, C. L. (2020). The History and Evolution of Elbow Medial Ulnar Collateral Ligament Reconstruction: From Tommy John to 2020. *Current Reviews in Musculoskeletal Medicine*, *13*(3), 349–360. https://doi.org/10.1007/s12178-020-09618-y

Kageyama, M., Sugiyama, T., Takai, Y., Kanehisa, H., & Maeda, A. (2014). Kinematic and Kinetic Profiles of Trunk and Lower Limbs during Baseball Pitching in Collegiate Pitchers. *Journal of Sports Science & Medicine*, *13*(4), 742–750.

Kang, B., Park, M., del Pobil, A. P., & Park, E. (2025). Data-driven approaches for predicting Tommy John Surgery risk in major league baseball pitchers. *Journal of Big Data*, *12*(1), 87. https://doi.org/10.1186/s40537-025-01138-1

Karnuta, J. M., Luu, B. C., Haeberle, H. S., Saluan, P. M., Frangiamore, S. J., Stearns, K. L., Farrow, L. D., Nwachukwu, B. U., Verma, N. N., Makhni, E. C., Schickendantz, M. S., & Ramkumar, P. N. (2020). Machine learning outperforms regression analysis to predict next-season Major League Baseball player injuries: Epidemiology and validation of 13,982 player-years from performance and injury profile trends, 2000-2017. *Orthopaedic Journal of Sports Medicine*, *8*(11), 2325967120963046. https://doi.org/10.1177/2325967120963046

Kato, Y., Yamada, S., & Chavez, J. (2019). Can platelet-rich plasma therapy save patients with ulnar collateral ligament tears from surgery? *Regenerative Therapy*, *10*, 123–126. https://doi.org/10.1016/j.reth.2019.02.004

Keller, R. A., Marshall, N. E., Guest, J.-M., Okoroha, K. R., Jung, E. K., & Moutzouros, V. (2016). Major League Baseball pitch velocity and pitch type associated with risk of ulnar collateral ligament injury. *Journal of Shoulder and Elbow Surgery*, *25*(4), 671–675. https://doi.org/10.1016/j.jse.2015.12.027

Keller, R. A., Steffes, M. J., Zhuo, D., Bey, M. J., & Moutzouros, V. (2014). The effects of medial ulnar collateral ligament reconstruction on Major League pitching performance. *Journal of Shoulder and Elbow Surgery*, *23*(11), 1591–1598. https://doi.org/10.1016/j.jse.2014.06.033

Kneifl, J., Rosin, D., Avci, O., Röhrle, O., & Fehr, J. (2023). Low-dimensional data-based surrogate model of a continuum-mechanical musculoskeletal system based on non-intrusive model order reduction. *Archive of Applied Mechanics*, *93*(9), 3637–3663. https://doi.org/10.1007/s00419-023-02458-5

Ko, J. U., Na, K., Oh, J.-S., Kim, J., & Youn, B. D. (2022). A new auto-encoder-based dynamic threshold to reduce false alarm rate for anomaly detection of steam turbines. *Expert Systems with Applications*, *189*, 116094. https://doi.org/10.1016/j.eswa.2021.116094

Kooima, C. L., Anderson, K., Craig, J. V., Teeter, D. M., & van Holsbeeck, M. (2004). Evidence of Subclinical Medial Collateral Ligament Injury and Posteromedial Impingement in Professional Baseball Players. *The American Journal of Sports Medicine*, *32*(7), 1602–1606. https://doi.org/10.1177/0363546503262646

Lapenda, L. V. N., Monteiro, R. P., & Bastos-Filho, C. J. A. (2020). Autoencoder latent space: An empirical study. *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2453–2460. https://doi.org/10.1109/SSCI47803.2020.9308551

Leland, D. P., Conte, S., Flynn, N., Conte, N., Crenshaw, K., Wilk, K. E., & Camp, C. L. (2019). Prevalence of medial ulnar collateral ligament surgery in 6135 current professional baseball players: A 2018 update. *Orthopaedic Journal of Sports Medicine*, *7*(9), 2325967119871442. https://doi.org/10.1177/2325967119871442

Lerch, B. G., Fleisig, G. S., Slowik, J. S., & Oliver, G. D. (2025). Variability of in-game markerless and laboratory marker-based baseball pitching biomechanics. *Journal of Biomechanics*, *188*, 112775. https://doi.org/10.1016/j.jbiomech.2025.112775

Lewis, C. L., Foch, E., Luko, M. M., Loverro, K. L., & Khuu, A. (2015). Differences in Lower Extremity and Trunk Kinematics between Single Leg Squat and Step Down Tasks. *PLOS ONE*, *10*(5), e0126258. https://doi.org/10.1371/journal.pone.0126258

Li, G., & Jung, J. J. (2023). Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges. *Information Fusion*, *91*, 93–102. https://doi.org/10.1016/j.inffus.2022.10.008

Lipa, D., Shivdasani, K., Scheidt, M., Anderson, J., Salazar, D., & Garbis, N. (2025). Major League Baseball pitchers' arm angles measured on game videos were not associated with an increased risk of ulnar collateral ligament injury. *Arthroscopy, Sports Medicine, and Rehabilitation*, *7*(1), 100979. https://doi.org/10.1016/j.asmr.2024.100979

Lizzio, V. A., Gulledge, C. M., Smith, D. G., Meldau, J. E., Borowsky, P. A., Moutzouros, V., & Makhni, E. C. (2020). Predictors of elbow torque among professional baseball pitchers. *Journal of Shoulder and Elbow Surgery*, *29*(2), 316–320. https://doi.org/10.1016/j.jse.2019.07.037

Luera, M. J., Dowling, B., Magrini, M. A., Muddle, T. W. D., Colquhoun, R. J., & Jenkins, N. D. M. (2018). Role of rotational kinematics in minimising elbow varus torques for professional versus high school pitchers. *Orthopaedic Journal of Sports Medicine*, *6*(3), 2325967118760780. https://doi.org/10.1177/2325967118760780

Major League Baseball. (2024, December 17). *MLB Injury Report 2024: Fractures, strains, Tommy John, and more*. https://www.mlb.com/injury-report

Makhni, E. C., Lee, R. W., Morrow, Z. S., Gualtieri, A. P., Gorroochurn, P., & Ahmad, C. S. (2014). Performance, return to competition, and reinjury after Tommy John surgery in Major League Baseball pitchers: A review of 147 cases. *The American Journal of Sports Medicine*, *42*(6), 1323–1332. https://doi.org/10.1177/0363546514528864

Manzi, J. E., Dowling, B., Trauger, N., Fu, M. C., Hansen, B. R., & Dines, J. S. (2022). The influence of shoulder abduction and external rotation on throwing arm kinetics in professional baseball pitchers. *Shoulder & Elbow*, *14*(1_suppl), 90–98. https://doi.org/10.1177/17585732211010300

Manzi, J. E., Dowling, B., Wang, Z., Luzzi, A., Thacher, R., Rauck, R. C., & Dines, J. S. (2022). Pitching mechanics and the relationship to accuracy in professional baseball pitchers. *The American Journal of Sports Medicine*, *50*(3), 814–822. https://doi.org/10.1177/03635465211067824

Manzi, J. E., Ruzbarsky, J. J., Krichevsky, S., Sudah, S. Y., Estrada, J., Wang, Z., Moran, J., Kunze, K. N., Ciccotti, M. C., Chen, F. R., & Dines, J. S. (2023). Kinematic and kinetic comparisons of arm slot position between high school and professional pitchers. *Orthopaedic Journal of Sports Medicine*, *11*(10), 23259671221147874. https://doi.org/10.1177/23259671221147874

Markovic, T., Moricz, S., & Leon, M. (2024). Comparative Evaluation of Autoencoders for Semi-Supervised Anomaly Detection on Univariate Time Series Data. *2024 International Conference on Machine Learning and Applications (ICMLA)*, 1321–1328. https://doi.org/10.1109/ICMLA61862.2024.00206

Mastroianni, M. A., Dillon, M. R., Frappa, N., Luzzi, A. J., Muscat, J., Chernov, D., Alexander, F. J., Nicholson, K. F., Ablove, R., & Ahmad, C. S. (2026). Pitch-Tracking Risk Factors and Warning Signs for Ulnar Collateral Ligament Injuries in Major League Baseball Pitchers. *The American Journal of Sports Medicine*, *54*(3), 694–704. https://doi.org/10.1177/03635465251411298

Mastroianni, M. A., Kunes, J. A., El-Najjar, D. B., Obana, K. K., Desai, S. S., Morrissette, C. R., Alexander, F. J., Rondon, A. J., Trofa, D. P., Popkin, C. A., Levine, W. N., & Ahmad, C. S. (2025). Advanced analytic and pitch-tracking metrics associated with UCL surgery in Major League Baseball pitchers: A case-control study. *Orthopaedic Journal of Sports Medicine*, *13*(2), 23259671241302432. https://doi.org/10.1177/23259671241302432

Mastroianni, M. A., Kunes, J. A., Mueller, J. D., Obana, K. K., Confino, J., Luzzi, A. J., Rondon, A. J., Trofa, D. P., Popkin, C. A., Jobin, C. M., Levine, W. N., & Ahmad, C. S. (2025). Pitch-Specific Advanced Analytic and Pitch-Tracking Risk Factors for Ulnar Collateral Ligament Injuries in Major League Baseball Pitchers. *The American Journal of Sports Medicine*, *53*(6), 1440–1449. https://doi.org/10.1177/03635465251330564

Mastroianni, M. A., Kunes, J., El-Najjar, D., Desai, S., Morrissette, C., Alexander, F., Rondon, A., Popkin, C., & Ahmad, C. (2024). Poster 165: Advanced Analytic and Biomechanical Risk Factors for UCL Injury in Major League Baseball Pitchers. *Orthopaedic Journal of Sports Medicine*, *12*(7_suppl2). https://doi.org/10.1177/2325967124s00134

Matsuo, T., Escamilla, R. F., Fleisig, G. S., Barrentine, S. W., & Andrews, J. R. (2001). Comparison of Kinematic and Temporal Parameters between Different Pitch Velocity Groups. *Journal of Applied Biomechanics, 17(1), 1-13.* https://doi.org/10.1123/jab.17.1.1

Matsuo, T., Fleisig, G. S., Zheng, N., & Andrews, J. R. (2006). Influence of Shoulder Abduction and Lateral Trunk Tilt on Peak Elbow Varus Torque for College Baseball Pitchers during Simulated Pitching. *Journal of Applied Biomechanics, 22(2), 93-106.* https://doi.org/10.1123/jab.22.2.93

Mayberry, J., Mullen, S., & Murayama, S. (2020). What can a jump tell us about elbow injuries in professional baseball pitchers. *The American Journal of Sports Medicine*, *48*(5), 1220–1225. https://doi.org/10.1177/0363546520905543

Mayo, B. C., Miller, A., Patetta, M. J., Schwarzman, G. R., Chen, J. W., Haden, M., Secretov, E., & Hutchinson, M. R. (2021). Preventing Tommy John Surgery: The identification of trends in pitch selection, velocity, and spin rate before ulnar collateral ligament reconstruction in Major League Baseball pitchers. *Orthopaedic Journal of Sports Medicine*, *9*(6), 23259671211012364. https://doi.org/10.1177/23259671211012364

McCutcheon, T. W., Slowik, J. S., & Fleisig, G. S. (2025). Kinematic parameters associated with elbow varus torque in elite adult baseball pitchers. *Orthopaedic Journal of Sports Medicine*, *13*(2), 23259671241300560. https://doi.org/10.1177/23259671241300560

McGraw, M. A., Kremchek, T. E., Hooks, T. R., & Papangelou, C. (2013). Biomechanical Evaluation of the Docking Plus Ulnar Collateral Ligament Reconstruction Technique Compared With the Docking Technique. *The American Journal of Sports Medicine*, *41*(2), 313–320. https://doi.org/10.1177/0363546512466375

McNally, M. P., Borstad, J. D., Oñate, J. A., & Chaudhari, A. M. W. (2015). Stride Leg Ground Reaction Forces Predict Throwing Velocity in Adult Recreational Baseball Pitchers. *The Journal of Strength & Conditioning Research*, *29*(10), 2708. https://doi.org/10.1519/JSC.0000000000000937

Meldau, J. E., Srivastava, K., Okoroha, K. R., Ahmad, C. S., Moutzouros, V., & Makhni, E. C. (2020). Cost analysis of Tommy John surgery for Major League Baseball teams. *Journal of Shoulder and Elbow Surgery*, *29*(1), 121–125. https://doi.org/10.1016/j.jse.2019.07.019

Mienye, I. D., & Swart, T. G. (2025). Deep Autoencoder Neural Networks: A Comprehensive Review and New Perspectives. *Archives of Computational Methods in Engineering*, *32*(7), 3981–4000. https://doi.org/10.1007/s11831-025-10260-5

Morrey, B. F., & An, K.-N. (1983). Articular and ligamentous contributions to the stability of the elbow joint. *The American Journal of Sports Medicine*, *11*(5), 315–319. https://doi.org/10.1177/036354658301100506

Mounayer, J., Rodriguez, S., Ghnatios, C., Farhat, C., & Chinesta, F. (2025). *Rank Reduction Autoencoders* (arXiv:2405.13980). arXiv. https://doi.org/10.48550/arXiv.2405.13980

Murray, T. A., Cook, T. D., Werner, S. L., Schlegel, T. F., & Hawkins, R. J. (2001). The Effects of Extended Play on Professional Baseball Pitchers. *The American Journal of Sports Medicine*, *29*(2), 137–142. https://doi.org/10.1177/03635465010290020501

Naito, K., Takagi, H., Yamada, N., Hashimoto, S., & Maruyama, T. (2014). Intersegmental dynamics of 3D upper arm and forearm longitudinal axis rotations during baseball pitching. *Human Movement Science*, *38*, 116–132. https://doi.org/10.1016/j.humov.2014.08.010

Nawaz, A., Khan, S. S., & Ahmad, A. (2024). Ensemble of Autoencoders for Anomaly Detection in Biomedical Data: A Narrative Review. *IEEE Access*, *12*, 17273–17289. https://doi.org/10.1109/ACCESS.2024.3360691

Ndubuaku, M. U., Anjum, A., & Liotta, A. (2019). Unsupervised Anomaly Thresholding from Reconstruction Errors. In R. Montella, A. Ciaramella, G. Fortino, A. Guerrieri, & A. Liotta (Eds.), *Internet and Distributed Computing Systems* (pp. 123–129). Springer International Publishing. https://doi.org/10.1007/978-3-030-34914-1_12

Nicholson, K. F., Collins, G. S., Waterman, B. R., & Bullock, G. S. (2022). Machine learning and statistical prediction of fastball velocity with biomechanical predictors. *Journal of Biomechanics*, *134*, 110999. https://doi.org/10.1016/j.jbiomech.2022.110999

Nicholson, K. F., Hulburt, T. C., Beck, E. C., Waterman, B. R., & Bullock, G. S. (2020). The relationship between pitch velocity and shoulder distraction force and elbow valgus torque in collegiate and high school pitchers. *Journal of Shoulder and Elbow Surgery*, *29*(12), 2661–2667. https://doi.org/10.1016/j.jse.2020.04.046

Oeding, J. F., Boos, A. M., Kalk, J. R., Sorenson, D., Verhooven, F. M., Moatshe, G., & Camp, C. L. (2024). Pitch-Tracking Metrics as a Predictor of Future Shoulder and Elbow Injuries in Major League Baseball Pitchers: A Machine-Learning and Game-Theory Based Analysis. *Orthopaedic Journal of Sports Medicine*, *12*(8), 23259671241264260. https://doi.org/10.1177/23259671241264260

Oi, T., Yoshiya, S., Slowik, J., Diffendaffer, A., Takagi, Y., Tanaka, H., Nobuhara, K., & Fleisig, G. S. (2019). Biomechanical differences between Japanese and American professional baseball pitchers. *Orthopaedic Journal of Sports Medicine*, *7*(2), 2325967119825625. https://doi.org/10.1177/2325967119825625

Oliver, G. D., & Keeley, D. W. (2010). Gluteal Muscle Group Activation and its Relationship With Pelvis and Torso Kinematics in High-School Baseball Pitchers. *The Journal of Strength & Conditioning Research*, *24*(11), 3015. https://doi.org/10.1519/JSC.0b013e3181c865ce

Oyama, S., Yu, B., Blackburn, J. T., Padua, D. A., Li, L., & Myers, J. B. (2013). Effect of Excessive Contralateral Trunk Tilt on Pitching Biomechanics and Performance in High School Baseball Pitchers. *The American Journal of Sports Medicine*, *41*(10), 2430–2438. https://doi.org/10.1177/0363546513496547

Oyama, S., Yu, B., Blackburn, J. T., Padua, D. A., Li, L., & Myers, J. B. (2014). Improper Trunk Rotation Sequence Is Associated With Increased Maximal Shoulder External Rotation Angle and Shoulder Joint Force in High School Baseball Pitchers. *The American Journal of Sports Medicine*, *42*(9), 2089–2094. https://doi.org/10.1177/0363546514536871

Pappas, A. M., Zawacki, R. M., & Sullivan, T. J. (1985). Biomechanics of baseball pitching. A preliminary report. *The American Journal of Sports Medicine*, *13*(4), 216–222. https://doi.org/10.1177/036354658501300402

Pereira, J., & Silveira, M. (2019). Learning Representations from Healthcare Time Series Data for Unsupervised Anomaly Detection. *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 1–7. https://doi.org/10.1109/BIGCOMP.2019.8679157

Petti, B., & Gilani, S. (2024). *baseballr: Acquiring and Analyzing Baseball Data* [Computer software]. https://github.com/BillPetti/baseballr

Pham, C.-H., Ladjal, S., & Newson, A. (2022). PCA-AE: Principal Component Analysis Autoencoder for Organising the Latent Space of Generative Networks. *Journal of Mathematical Imaging and Vision*, *64*(5), 569–585. https://doi.org/10.1007/s10851-022-01077-z

Piergiovanni, A. J., & Ryoo, M. S. (2019). Early Detection of Injuries in MLB Pitchers from Video. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2431–2438. https://doi.org/10.1109/CVPRW.2019.00298

Pinaya, W. H. L., Vieira, S., Garcia-Dias, R., & Mechelli, A. (2020). Chapter 11—Autoencoders. In A. Mechelli & S. Vieira (Eds.), *Machine Learning* (pp. 193–208). Academic Press. https://doi.org/10.1016/B978-0-12-815739-8.00011-0

Platt, B. N., Zacharias, A. V., Conley, C., Hockensmith, L., Stockwell, N., Sciascia, A., & Stone, A. V. (2021). Association between pitch break on the 4-seam fastball and slider and shoulder injury in Major League Baseball pitchers: A case-control study. *Orthopaedic Journal of Sports Medicine*, *9*(10), 23259671211038961. https://doi.org/10.1177/23259671211038961

Podesta, L., Crow, S. A., Volkmer, D., Bert, T., & Yocum, L. A. (2013). Treatment of Partial Ulnar Collateral Ligament Tears in the Elbow With Platelet-Rich Plasma. *The American Journal of Sports Medicine*, *41*(7), 1689–1694. https://doi.org/10.1177/0363546513487979

Portney, D. A., Buchler, L. T., Lazaroff, J. M., Gryzlo, S. M., & Saltzman, M. D. (2019). Influence of pitching release location on ulnar collateral ligament reconstruction risk among Major League Baseball pitchers. *Orthopaedic Journal of Sports Medicine*, *7*(2), 2325967119826540. https://doi.org/10.1177/2325967119826540

Post, E. G., Laudner, K. G., McLoda, T. A., Wong, R., & Meister, K. (2015). Correlation of shoulder and elbow kinetics with ball velocity in collegiate baseball pitchers. *Journal of Athletic Training*, *50*(6), 629–633.

Pota, M., De Pietro, G., & Esposito, M. (2023). Real-time anomaly detection on time series of industrial furnaces: A comparison of autoencoder architectures. *Engineering Applications of Artificial Intelligence*, *124*, 106597. https://doi.org/10.1016/j.engappai.2023.106597

Prodromo, J., Patel, N., Kumar, N., Denehy, K., Tabb, L. P., & Tom, J. (2016). Pitch characteristics before ulnar collateral ligament reconstruction in Major League pitchers compared with age-matched controls. *Orthopaedic Journal of Sports Medicine*, *4*(6), 2325967116653946. https://doi.org/10.1177/2325967116653946

Putnam, C. A. (1993). Sequential motions of body segments in striking and throwing skills: Descriptions and explanations. *Journal of Biomechanics, Proceedings of the XIIIth Congress of the International Society of Biomechanics*, *26*, 125–135. https://doi.org/10.1016/0021-9290(93)90084-R

Qiao, M., Crotin, R. L., & Szymanski, D. J. (2025). An inferential investigation into countermovement jump determinants of ulnar collateral ligament injuries in collegiate baseball pitchers. *The American Journal of Sports Medicine*, *53*(5), 1202–1209. https://doi.org/10.1177/03635465251322913

Quan, H., & Shih, W. J. (1996). Assessing Reproducibility by the Within-Subject Coefficient of Variation with Random Effects Models. *Biometrics*, *52*(4), 1195–1203. JSTOR. https://doi.org/10.2307/2532835

R Core Team. (2024). *R: A language and environment for statistical computing* (Version 4.4.2 (Pile of Leaves)) [Computer software]. R Foundation for Statistical Computing. https://www.r-project.org/

Roach, N. T., Venkadesan, M., Rainbow, M. J., & Lieberman, D. E. (2013). Elastic energy storage in the shoulder and the evolution of high-speed throwing in Homo. *Nature*, *498*(7455), 483–486. https://doi.org/10.1038/nature12267

Roegele, J. (2012). *Tommy John surgery list* [Dataset]. https://docs.google.com/spreadsheets/u/0/d/1gQujXQQGOVNaiuwSN680HqFDVsCwvN-3AazykOBON0/htmlview.

Sak, H., Senior, A., & Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling. *Proceedings of Interspeech 2014, pp. 338–342.* https://doi.org/10.21437/Interspeech.2014-80

Sakurai, M., Barrack, A. J., Lobb, N. J., Wee, C. P., Diaz, P. R., Michener, L. A., & Karduna, A. R. (2024). Collegiate baseball pitchers demonstrate a relationship between ball velocity and elbow varus torque, both within and across pitchers. *Sports Biomechanics*, *23*(12), 3103–3111.

Slowik, J. S., Aune, K. T., Diffendaffer, A. Z., Cain, E. L., Dugas, J. R., & Fleisig, G. S. (2019). Fastball velocity and elbow-varus torque in professional baseball pitchers. *Journal of Athletic Training*, *54*(3), 296–301. https://doi.org/10.4085/1062-6050-558-17

Slowik, J. S., & Fleisig, G. S. (2019). The influence of age-related factors on the varus/valgus moment capacity of the elbow muscles of baseball pitchers. *ISBS Proceedings Archive*, *37*(1), 200.

Solomito, M. J., Garibay, E. J., Cohen, A., & Nissen, C. W. (2022). Lead knee flexion angle is associated with both ball velocity and upper extremity joint moments in collegiate baseball pitchers. *Sports Biomechanics*, *23*(12), 2626–2636. https://doi.org/10.1080/14763141.2022.2046143

Stodden, D. F., Fleisig, G. S., McLean, S. P., & Andrews, J. R. (2005). Relationship of Biomechanical Factors to Baseball Pitching Velocity: Within Pitcher Variation. *Journal of Applied Biomechanics, 21(1), 44-45.* https://doi.org/10.1123/jab.21.1.44

Stone, B. L., & Schilling, B. K. (2020). Neuromuscular Fatigue in Pitchers Across a Collegiate Baseball Season. *The Journal of Strength & Conditioning Research*, *34*(7), 1937. https://doi.org/10.1519/JSC.0000000000003663

Tanaka, Y., Ishida, T., Ino, T., Suzumori, Y., Samukawa, M., Kasahara, S., & Tohyama, H. (2024). The effects of relative trunk rotation velocity on ball speed and elbow and shoulder joint torques during baseball pitching. *Sports Biomechanics*, *23*(12), 3551–3563. https://doi.org/10.1080/14763141.2022.2129431

Udall, J. H., Fitzpatrick, M. J., McGarry, M. H., Leba, T.-B., & Lee, T. Q. (2009). Effects of flexor-pronator muscle loading on valgus stability of the elbow with an intact, stretched, and resected medial ulnar collateral ligament. *Journal of Shoulder and Elbow Surgery*, *18*(5), 773–778. https://doi.org/10.1016/j.jse.2009.03.008

Way, G. P., Zietz, M., Rubinetti, V., Himmelstein, D. S., & Greene, C. S. (2020). Compressing gene expression data using multiple latent space dimensionalities learns complementary biological representations. *Genome Biology*, *21*(1), 109. https://doi.org/10.1186/s13059-020-02021-3

Weng, Y.-H., Chang, P.-H., Wu, K.-P., Lin, J.-J., & Huang, T.-S. (2025). Enhanced personalized prediction of baseball-related upper extremity injuries through novel features and explainable artificial intelligence. *Journal of Sports Sciences*, *43*(7), 719–727. https://doi.org/10.1080/02640414.2025.2474328

Werner, S. L., Fleisig, G. S., Dillman, C. J., & Andrews, J. R. (1993). Biomechanics of the Elbow During Baseball Pitching. *Journal of Orthopaedic & Sports Physical Therapy*, *17*(6), 274–278. https://doi.org/10.2519/jospt.1993.17.6.274

Werner, S. L., Murray, T. A., Hawkins, R. J., & Gill, T. J. (2002). Relationship between throwing mechanics and elbow valgus in professional baseball pitchers. *Journal of Shoulder and Elbow Surgery*, *11*(2), 151–155. https://doi.org/10.1067/mse.2002.121481

Whiteley, R. (2007). Baseball Throwing Mechanics as They Relate to Pathology and Performance—A Review. *Journal of Sports Science & Medicine*, *6*(1), 1–20.

Whiteside, D., Martini, D. N., Lepley, A. S., Zernicke, R. F., & Goulet, G. C. (2016). Predictors of ulnar collateral ligament reconstruction in Major League Baseball pitchers. *The American Journal of Sports Medicine*, *44*(9), 2202–2209. https://doi.org/10.1177/0363546516643812

Wilk, K. E., Macrina, L. C., & Arrigo, C. (2012). Passive Range of Motion Characteristics in the Overhead Baseball Pitcher and Their Implications for Rehabilitation. *Clinical Orthopaedics and Related Research*, *470*(6), 1586. https://doi.org/10.1007/s11999-012-2265-z

Yanagisawa, O. (2024). Alterations in pitching biomechanics and performance with an increasing number of pitches in baseball pitchers: A narrative review. *PM&R*, *16*(6), 632–643. https://doi.org/10.1002/pmrj.13054

Yanai, T., Onuma, K., & Nagami, T. (2025). Varus Strength of the Medial Elbow Musculature for Stress Shielding of the Ulnar Collateral Ligament in Competitive Baseball Pitchers. *Medicine & Science in Sports & Exercise*, *57*(4), 791. https://doi.org/10.1249/MSS.0000000000003614

Appendices

Appendix A. R Code Excerpts — Autoencoder-Based Anomaly Detection

Core code excerpts from the autoencoder-based anomaly detection pipeline are presented below, organised by functional component. For each pitcher, a vanilla AE is trained on baseline data (200–600 days before the last game appearance) and applied to the detection window (0–200 days before the last game appearance) to compute per-game reconstruction errors. Anomaly thresholds are derived empirically from each pitcher's baseline reconstruction error distribution at five percentile levels. The complete script and full analysis pipeline are available at https://github.com/ozakiryotaro/uclr-biomechanical-anomaly-detection.

Hyperparameter Settings

```
train_ratio           <- 0.75
latent_dim            <- 12
learning_rate         <- 0.001
batch_size            <- 4
dropout_rate          <- 0.1
threshold_percentiles <- c(90, 92.5, 95, 97.5, 99)
```

Autoencoder Architecture

```
build_autoencoder <- function(n_features, latent_dim, learning_rate,
dropout_rate) {

  input   <- layer_input(shape = n_features)
  encoded <- input %>%
    layer_dense(units = 128, activation = "relu") %>%
    layer_dropout(rate = dropout_rate) %>%
    layer_dense(units = 64,  activation = "relu") %>%
    layer_dropout(rate = dropout_rate)

  latent <- encoded %>%
    layer_dense(units = latent_dim, activation = "relu")

  decoded <- latent %>%
    layer_dense(units = 64,  activation = "relu") %>%
    layer_dropout(rate = dropout_rate) %>%
    layer_dense(units = 128, activation = "relu") %>%
    layer_dropout(rate = dropout_rate)

  output <- decoded %>%
    layer_dense(units = n_features, activation = "linear")

  model <- keras_model(input, output)
  model %>% compile(
    optimizer = optimizer_adam(learning_rate = learning_rate),
    loss      = "mse"
  )
```

```
  return(model)
}
```

Zero-Handling for Missing Pitch Types

```
apply_zero_handling <- function(scaled_matrix, raw_matrix, train_means,
pitch_cols) {
  for (col in pitch_cols) {
    zero_idx <- raw_matrix[, col] == 0 & train_means[col] != 0
    scaled_matrix[zero_idx, col] <- 0
  }
  return(scaled_matrix)
}
```

Reconstruction Error Functions

```
compute_reconstruction_error <- function(model, scaled_matrix) {
  preds <- model %>% predict(scaled_matrix, verbose = 0)
  rowMeans((scaled_matrix - preds)^2)
}

compute_feature_errors <- function(model, scaled_matrix) {
  preds <- model %>% predict(scaled_matrix, verbose = 0)
  err   <- (scaled_matrix - preds)^2
  as_tibble(err, .name_repair = "minimal")
}
```

Data Preparation: Variance Filtering, Scaling, and Zero Handling

```
# Remove zero-variance columns from training data only
train_vars <- apply(train_matrix, 2, var, na.rm = TRUE)
keep_cols  <- !is.na(train_vars) & train_vars > 1e-6

train_matrix     <- train_matrix[,     keep_cols, drop = FALSE]
val_matrix       <- val_matrix[,       keep_cols, drop = FALSE]
detection_matrix <- detection_matrix[, keep_cols, drop = FALSE]

# Scale using training statistics only
train_means <- colMeans(train_matrix, na.rm = TRUE)
train_sds   <- apply(train_matrix, 2, sd, na.rm = TRUE)
train_sds[train_sds == 0] <- 1

train_scaled     <- scale(train_matrix,     center = train_means, scale =
train_sds)
val_scaled       <- scale(val_matrix,       center = train_means, scale =
train_sds)
detection_scaled <- scale(detection_matrix, center = train_means, scale =
train_sds)

train_scaled[is.na(train_scaled)]         <- 0
val_scaled[is.na(val_scaled)]             <- 0
detection_scaled[is.na(detection_scaled)] <- 0

# Apply zero handling to all three splits
pitch_cols <- grep("^(FF|SI|SL|CH|CU|FC)_", colnames(train_matrix))
```

```
train_scaled     <- apply_zero_handling(train_scaled,     train_matrix,
train_means, pitch_cols)
val_scaled       <- apply_zero_handling(val_scaled,       val_matrix,
train_means, pitch_cols)
detection_scaled <- apply_zero_handling(detection_scaled,
detection_matrix, train_means, pitch_cols)
```

Model Training

```
history <- model %>% fit(
  x               = train_scaled,
  y               = train_scaled,
  epochs          = 200,
  batch_size      = batch_size,
  validation_data = list(val_scaled, val_scaled),
  callbacks       = list(
    callback_early_stopping(
      monitor              = "val_loss",
      patience             = 20,
      restore_best_weights = TRUE
    )
  ),
  verbose = 0
)
```

Threshold Computation and Anomaly Flagging

```
# Derive percentile thresholds from baseline reconstruction errors
thresholds <- quantile(baseline_errors, probs = threshold_percentiles /
100)

# Flag detection window games at each threshold
for (i in seq_along(threshold_percentiles)) {
  col_name <- paste0("anomaly_p", threshold_percentiles[i])
  detection_results[[col_name]] <- detection_results$reconstruction_error
> thresholds[i]
}
```

Appendix B. R Code Excerpts — Hyperparameter Optimisation via Grid Search

Core code excerpts from the two-stage grid search used to identify optimal AE hyperparameters are presented below, organised by functional component. An initial pilot across five pitchers evaluated 243 configurations and fixed the encoder architecture (128-64) and dropout rate (0.1). The reduced 27-configuration grid was then applied across 34 representative pitchers, processed in batches of ten to manage computation time. The optimal configuration was selected by mean validation loss across pitchers.

Reduced Grid Definition

```
grid <- expand.grid(
  latent_dim    = c(4, 8, 12),
  encoder_type  = c("128_64"),
  learning_rate = c(0.001, 0.0005, 0.0001),
  batch_size    = c(4, 8, 16),
  dropout_rate  = c(0.1),
  stringsAsFactors = FALSE
)
```

Architecture Layer Configuration

```
get_layers <- function(encoder_type) {
  switch(encoder_type,
         "128_64"    = list(encoder = c(128, 64),     decoder = c(64,
128)),
         "64_32"     = list(encoder = c(64, 32),      decoder = c(32,
64)),
         "128_64_32" = list(encoder = c(128, 64, 32), decoder = c(32, 64,
128))
  )
}
```

Autoencoder Construction

```
build_autoencoder <- function(n_features, latent_dim, encoder_type,
                              learning_rate, dropout_rate) {

  layers  <- get_layers(encoder_type)
  input   <- layer_input(shape = n_features)
  encoded <- input

  for (units in layers$encoder) {
    encoded <- encoded %>%
      layer_dense(units = units, activation = "relu") %>%
      layer_dropout(rate = dropout_rate)
  }

  latent  <- encoded %>%
```

```
    layer_dense(units = latent_dim, activation = "relu")
  decoded <- latent

  for (units in layers$decoder) {
    decoded <- decoded %>%
      layer_dense(units = units, activation = "relu") %>%
      layer_dropout(rate = dropout_rate)
  }

  output <- decoded %>%
    layer_dense(units = n_features, activation = "linear")

  model <- keras_model(input, output)
  model %>% compile(
    optimizer = optimizer_adam(learning_rate = learning_rate),
    loss      = "mse"
  )

  return(model)
}
```

Data Preparation: Scaling and Zero Handling

```
# Chronological 75/25 split
split_point  <- floor(n_games * train_ratio)
train_matrix <- data_matrix[1:split_point, , drop = FALSE]
val_matrix   <- data_matrix[(split_point + 1):n_games, , drop = FALSE]

# Remove zero-variance columns from training data only
train_vars   <- apply(train_matrix, 2, var, na.rm = TRUE)
keep_cols    <- train_vars > 1e-6
train_matrix <- train_matrix[, keep_cols, drop = FALSE]
val_matrix   <- val_matrix[,   keep_cols, drop = FALSE]

# Scale using training statistics only
train_means  <- colMeans(train_matrix)
train_sds    <- apply(train_matrix, 2, sd)
train_sds[train_sds == 0] <- 1

train_scaled <- scale(train_matrix, center = train_means, scale =
train_sds)
val_scaled   <- scale(val_matrix,   center = train_means, scale =
train_sds)

train_scaled[is.na(train_scaled)] <- 0
val_scaled[is.na(val_scaled)]     <- 0

# Zero handling for pitch-type columns
pitch_cols <- grep("^(FF|SI|SL|CH|CU|FC)_", colnames(train_matrix))
for (col in pitch_cols) {
  zero_train <- train_matrix[, col] == 0 & train_means[col] != 0
  train_scaled[zero_train, col] <- 0
  zero_val <- val_matrix[, col] == 0 & train_means[col] != 0
  val_scaled[zero_val, col] <- 0
}
```

Model Training

```
history <- model %>% fit(
  x               = train_scaled,
  y               = train_scaled,
  epochs          = 200,
  batch_size      = grid$batch_size[i],
  validation_data = list(val_scaled, val_scaled),
  callbacks       = list(
    callback_early_stopping(
      monitor              = "val_loss",
      patience             = 20,
      restore_best_weights = TRUE
    )
  ),
  verbose = 0
)
```

Configuration Summary and Optimal Selection

```
config_summary <- all_results %>%
  filter(status == "completed") %>%
  group_by(config, latent_dim, encoder_type, learning_rate,
           batch_size, dropout_rate) %>%
  summarise(
    n_pitchers_completed = n(),
    mean_val_loss        = mean(val_loss,    na.rm = TRUE),
    sd_val_loss          = sd(val_loss,      na.rm = TRUE),
    mean_train_loss      = mean(train_loss,  na.rm = TRUE),
    mean_overfit_gap     = mean(overfit_gap, na.rm = TRUE),
    median_val_loss      = median(val_loss,  na.rm = TRUE),
    mean_epochs          = mean(epochs_run,  na.rm = TRUE),
    .groups = "drop"
  ) %>%
  arrange(mean_val_loss)
```

Appendix C. Propensity Score Matching — Covariate Balance Diagnostics

Covariate balance diagnostics for the propensity score-matched control group are reported below. Pre-matching values reflect all 46 UCLR pitchers and the full candidate control pool of 580 eligible pitchers. Post-matching values reflect the final matched sample of 45 pairs. One UCLR pitcher was excluded due to insufficient active controls meeting minimum appearance thresholds, reducing the matched sample from 46 to 45 pairs. All post-matching standardised mean differences (SMD) fall below 0.10, indicating negligible imbalance (Austin, 2011). The marginal increase in age SMD from pre- (0.142) to post-matching (0.174) is an artefact of the one excluded UCLR pitcher and does not indicate a matching failure. The value warrants acknowledgement given it is the only covariate where matching did not improve balance.

Table 0.1 Covariate balance before and after propensity score matching

	Pre-matching			**Post-matching**		
Covariate	UCLR (n = 46)	Control pool (n = 580)	SMD	UCLR (n = 45)	Matched controls (n = 45)	SMD
Age at reference (years)	29.1 ± 3.1	29.5 ± 3.2	0.142	29.0 ± 3.0	29.5 ± 3.3	0.174
Right-handed (%)	71.7	72.6	0.019	73.3	71.1	0.049
Mean PV (kph)	142.8 ± 5.4	143.4 ± 4.8	0.129	142.9 ± 5.3	143.2 ± 6.2	0.047
Cumulative pitch count	1696 ± 902	1514 ± 787	0.215	1705 ± 910	1679 ± 1102	0.025

Note. Values represent mean ± SD unless otherwise indicated. UCLR = ulnar collateral ligament reconstruction. SMD = standardised mean difference, calculated as the absolute difference in group means divided by the pooled standard deviation. kph = kilometres per hour. PV = pitch velocity. Right-handed (%) represents the proportion of right-handed pitchers in each group. Pre-matching control pool comprises all candidate pitchers meeting minimum appearance thresholds (≥20 baseline games, ≥5 detection window games) anchored to each UCLR pitcher's reference date, deduplicated to one row per pitcher with covariates averaged across eligible reference dates.

Appendix D. UCLR Cohort — Pitcher Characteristics and Data Availability

Characteristics and data availability for all 46 MLB pitchers in the analytical sample are listed below, ordered by surgery year. Reference date corresponds to each pitcher's last game appearance before UCLR surgery. Baseline pitch count reflects total pitches thrown during the baseline training window (200–600 days before the reference date). Detection window games reflects the number of game appearances in the detection window (0–200 days before the reference date). One pitcher (Zach Duke) was excluded from the propensity score-matched control comparison due to insufficient active controls meeting minimum appearance thresholds within his baseline window.

Table 0.2 Ulnar collateral ligament cohort pitcher characteristics and data availability (n = 46)

Pitcher	Throws	Surgery year	Baseline games	Baseline pitch count	Detection window games
Zach Duke	L	2016	73	1313	80
Drew Storen	R	2017	57	1004	58
Grant Dayton	L	2017	34	697	29
Michael Pineda	R	2017	33	3250	17
Zach Putnam	R	2017	25	517	7
Brent Suter	L	2018	23	1509	20
Chad Kuhl	R	2018	32	2984	16
Johnny Cueto	R	2018	25	2703	9
Jordan Montgomery	L	2018	30	2743	6
Kevin Shackelford	R	2018	26	569	5
Keynan Middleton	R	2018	64	1122	16
Lance McCullers Jr.	R	2018	27	2424	30
Taijuan Walker	R	2018	27	2855	5
Adam Warren	R	2019	49	1197	25
Chasen Bradford	R	2019	46	973	13
David Robertson	R	2019	69	1358	9
Jordan Hicks	R	2019	73	1541	28
Steven Wright	R	2019	21	931	6
Tim Mayza	L	2019	38	718	68
Dakota Hudson	R	2020	37	3351	8
Jalen Beeks	L	2020	34	1982	12
Luis Perdomo	R	2020	48	1224	10
Evan Marshall	R	2021	27	532	27
Andrew Kittredge	R	2022	60	1236	17
Anthony Bender	R	2022	63	1232	22
Chad Green	R	2022	71	1636	14
Justin Wilson	L	2022	44	749	5
Tanner Rainey	R	2022	42	865	29

Tyler Matzek	L	2022	83	1527	42
Vladimir Gutierrez	R	2022	22	2043	10
Jeffrey Springs	L	2023	31	2171	5
Johan Oviedo	R	2023	21	1073	32
Liam Hendriks	R	2023	60	1153	5
Luis Garcia	R	2023	28	2624	6
Penn Murfee	R	2023	66	1206	16
Reiver Sanmartin	L	2023	45	1173	14
Robbie Ray	L	2023	29	2926	5
Sandy Alcantara	R	2023	33	3457	28
Tony Gonsolin	R	2023	26	2202	20
Tyler Mahle	R	2023	24	2312	5
Brooks Raley	L	2024	70	1299	8
Cristian Javier	R	2024	34	3367	8
Daniel Duarte	R	2024	29	631	6
Kyle Bradish	R	2024	31	2972	8
Luis Medina	R	2024	23	2155	8
Ray Kerr	L	2024	22	511	10

Note. UCLR = ulnar collateral ligament reconstruction. Reference date = last game appearance before surgery. Baseline games and baseline pitch count reflect the baseline training window (200–600 days before the reference date). Detection window games = number of game appearances in the detection window (0–200 days before the reference date). Eligibility thresholds: ≥20 baseline games and ≥5 detection window games. One pitcher (Zach Duke) was excluded from the propensity score-matched control comparison due to insufficient active controls meeting minimum appearance thresholds within his baseline window.