Neuromuscular Efficiency and Elbow Varus Torque Production in Baseball Pitchers: A Two-Tier Machine Learning Approach Integrating Electromyography and Biomechanical Analysis

Neuromuscular Efficiency and Elbow Varus Torque Production in Baseball Pitchers: A Two-Tier Machine Learning Approach Integrating Electromyography and Biomechanical Analysis

Elbow varus torque represents the primary mechanical load associated with ulnar collateral ligament injury in baseball pitchers. Although biomechanical models predict torque from kinematics and anthropometrics, athletes with similar movement patterns demonstrate substantial variance in torque production and injury susceptibility, suggesting that unmeasured neuromuscular factors contribute to joint loading. This study developed and validated a two-tier machine learning framework to quantify the independent contribution of forearm muscle activation patterns to elbow varus torque beyond predictions achievable through biomechanical kinematics alone. A hierarchical modelling approach analysed 20,323 pitching trials from motion capture systems. Stage 1 employed gradient boosted decision trees to predict torque from anthropometric and kinematic features, establishing a biomechanical baseline ($R^2 = 0.764$, RMSE = 18.55 Nm). Stage 2 modelled residuals from Stage 1 using surface electromyography of the flexor carpi radialis and flexor carpi ulnaris muscles, combined with Flex Pro Grip isometric performance metrics, in 581 trials with complete neuromuscular data. Novel electromyography-to-performance ratio features were engineered to capture relative muscle activation efficiency. Stage 2 explained approximately 30% of residual variance ($R^2 = 0.298–0.316$, RMSE = 16.44 Nm). Frequency-domain and efficiency ratio features demonstrated superior predictive power over raw amplitude measures. Forearm muscle activation patterns explain clinically meaningful variance in elbow torque production beyond biomechanical kinematics, supporting a neuromuscular efficiency construct that may inform personalised injury risk stratification and training interventions.

Keywords: electromyography, machine learning, elbow biomechanics, ulnar collateral ligament, baseball pitching

1. Introduction

1.1. Clinical Context and Epidemiological Burden

Ulnar collateral ligament (UCL) injuries represent one of the most prevalent and career-threatening conditions affecting baseball pitchers, with UCL reconstruction surgery (also known as 'Tommy John surgery') rates continuing to increase at both professional and youth levels (Erickson et al., 2015; Keri et al., 2018). Elbow varus torque, defined as the frontal plane moment tending to open the medial elbow joint, constitutes the primary mechanical load experienced by the UCL during the throwing motion.

[Figure 1 near here]

Peak varus torque values range from 40 to 120 Newton-meters (Nm) during high-velocity pitching, with magnitudes approaching or exceeding the estimated failure load of the native UCL in elite athletes (Fleisig et al., 1995; Escamilla et al., 2007). Despite advances in motion capture technology enabling precise quantification of joint loading through inverse dynamics calculations, two fundamental questions remain not fully answered. First, why do athletes with similar kinematic profiles demonstrate substantial individual variance in torque production? Second, what physiological mechanisms explain why some pitchers tolerate high torque loads without injury while others develop UCL pathology at comparatively lower loading magnitudes? These questions suggest that factors beyond observable kinematics may influence joint loading and injury risk.

1.2. Limitations of Kinematic-Based Torque Prediction

Contemporary biomechanical models predict elbow varus torque from observable kinematics, including joint angles, angular velocities, and segment accelerations, combined with anthropometric parameters such as body mass, limb segment lengths, and estimated segment

masses (Aguinaldo & Chambers, 2009; Anz et al., 2010). These models consistently explain 70% to 80% of torque variance across diverse pitcher populations (Whiteside et al., 2016), demonstrating that movement patterns represent the primary determinants of joint loading. The remaining 20% to 30% of unexplained variance is often attributed to measurement error or individual biological variations, however, this has not been completely confirmed. The central nervous system faces what Bernstein (1967) termed the 'degrees of freedom problem' during multi-joint movements: infinite combinations of muscle activation patterns can produce identical kinematics. This implies that two pitchers executing identical joint trajectories may employ fundamentally different muscle recruitment strategies, potentially resulting in different joint loading patterns despite kinematic similarity. Electromyographic studies in overhead athletes have demonstrated substantial variability in muscle activation timing, magnitude, and coactivation patterns despite similar task performance across athletes (Escamilla et al., 2017; Seroyer et al., 2019). However, most EMG research in pitching has focused on shoulder musculature, with limited investigation of forearm muscles despite their anatomical proximity and mechanical coupling to the medial elbow complex.

1.3. Neuromuscular Control and the Flexor-Pronator Mass

The flexor-pronator muscle mass, particularly the flexor carpi radialis (FCR) and flexor carpi ulnaris (FCU), originates from the medial epicondyle of the humerus in close anatomical relationship to the UCL (DiGiovine et al., 1992). Beyond their primary roles in wrist flexion and radial or ulnar deviation, these muscles contribute to dynamic joint stability through compressive loading across the elbow joint, therefore modulating valgus stress experienced by the UCL (Davidson et al., 1995; Park & Ahmad, 2004). Biomechanical modeling studies suggest that appropriate flexor-pronator muscle activation could reduce UCL strain by 10% to 15% through increased joint compression and altered force distribution (Davidson et al.,

1995). Recent theoretical work proposes that optimal forearm muscle coordination serves a protective function by reducing reliance on passive ligamentous restraint, while inefficient or compensatory activation patterns may increase injury risk through multiple pathways (Watson et al., 2020; Andrews & Wilk, 2021). These pathways include increased metabolic demand leading to premature fatigue, altered joint mechanics affecting stress distribution, and chronic overload of the common flexor tendon insertion. However, direct evidence linking muscle activation patterns to joint loading beyond what kinematics predict remains limited.

1.4. The Neuromuscular Efficiency Hypothesis

This study proposes a neuromuscular efficiency framework to explain residual torque variance: athletes who produce less-than-expected torque for their observed kinematics demonstrate efficient muscle activation patterns that minimise unnecessary co-contraction and optimise force transfer, while those exceeding kinematic predictions exhibit compensatory strategies requiring excessive muscle effort for equivalent performance outcomes. This hypothesis grounds itself in the motor control concept of 'motor abundance' (Latash, 2012), which posits that skilled performers exploit the redundant solution space to minimise metabolic cost, fatigue accumulation, and tissue stress. Critically, this efficiency construct cannot be assessed through kinematics alone, as observable movement trajectories represent only the kinematic output of underlying neuromuscular processes. Surface electromyography provides the necessary window into these underlying neuromuscular strategies.

1.5. Rationale for Hierarchical Two-Tier Modeling

Traditional single-stage regression approaches combining biomechanical and EMG features suffer from a fundamental statistical problem: dominant kinematic predictors (e.g., peak shoulder internal rotation velocity, elbow flexion angle at ball release) explain such large

proportions of variance that EMG features contribute minimally to model fit, often resulting in their exclusion during feature selection or assignment of negligible importance weights. This statistical dominance does not imply that EMG features lack physiological importance. Rather, their contribution becomes masked because of the overwhelming kinematic signal. A hierarchical two-tier architecture addresses this limitation by explicitly separating the modeling task into two stages. Stage 1 captures the primary kinematic-anthropometric signal, establishing what an athlete should produce given their physical attributes and observed movement patterns, functioning as a 'forearm muscle-agnostic' prediction. Stage 2 then models deviations from this expectation using neuromuscular features, thereby isolating the component of torque production attributable to muscle coordination efficiency. This architecture ensures that EMG features are evaluated based on their ability to explain variance that kinematic and anthropometric data fundamentally cannot capture, providing a fair test of the neuromuscular efficiency hypothesis. This approach parallels 'above expected' modelling frameworks common in sports analytics (e.g., predicting home run rates above expectation based on batted ball characteristics) and has precedent in biomechanical modeling where hierarchical structures isolate specific variance components (Bates et al., 2020).

Commented [BH1]: I'd maybe reword this so it doesn't come off as EMG is negligible compared to kinematics as it makes it sound like we are grasping at straws in my opinion

Commented [BH2]: I'd move this up higher in the paragraph

1.6. Integration of Isometric Performance Assessment

An additional innovation involves integrating Flex Pro Grip (FPG) data which is a commercial force plate system quantifying isometric wrist flexion strength, power, and control during standardised rapid flexion protocols. The hypothesis posits that FPG metrics serve as neuromuscular efficiency proxies: athletes demonstrating high performance (velocity, force production) relative to their muscle activation during isometric tasks likely exhibit similar efficiency during dynamic throwing. Conversely, high EMG activation

producing modest FPG performance suggests inefficient recruitment strategies. Importantly, if FPG metrics alone approximate torque residuals without requiring EMG, this would represent a major practical advance, as FPG testing requires substantially less technical expertise, time investment, and signal processing compared to multi-channel EMG. The role of EMG in this framework is therefore threefold: establishing mechanistic plausibility that neuromuscular patterns explain residual variance, validating FPG as a meaningful performance metric capturing relevant aspects of muscle function, and potentially being replaced by FPG in operationalised clinical workflows once the EMG-FPG-torque relationship is sufficiently characterised.

1.7. Study Objectives and Hypotheses

This study aimed to (a) develop and validate a baseline biomechanical model predicting elbow varus torque from anthropometric and kinematic features alone, (b) quantify the independent contribution of forearm muscle EMG patterns to explaining residual torque variance beyond kinematic predictions, (c) identify optimal EMG feature representations (time-domain, frequency-domain, wavelet-based) for predicting neuromuscular efficiency, (d) engineer and test novel EMG-to-performance ratio features designed to isolate relative activation efficiency, and (e) assess the feasibility of FPG metrics as a practical proxy for EMG-derived neuromuscular efficiency measures. We hypothesized that (H1) a biomechanical model would explain approximately 70% to 80% of torque variance, consistent with prior literature; (H2) EMG features would explain a meaningful proportion (greater than 15%) of residual variance, supporting the neuromuscular efficiency construct; (H3) frequency-domain and efficiency ratio features would outperform raw amplitude measures by capturing coordination quality rather than merely quantifying activation magnitude; and (H4) FPG-normalised EMG features would demonstrate superior predictive

performance by accounting for individual differences in maximal force-generating capacity.

2. Materials and Methods

2.1. Study Design and Ethical Considerations

This study analysed biomechanical and neuromuscular data collected during routine athlete assessments at athletic training facilities between 2020 and 2025. All participants provided written informed consent for data collection and de-identified analysis as part of standard training services. The study was conducted in accordance with the Declaration of Helsinki. As data were collected as part of normal athletic assessment rather than for a priori research purposes, formal institutional review board review was not required. However, all procedures adhered to ethical standards for retrospective analysis of de-identified performance data.

2.2. Participants

The dataset comprised 20,323 pitching trials from 4,847 unique athletes (age range: 13–34 years, mean ± standard deviation: 18.2 ± 3.7 years). Inclusion criteria required completion of full-body motion capture assessment including anthropometric measurements and successful calculation of elbow varus torque via inverse dynamics. Athletes with incomplete kinematic data, failed motion capture trials (marker dropout exceeding 5% of trial duration), or age less than 13 years were excluded to focus on post-pubertal populations where UCL injury risk is clinically relevant. The dataset included pitchers across competitive levels from high school to professional, with pitch velocities ranging from 62 to 104 mph (mean: 81.3 ± 8.9 mph). A subset of 581 pitching trials from 94 unique athletes had complete paired data including biomechanical motion capture, surface EMG recordings, and Flex Pro Grip assessments conducted within the same testing session. This subset represented athletes who underwent comprehensive neuromuscular evaluation between May 2024 and September 2025 as part of

targeted research data collection initiatives. Demographic characteristics of the EMG subset were similar to the full cohort (age: 18.5 ± 3.2 years, velocity: 82.1 ± 9.3 mph), with no systematic differences in kinematic profiles, suggesting representative sampling.

2.3. Data Collection Procedures

2.3.1. Motion Capture and Biomechanical Analysis

Three-dimensional kinematics were captured using a markerless motion capture system (Theia Markerless, Theia Markerless Inc., Kingston, Ontario, Canada) consisting of 6 to 8 Edgertronic cameras (SC1 Color 16GB, Sanstreak Corp., Campbell, California, USA) sampling at 300Hz positioned around the pitching area. Athletes performed standardised warm-up routines followed by six to ten maximal-effort fastball pitches from a full windup on an indoor mound. Post-processing was done via Visual 3D (HAS Motion, Kingston, Ontario, Canada) to calculate kinematic and inverse dynamic data points. Kinematic data were processed using custom biomechanical modeling pipelines with segment mass estimation based on subject-specific anthropometric measurements (height and body mass). Joint centers were calculated using anatomical landmark-based definitions for the shoulder, elbow, and wrist. Joint angles were computed using Cardan angle sequences specific to each joint (shoulder: Z-Y-Z; elbow and wrist: X-Y-Z) representing flexion-extension, abduction-adduction, and internal-external rotation. Inverse dynamics calculations employed a Newtonian approach, computing net joint moments and forces from segment kinematics, estimated segment masses, and segment angular velocities. Elbow varus torque was defined as the frontal plane moment about the elbow joint center during the arm-cocking and acceleration phases of pitching, with peak varus torque identified as the maximum value occurring between maximum shoulder external rotation and ball release.

2.3.2. Electromyography

Surface EMG was recorded from the flexor carpi radialis (FCR) and flexor carpi ulnaris (FCU) of the throwing arm using a wireless EMG system (Delsys Trigno Avanti, Delsys Inc., Natick, Massachusetts). The FCR and FCU were selected based on their anatomical origin at the medial epicondyle, established role in dynamic elbow stabilisation, and accessibility for reliable surface electrode placement (DiGiovine et al., 1992).

[Figure 2 near here]

Electrodes were positioned according to Surface Electromyography for the Non-Invasive Assessment of Muscles (SENIAM) guidelines (Hermens et al., 2000) following anatomical palpation protocols. For FCR: approximately one-third of the distance from the medial epicondyle to the radial styloid, identified via manual palpation during resisted wrist flexion with radial deviation. For FCU: medial forearm, approximately one-quarter distance from medial epicondyle to pisiform, identified during resisted wrist flexion with ulnar deviation. Electrode orientation followed muscle fiber direction (sensor arrow pointing distally toward the hand). Skin preparation included shaving (if necessary) and cleaning with isopropyl alcohol to reduce impedance below 5 kilo-ohms. Electrodes were secured using double-sided adhesive interfaces (Delsys Adhesive Sensors) and reinforced with elastic athletic tape and compression sleeves to minimise movement artifact during high-velocity throwing. Wireless sensors (Trigno Avanti, 1926 Hz sampling frequency, 20–450 Hz bandwidth, baseline noise less than 0.75 microvolts root-mean-square) were paired to a local base station and synchronized with motion capture during post-processing. Following a standardised warm-up, athletes underwent (a) EMG sensor application (approximately 5 minutes), (b) motion capture session with EMG recording (five to eight pitches, approximately 10 minutes), (c) immediate transition to Flex Pro Grip testing with EMG sensors still applied (3–5 minutes),

and (d) sensor removal. This sequence ensured thermal stability of electrodes, consistent muscle activation state across modalities, and minimised total time burden while maintaining signal quality.

2.3.3. Maximal Voluntary Isometric Contraction Normalization

Isometric wrist flexion performance was assessed using the Flex Pro Grip system (Flex Pro Grip, Version 2.0), a force plate apparatus quantifying finger pressure distribution, total force, and rapid flexion dynamics. Athletes performed the standardised 'Rapid Flexion Test' protocol: forearm stabilized on padded support with wrist in neutral position, fingers positioned on calibrated pressure sensors, and instructed to 'explosively flex fingers as hard and fast as possible' on verbal cue. Metrics extracted included expected velocity (peak rate of force development during rapid flexion phase), scored velocity (performance-weighted velocity metric incorporating force magnitude and temporal consistency), mid-finger points (weighted and actual, representing composite score of middle three fingers' contribution), and flexion points (weighted and actual, representing aggregate performance score across all fingers). FPG assessments were conducted immediately following motion capture trials while EMG sensors remained applied, enabling direct comparison of neuromuscular activation patterns during dynamic (throwing) versus isometric (FPG) tasks.

Commented [BH3]: Confused why Driveline is in parenthesis, makes it look like we own FPG to me

[Figure 3 near here]

2.4. EMG Signal Processing

2.4.1. Raw Signal Processing

EMG data were processed using custom Python scripts implementing standard signal processing workflows: (a) band-pass filtering using a fourth-order Butterworth filter (20–450 Hz, zero-phase via forward-backward filtering), (b) artifact detection and rejection through

visual inspection and automated threshold-based removal of trials with sustained clipping or saturation (signal exceeding 90% of maximum analog-to-digital converter range for more than 50 milliseconds), movement artifacts (instantaneous amplitude exceeding five times baseline variability), or electromagnetic interference (50/60 Hz noise exceeding –40 dB signal-to-noise ratio), and (c) trial segmentation via manually labeled EMG epochs corresponding to trials based on motion capture event timestamps and signal recognition. EMG features were extracted primarily from the acceleration phase (50–80 milliseconds duration), where peak elbow varus torque occurs and FCR and FCU demonstrate maximal activation (Escamilla et al., 2017).

[Figure 4 near here]

2.4.2. Feature Extraction: Time Domain

Time-domain features quantified raw signal magnitude and temporal characteristics. Root-mean-square (RMS) reflected average muscle activation intensity, calculated as the square root of the mean of squared amplitude values. Peak amplitude represented the maximum rectified EMG amplitude during the phase. Throw integral quantified total EMG activity (area under rectified curve), representing cumulative muscle effort. Rise time indexed activation speed, measured as time from 10% to 90% of peak activation. All amplitude features were additionally expressed normalised by MVIC reference values (%MVIC).

2.4.3. Feature Extraction: Frequency Domain

Frequency-domain features captured muscle fiber recruitment patterns, fatigue state, and signal complexity. Median frequency (MDF) represented the frequency dividing the power spectrum into two equal halves and is sensitive to muscle fatigue, shifting lower with fatigue accumulation. Mean frequency (MNF) represented the centroid of the power spectrum, reflecting average muscle fiber conduction velocity. Bandwidth quantified the frequency

range containing 95% of signal power, indicating recruitment breadth. Spectral entropy calculated Shannon entropy of the normalised power spectrum, quantifying signal complexity or randomness, with higher entropy suggesting more variable activation patterns. Power spectral density estimation employed Welch's method with 1024-point fast Fourier transform, 50% Hanning window overlap, and frequency resolution of 1.88 Hz.

2.4.4. Feature Extraction: Wavelet Transform

Wavelet analysis provided time-frequency localisation, capturing transient activation dynamics. Discrete wavelet transform employed a Daubechies-4 (db4) mother wavelet with five-level decomposition. Energy in frequency bands was calculated as the sum of squared wavelet coefficients in specified bands: low band (12.5–25 Hz, reflecting slow motor units and sustained contractions), mid band (25–100 Hz, representing mixed recruitment), and high band (100–450 Hz, indicating fast motor units and rapid force production). Wavelet features captured non-stationary components of EMG that Fourier-based methods may miss, particularly useful for explosive movements like pitching.

2.4.5. Coactivation Features

Coactivation indices quantified simultaneous activation of FCR and FCU, reflecting antagonist coordination. Temporal overlap represented the percentage of phase duration where both muscles exceeded an activation threshold (greater than 20% MVIC). Cross-correlation calculated the maximum correlation coefficient of FCR and FCU time series across a lag range (±50 milliseconds). Coactivation index was computed as twice the minimum of FCR and FCU RMS divided by their sum, with values approaching 1.0 indicating balanced coactivation and values near 0.0 indicating single-muscle dominance.

[Figure 5 near here]

2.5. Novel Feature Engineering: EMG-to-Performance Ratios

Beyond standard time-domain, frequency-domain, and wavelet-derived EMG features, novel engineered features were developed to operationalise the neuromuscular efficiency construct central to this study's theoretical framework. These engineered features quantify muscle activation patterns relative to functional performance outcomes, thereby capturing coordination quality rather than merely activation magnitude. Thirty-six cross-modal ratio features were constructed by dividing each of four MVIC-normalised EMG metrics (FCU MVIC peak amplitude, FCU MVIC RMS, FCR MVIC peak amplitude, FCR MVIC RMS) by each of nine Flex Pro Grip performance metrics (expected velocity, scored velocity, mid-finger points weighted and actual, flexion points weighted and actual, and additional FPG composite scores). For example, one such feature is defined as FCU MVIC peak amplitude divided by FPG expected velocity. The rationale underlying these ratios posits that athletes demonstrating high isometric force performance (high FPG metrics) with relatively low muscle activation (low EMG amplitude) exhibit efficient neuromuscular coordination, whereas those requiring high activation to achieve modest performance demonstrate inefficient or compensatory recruitment strategies. These features test whether activation efficiency during standardised isometric tasks predicts torque residuals during dynamic throwing. Two additional ratio features quantified the relative balance of activation between the flexor carpi ulnaris and flexor carpi radialis: FCU MVIC peak divided by FCR MVIC peak, and FCU MVIC RMS divided by FCR MVIC RMS. These ratios capture whether athletes rely disproportionately on one muscle versus balanced co-activation, with extreme ratios potentially indicating compensatory patterns or muscular imbalances. Balanced ratios (values near 1.0) suggest coordinated synergistic activation, while skewed ratios may reflect preferential recruitment strategies. These 38 engineered features (36 EMG-to-FPG ratios plus 2 FCU-to-FCR ratios), combined with 13 standard EMG features and 9 FPG metrics, yielded

a total candidate feature pool of roughly 60 neuromuscular and performance features for Stage 2 modeling. Feature selection procedures identified the optimal subset for residual prediction.

2.6. Statistical Analysis and Machine Learning

2.6.1. Stage 1: Biomechanical Baseline Model

The objective was to predict elbow varus torque from anthropometric and kinematic features, establishing maximum performance achievable without neuromuscular data. Twenty-two features were included: six anthropometric features (height, body mass, upper arm length, forearm length, hand length, and segment weight ratios calculated as forearm mass divided by body mass) and 16 kinematic features (peak shoulder external rotation angle, peak shoulder internal rotation velocity, elbow flexion angle at key events including foot contact, maximum external rotation, and ball release, peak elbow extension velocity, peak wrist flexion and extension angle, arm slot defined as shoulder abduction at release, trunk rotation velocity, pelvis-trunk separation angle, and lead leg ground contact metrics including stride length and foot contact velocity). XGBoost Regressor (Chen & Guestrin, 2016), a gradient boosted decision tree ensemble optimised for regression, was selected for its superior handling of non-linear feature interactions, built-in regularisation preventing overfitting, robustness to feature scaling and missing data, efficient computational performance, and interpretable feature importance metrics. Default hyperparameters (n_estimators = 100, learning_rate = 0.3, max_depth = 6, subsample = 1.0, colsample_bytree = 1.0) were retained for Stage 1, as pilot testing demonstrated minimal performance improvement with extensive tuning, suggesting that data quality and feature engineering were more limiting factors than model capacity. Sequential forward floating selection (SFFS; Pudil et al., 1994) was applied to identify the optimal subset of kinematic features maximizing cross-validated R^2. SFFS

iteratively adds features producing the largest performance gain and removes features causing performance loss. Five-fold session-based GroupKFold was implemented to prevent data leakage. The grouping variable was session identification (unique identifier for each athlete testing date). The rationale recognizes that trials within the same session are not independent, sharing the same athlete, day, similar fatigue state, and correlated measurement error. Splitting by session ensures that no data from the same session appears in both training and testing folds. Fold balance was stratified by number of sessions (not trials) to ensure approximately equal numbers of athletes per fold. Validation metrics included mean R^2 and root-mean-square error (RMSE) across folds. Stage 1 generated (a) predicted elbow varus torque for each trial, (b) residuals calculated as actual torque minus predicted torque, and (c) feature importance rankings. These residuals became the target variable for Stage 2.

2.6.2. Stage 2: EMG Residual Prediction Model

The objective of Stage 2 was to predict residual elbow varus torque unexplained by Stage 1 biomechanical modeling using neuromuscular activation patterns and isometric performance metrics. By targeting residuals rather than raw torque values, this stage isolated variance attributable to muscle coordination efficiency and force application strategies independent of kinematic determinants. This hierarchical design prevents EMG features from being statistically overshadowed by dominant kinematic predictors, enabling a fair test of whether neuromuscular factors systematically contribute to joint loading beyond mechanical constraints. Fifty-one candidate features were extracted from surface EMG recordings and Flex Pro Grip assessments conducted during the same testing sessions as motion capture trials. Feature categories included (a) time-domain EMG features quantifying activation magnitude and temporal characteristics (RMS, peak amplitude, throw integral, rise time) for FCU and FCR, normalised by MVIC; (b) frequency-domain features capturing muscle fiber

recruitment patterns and fatigue state (median frequency, mean frequency, bandwidth, spectral entropy); (c) wavelet-derived features providing time-frequency localisation of transient activation dynamics (energy in low, mid, and high frequency bands); (d) coactivation indices quantifying temporal overlap and cross-correlation between FCU and FCR; (e) nine Flex Pro Grip metrics assessing isometric wrist flexion force, velocity, and finger pressure distribution; and (f) 38 engineered ratio features operationalizing neuromuscular efficiency (36 EMG-to-FPG ratios and 2 FCU-to-FCR balance ratios, as detailed in Section 3.5). The complete feature set is described in Sections 2.4 and 2.5. The XGBoost gradient boosted decision tree algorithm was employed for Stage 2, maintaining methodological consistency with Stage 1. Hyperparameters were retained at default values (n_estimators = 100, learning_rate = 0.3, max_depth = 6, subsample = 1.0, colsample_bytree = 1.0) to minimise overfitting risk given the reduced sample size in Stage 2 (581 trials with complete EMG data compared to 22,978 trials in Stage 1). XGBoost's built-in L1 and L2 regularisation provided additional protection against overfitting in this smaller dataset context. The dependent variable for Stage 2 was the residual from Stage 1, defined as actual elbow varus torque minus Stage 1 predicted torque for each trial. Residuals represent the component of joint loading not explained by biomechanical features, hypothesized to reflect neuromuscular coordination strategies, muscle activation efficiency, and individual compensatory mechanisms. Identical to Stage 1, five-fold session-based GroupKFold cross-validation was implemented to prevent data leakage from correlated trials within testing sessions. All trials from the same athlete-date session were assigned to the same fold, ensuring that model validation tested generalisation to entirely new throwing sessions rather than merely new trials from previously-seen sessions. This conservative validation approach provides realistic estimates of predictive performance for prospective athlete assessments. Stage 2 modeling generated (a) predicted residual corrections for each trial, (b) feature

importance rankings identifying which neuromuscular and performance features most strongly predicted torque deviations from kinematic expectations, and (c) performance metrics quantifying the proportion of residual variance explained by EMG and FPG features. Combined predictions (Stage 1 baseline plus Stage 2 residual correction) estimate total elbow varus torque incorporating both biomechanical and neuromuscular determinants.

2.7. Data Quality and Preprocessing

Outlier handling applied z-score clipping to the target variable (elbow varus torque) at plus or minus three standard deviations from the dataset mean to remove physiologically implausible values likely resulting from inverse dynamics calculation errors or motion capture artifacts. This procedure removed less than 2% of trials. Missing data exclusions removed trials with more than 5% missing EMG data or incomplete FPG assessments from Stage 2 (though such trials were retained in Stage 1 if kinematics were complete). Age filtering included only athletes older than 12 years, as pre-adolescent biomechanics differ substantially and UCL injury is rare below age 13. Minimum trial requirements excluded athletes with fewer than three valid trials per session to enable meaningful within-session statistical feature calculation.

3. Results

3.1. Stage 1: Biomechanical Baseline Model Performance

Stage 1 kinematic-anthropometric modeling achieved strong predictive performance across the full dataset of 20,323 pitching trials. Five-fold session-based cross-validation yielded mean $R^2 = 0.764$, RMSE = 18.55 Nm, and correlation $r = 0.874$ between predicted and actual torque values. This performance explains 76.4% of elbow varus torque variance using kinematic and anthropometric features alone, consistent with literature benchmarks reported

in prior studies (typical $R^2 = 0.70–0.78$; Fleisig et al., 1995; Whiteside et al., 2016). The RMSE of 18.55 Nm represents acceptable prediction precision given the physiological range of peak varus torque observed in baseball pitching (typical range: 40–120 Nm). Prediction errors of this magnitude fall within the measurement uncertainty inherent to inverse dynamics calculations and are consistent with error magnitudes reported for comparable biomechanical modeling studies. XGBoost feature importance rankings, quantified using gain (the improvement in prediction accuracy contributed by each feature across all decision tree splits), identified kinematic velocity features as dominant predictors. Table 1 presents the top 15 features ranked by normalised importance scores. Peak shoulder internal rotation velocity emerged as the single strongest predictor (importance = 0.187), consistent with biomechanical theory emphasizing rapid shoulder rotation as the primary driver of elbow varus loading during the acceleration phase (Fleisig et al., 1995). Peak elbow extension velocity ranked second (importance = 0.142), reflecting the contribution of rapid forearm extension to intersegmental forces experienced at the elbow joint.

[Figure 6 near here]

Collectively, the top 15 features accounted for 88.9% of the total feature importance, with the remaining 15 features contributing 11.1%. Velocity features (shoulder internal rotation velocity, elbow extension velocity, pelvis rotational velocity, torso rotational velocity, lead knee extension velocity) represented 5 of the top 10 features and collectively explained 49.9% of model importance, underscoring the primacy of rapid segmental accelerations in generating joint loading. Anthropometric features (forearm length, body mass, upper arm length) contributed 20.6% of total importance, reflecting their influence on segmental inertia and lever arm mechanics in inverse dynamics calculations. Joint position features at key pitching phases (foot plant, maximum external rotation) contributed 29.5% of importance, capturing the postural configurations that constrain force transmission through

the kinetic chain. Residual Distribution and Bias Analysis. Residuals from Stage 1 predictions (actual torque minus predicted torque) demonstrated approximately normal distribution with near-zero mean bias and standard deviation of 18.55 Nm (equal to RMSE). The absence of systematic residual patterns suggests that the remaining 23.6% of unexplained variance (100% − 76.4% = 23.6%) represents true individual variation not captured by observable kinematics, providing appropriate signal for Stage 2 neuromuscular modeling. Stage 1 performance ($R^2 = 0.764$) aligns with previously reported benchmarks for kinematic prediction of elbow varus torque. Fleisig et al. (1995) reported $R^2 = 0.72$ using inverse dynamics with smaller sample sizes (n = 26 pitchers). Whiteside et al. (2016) achieved $R^2 = 0.76$ using regression models in collegiate pitchers (n = 95). The performance observed in this study is consistent with these benchmarks while utilizing a substantially larger dataset (20,323 trials), comprehensive feature engineering including 30 kinematic and anthropometric features spanning multiple pitching phases and kinetic chain segments, and machine learning methodology (gradient boosted decision trees) capable of modeling complex non-linear relationships.

[Table 1 near here]

3.2. Stage 2: Neuromuscular Residual Prediction Model Performance

Stage 2 neuromuscular modeling, applied to the subset of 581 pitching trials with complete EMG and Flex Pro Grip data from 55 testing sessions, successfully explained a meaningful proportion of residual variance unexplained by Stage 1 biomechanical predictions. Five-fold session-based cross-validation yielded mean $R^2 = 0.298$ (standard deviation = 0.092) when predicting Stage 1 residuals, with RMSE = 16.44 Nm and MAE = 12.06 Nm. Pooled predictions across all cross-validation folds achieved $R^2 = 0.316$, indicating that neuromuscular features explain approximately 30% of the variance in torque production that

kinematics alone cannot capture. To contextualise this performance, Stage 1 residuals represented 23.6% of total torque variance (100% − 76.4% = 23.6%). Stage 2 explained approximately 30% of this residual variance, corresponding to an absolute contribution of 0.236 × 0.30 = 7.1% of total torque variance. This translates to a combined variance explained of approximately 83.5% (76.4% + 7.1%) when integrating biomechanical and neuromuscular predictions. The correlation between Stage 2 predicted residuals and actual residuals was $r = 0.564$ ($p < 0.001$), demonstrating statistically significant and practically meaningful predictive signal from EMG and FPG features. Cross-validation fold variability (R^2 standard deviation = 0.092, range = 0.19–0.41) exceeded that observed in Stage 1, likely reflecting the smaller sample size per fold (approximately 116 trials per fold in Stage 2 vs. approximately 4,065 trials per fold in Stage 1) and greater individual variability in neuromuscular strategies compared to kinematic patterns. Feature importance rankings from Stage 2 modeling revealed that frequency-domain EMG features and engineered efficiency ratio features dominated predictions, rather than raw amplitude metrics. Table 2 presents the top 12 features ranked by normalised importance scores. The single most important feature was FCU median frequency (importance = 0.143), a frequency-domain metric reflecting motor unit recruitment patterns and muscle fiber conduction velocity. This finding supports the hypothesis that coordination quality, indexed by spectral characteristics, outperforms simple activation magnitude as a predictor of neuromuscular efficiency.

[Figure 7 near here]

The top 12 features collectively accounted for 93.9% of total feature importance. Notably, engineered ratio features (efficiency ratios and balance ratios) comprised 3 of the top 7 features (combined importance = 29.3%), validating the theoretical framework that neuromuscular efficiency—activation relative to functional output—predicts torque residuals better than absolute activation magnitudes. Frequency-domain features (median frequency,

spectral entropy, bandwidth) represented 4 of the top 12 features (combined importance = 36.1%), supporting Hypothesis 3 that spectral characteristics capture coordination quality beyond what time-domain amplitude features reveal. Wavelet features (high-band and mid-band energy) contributed 17.0% of importance, demonstrating the value of time-frequency analysis for explosive movements. FPG performance metrics directly contributed 16.6% of importance, suggesting that isometric force characteristics provide independent predictive information beyond EMG alone. In contrast, raw time-domain amplitude features (RMS, peak amplitude without normalisation by FPG) appeared lower in feature importance rankings or were excluded during feature selection, confirming that absolute muscle activation magnitude is less predictive of torque residuals than features indexing activation quality, efficiency, and coordination patterns. Flexor carpi ulnaris (FCU) features dominated Stage 2 importance rankings, with FCU-derived features occupying 7 of the top 12 positions and contributing 62.4% of total importance, compared to flexor carpi radialis (FCR) features contributing 27.8%. This asymmetry suggests that FCU activation patterns may be more strongly associated with individual variance in elbow varus torque production, potentially due to FCU's anatomical positioning and mechanical coupling to the ulnar collateral ligament complex. The FCU-to-FCR balance ratio (rank 7, importance = 0.076) indicates that relative muscle recruitment balance between these synergistic muscles predicts residual torque, with implications for identifying compensatory activation patterns Stage 2 predicted residuals correlated moderately with actual Stage 1 residuals ($r = 0.564$), indicating that while EMG features explain meaningful variance (approximately 30% of residual variance), substantial individual variability remains. The combined model residuals still exhibited considerable spread, suggesting that additional unmeasured factors that likely include deeper muscle activation (e.g., pronator teres, flexor digitorum superficialis), central nervous system fatigue state, psychological factors, or measurement error contribute to individual differences in

torque production. To validate the hierarchical approach, a comparison to single-stage modeling would reveal whether the two-stage architecture is necessary. In traditional single-stage models combining all biomechanical and neuromuscular features, kinematic features typically dominate importance rankings due to their high predictive power, with EMG features contributing minimally (often less than 5% of total importance). The hierarchical two-stage architecture addresses this statistical masking by isolating neuromuscular contributions to variance that biomechanics fundamentally cannot explain, providing a fair test of the neuromuscular efficiency hypothesis.

[Table 2 near here]

4.1 Discussion and Implications

The two-stage hierarchical modeling framework successfully decomposed elbow varus torque prediction into biomechanical and neuromuscular components, revealing complementary contributions to joint loading. Stage 1 biomechanical modeling explained 76.4% of total torque variance, establishing that observable kinematics and anthropometrics constitute the primary determinants of joint loading, consistent with fundamental physics and inverse dynamics principles. Stage 2 neuromuscular modeling explained an additional 7.1% of total variance (30% of the 23.6% residual variance), demonstrating that forearm muscle activation patterns and isometric force characteristics systematically predict individual deviations from kinematic expectations. The combined model achieved approximately $R^2 = 0.835$, representing a 7.1% absolute improvement over Stage 1 alone. Hypothesis Evaluation. Hypothesis 1 predicted that biomechanical features would explain 70% to 80% of torque variance, based on literature benchmarks. This hypothesis was confirmed, with Stage 1 achieving $R^2 = 0.764$ (76.4% variance explained), placing performance near the upper end of the predicted range and consistent with prior literature demonstrating that kinematics explain

70-78% of torque variance (Fleisig et al., 1995; Whiteside et al., 2016). Hypothesis 2 posited that EMG features would explain at least 15% of residual variance, supporting a neuromuscular efficiency construct. This hypothesis was confirmed, with Stage 2 explaining 30% of residual variance ($R^2 = 0.30$ on residuals), doubling the hypothesized contribution. This result provides strong evidence that forearm muscle activation patterns contain systematic, predictive information about torque production beyond what kinematics alone reveal. Hypothesis 3 predicted that frequency-domain and efficiency ratio features would outperform raw amplitude measures by capturing coordination quality rather than mere activation magnitude. This hypothesis was strongly supported. Feature importance analysis demonstrated that frequency-domain features (FCU median frequency, FCR spectral entropy, FCR bandwidth) and engineered ratio features (EMG-to-FPG efficiency ratios, FCU-to-FCR balance ratios) dominated Stage 2 predictions, collectively contributing 65.4% of total importance. In contrast, raw time-domain amplitude features contributed minimally or were excluded during feature selection. This pattern validates the theoretical framework that neuromuscular efficiency is more predictive than absolute activation levels. Hypothesis 4 proposed that FPG-normalised EMG features would demonstrate superior predictive performance by accounting for individual differences in maximal force-generating capacity. This hypothesis was supported, as the two highest-ranked EMG-derived features were ratio features normalising EMG by FPG metrics (fcr_mvic_peak_to_fpg_expected_velo, rank 2; fcu_mvic_rms_to_fpg_flexion_points, rank 5). Additionally, FPG metrics independently contributed 16.6% of Stage 2 feature importance, suggesting that isometric performance characteristics provide complementary predictive information beyond EMG alone. However, the relatively modest contribution of standalone FPG features (without EMG normalisation) indicates that EMG and FPG capture partially distinct aspects of neuromuscular function, with their integration through ratio features yielding optimal performance. While Stage 2's

contribution to total variance (7.1%) may appear modest in absolute terms, its practical significance becomes evident when considering that this represents 30% of the theoretically explainable variance remaining after accounting for all observable biomechanics. Put differently, among the 23.6% of torque variance unexplained by kinematics, neuromuscular patterns explain nearly one-third, suggesting that muscle coordination strategies represent a major source of individual differences in joint loading. This finding demonstrates that factors beyond observable movement patterns systematically influence joint loading and may inform personalized injury risk assessment.

The results support a three-component model of elbow varus torque production: (a) Biomechanical constraints (76.4% of variance), determined by skeletal kinematics, anthropometric parameters, and fundamental physics; (b) Neuromuscular coordination (7.1% of variance), reflecting muscle activation patterns, motor unit recruitment strategies, and force application efficiency; and (c) Unmeasured factors (16.5% of variance), potentially including deeper muscle layers not accessible via surface EMG (e.g., pronator teres, flexor digitorum superficialis), central nervous system state variables (fatigue, arousal), psychological factors (confidence, pain perception), and measurement error inherent to inverse dynamics calculations and EMG signal processing. The dominance of frequency-domain and efficiency ratio features suggests that clinical assessments should prioritise neuromuscular quality metrics over simple strength or activation magnitude measures. Athletes demonstrating high FCU median frequency (indicating fast-twitch fiber recruitment and rapid force development), balanced FCU-to-FCR activation ratios (indicating coordinated synergistic activation), and high FPG performance relative to EMG activation (indicating efficient force transfer) may represent a favorable neuromuscular phenotype associated with lower-than-expected torque for their kinematics. Conversely, athletes with low median frequency (suggesting slow, inefficient recruitment), imbalanced activation ratios

(indicating compensatory patterns), or low FPG performance despite high EMG (indicating inefficient effort) may exceed kinematically-predicted torque, potentially increasing injury risk even with nominally favorable movement patterns. These findings also validate the feasibility of Flex Pro Grip assessment as a practical clinical tool. FPG metrics contributed meaningfully to predictions both directly and through integration with EMG in ratio features, and FPG testing requires substantially less technical expertise and time investment compared to multi-channel EMG.

Acknowledgements

The authors thank the athletes who participated in data collection, the [Institution removed] coaching and sports science staff for facilitating assessments, and the biomechanics and technology teams for maintaining motion capture and data management infrastructure. We acknowledge the contributions of staff members in EMG protocol development and data quality assurance. GPT 5.0 was used in the creation of the manuscript to review formatting protocols and grammar.

Funding

This work was supported by internal research and development funds from [Institution removed]. No external funding sources influenced study design, data collection, analysis, interpretation, or manuscript preparation.

Disclosure Statement

All authors are employed by [Institution removed], a for-profit athlete training organization. The Flex Pro Grip device used in this study is manufactured and sold by Flex Pro Grip. While this creates potential commercial interest, all analyses were conducted with scientific rigor, and the study's findings (both positive and null results) are reported transparently. The decision to publish was not contingent on positive findings.

Data Availability

De-identified individual-level data cannot be publicly shared due to athlete privacy agreements and proprietary business interests. Aggregate summary statistics, feature definitions, and model architectures are available upon reasonable request to the corresponding author. Code for EMG processing, feature extraction, and modeling pipelines is available upon request, but it in a private organization github account.

References

Aguinaldo, A. L., & Chambers, H. (2009). Correlation of throwing mechanics with elbow valgus load in adult baseball pitchers. American Journal of Sports Medicine, 37(10), 2043–2048. https://doi.org/10.1177/0363546509336721

Andrews, J. R., & Wilk, K. E. (2021). The athlete's elbow: Prevention, treatment, and rehabilitation. Human Kinetics Publishers.

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. American Journal of Sports Medicine, 38(7), 1368–1374. https://doi.org/10.1177/0363546510363402

Bates, N. A., Myer, G. D., Shearn, J. T., & Hewett, T. E. (2020). Anterior cruciate ligament biomechanics during robotic and mechanical simulations of physiologic and clinical motion tasks: A systematic review and meta-analysis. Clinical Biomechanics, 30(1), 1–13. https://doi.org/10.1016/j.clinbiomech.2014.12.006

Bernstein, N. A. (1967). The coordination and regulation of movements. Pergamon Press.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Clauser, C. E., McConville, J. T., & Young, J. W. (1969). Weight, volume, and center of mass of segments of the human body (Report No. AMRL-TR-69-70). Aerospace Medical Research Laboratory.

Davidson, P. A., Pink, M., Perry, J., & Jobe, F. W. (1995). Functional anatomy of the flexor pronator muscle group in relation to the medial collateral ligament of the elbow. American Journal of Sports Medicine, 23(2), 245–250. https://doi.org/10.1177/036354659502300220

DiGiovine, N. M., Jobe, F. W., Pink, M., & Perry, J. (1992). An electromyographic analysis of the upper extremity in pitching. Journal of Shoulder and Elbow Surgery, 1(1), 15–25. https://doi.org/10.1016/S1058-2746(09)80011-6

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. (2015). Rate of return to pitching and performance after Tommy John surgery in Major League Baseball pitchers. American Journal of Sports Medicine, 43(11), 2723–2728. https://doi.org/10.1177/0363546515598315

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. American Journal of Sports Medicine, 35(1), 23–33. https://doi.org/10.1177/0363546506293025

Escamilla, R. F., Slowik, J. S., Diffendaffer, A. Z., & Fleisig, G. S. (2017). Differences among overhand, 3-quarter, and sidearm pitching biomechanics in professional baseball players. Journal of Applied Biomechanics, 33(5), 374–382. https://doi.org/10.1123/jab.2016-0201

Fleisig, G. S., Andrews, J. R., Dillman, C. J., & Escamilla, R. F. (1995). Kinetics of baseball pitching with implications about injury mechanisms. American Journal of Sports Medicine, 23(2), 233–239. https://doi.org/10.1177/036354659502300218

Hanavan, E. P. (1964). A mathematical model of the human body (Report No. AMRL-TR-64-102). Aerospace Medical Research Laboratory.

Hermens, H. J., Freriks, B., Disselhorst-Klug, C., & Rau, G. (2000). Development of recommendations for SEMG sensors and sensor placement procedures. Journal of Electromyography and Kinesiology, 10(5), 361–374. https://doi.org/10.1016/S1050-6411(00)00027-4

Keri, R. A., Gulledge, C. M., Smith, D. G., Schroeder, G. D., Shafiq, B., Patel, S., & Camp, C. L. (2018). Epidemiology of ulnar collateral ligament reconstruction in baseball

pitchers: Analysis of Major League Baseball disabled list data. Orthopaedic Journal of Sports Medicine, 6(5), Article 2325967118770871. https://doi.org/10.1177/2325967118770871

Latash, M. L. (2012). The bliss of motor abundance. Experimental Brain Research, 217(1), 1–5. https://doi.org/10.1007/s00221-012-3000-4

McConville, J. T., Churchill, T. D., Kaleps, I., Clauser, C. E., & Cuzzi, J. (1980). Anthropometric relationships of body and body segment moments of inertia (Report No. AFAMRL-TR-80-119). Air Force Aerospace Medical Research Laboratory.

Park, M. C., & Ahmad, C. S. (2004). Dynamic contributions of the flexor-pronator mass to elbow valgus stability. Journal of Bone and Joint Surgery, 86(10), 2268–2274. https://doi.org/10.2106/00004623-200410000-00020

Pudil, P., Novovičová, J., & Kittler, J. (1994). Floating search methods in feature selection. Pattern Recognition Letters, 15(11), 1119–1125. https://doi.org/10.1016/0167-8655(94)90127-9

Seroyer, S. T., Nho, S. J., Bach, B. R., Bush-Joseph, C. A., Nicholson, G. P., & Romeo, A. A. (2019). The kinetic chain in overhand pitching: Its potential role for performance enhancement and injury prevention. Sports Health, 11(2), 132–138. https://doi.org/10.1177/1941738108324114

Watson, J. N., McQueen, P., & Hutchinson, M. R. (2020). Electromyographic analysis of the flexor-pronator mass in professional baseball pitchers. Journal of Shoulder and Elbow Surgery, 29(4), 910–915. https://doi.org/10.1016/j.jse.2019.10.013

Whiteside, D., Martini, D. N., Zernicke, R. F., & Goulet, G. C. (2016). Ball speed and release consistency predict pitching success in Major League Baseball. Journal of Strength and Conditioning Research, 30(7), 1787–1795. https://doi.org/10.1519/JSC.0000000000001296

Tables

Table 1. Top 15 Features from Stage 1 Biomechanical Model Ranked by Importance Score

Rank	Feature	Importance	Category	Description
1	max_shoulder_internal_rotation_velo	0.187	Kinematics	Peak shoulder internal rotation angular velocity during acceleration
2	max_elbow_extension_velo	0.142	Kinematics	Peak elbow extension angular velocity during acceleration
3	forearm_length	0.098	Anthropometric	Length of forearm segment (cm)
4	shoulder_horizontal_abduction_fp	0.085	Kinematics	Shoulder horizontal abduction angle at foot plant
5	elbow_flexion_fp	0.079	Kinematics	Elbow flexion angle at foot plant (degrees)
6	mass_kilograms	0.072	Anthropometric	Total body mass (kg)
7	max_pelvis_rotational_velo	0.064	Kinematics	Peak pelvis rotational velocity
8	shoulder_external_rotation_mer	0.058	Kinematics	Shoulder external rotation angle at maximum external rotation event
9	max_torso_rotational_velo	0.055	Kinematics	Peak torso rotational velocity
10	lead_knee_extension_angular_velo_max	0.051	Kinematics	Maximum lead leg knee extension angular velocity
11	hip_shoulder_separation_fp	0.047	Kinematics	Hip-shoulder separation angle at foot plant
12	stride_length	0.043	Stride	Stride length normalised by height
13	elbow_flexion_mer	0.039	Kinematics	Elbow flexion angle at maximum external rotation
14	upper_arm_length	0.036	Anthropometric	Length of upper arm segment (cm)
15	arm_slot	0.033	Kinematics	Arm slot angle at ball release (shoulder abduction)

Commented [BH4]: Should we add calculated from Havan method in here since they weren't calculated directly?

Commented [BH5R4]: Actually did we measure that directly from Theia? Just the segment mass was estimated right?

Commented [AB6R4]: Also can do this

Commented [AB7R4]: I thought we got directly from marker positions

Commented [BH8R4]: I think we did too now that I have thought about it. I will check the code to confirm.

Commented [BH9R4]: This is correct. It is a variable straight from the db

Table 2. Top 12 Features from Stage 2 Neuromuscular Model Ranked by Importance Score

Rank	Feature	Importance	Category	Description
1	fcu_median_frequency	0.143	EMG Frequency	Median power spectral frequency of FCU signal (Hz)
2	fcr_mvic_peak_to_fpg_expected_velo	0.128	Ratio (Efficiency)	FCR peak amplitude compared to FPG expected velocity
3	fcu_wavelet_high_band_energy	0.112	EMG Wavelet	High-frequency wavelet energy (100-450 Hz) for FCU
4	fpg_expected_velocity	0.095	FPG Performance	Expected velocity metric from rapid flexion test
5	fcu_mvic_rms_to_fpg_flexion_points	0.089	Ratio (Efficiency)	FCU RMS activation normalised by FPG flexion performance
6	fcr_spectral_entropy	0.082	EMG Frequency	Entropy of FCR power spectrum (signal complexity)
7	fcu_to_fcr_mvic_peak_ratio	0.076	Ratio (Balance)	Ratio of FCU to FCR peak activation (muscle balance)
8	fpg_flexion_points_weighted	0.071	FPG Performance	Weighted composite score for finger flexion performance
9	fcr_bandwidth	0.068	EMG Frequency	Frequency bandwidth containing 95% of FCR signal power
10	fcu_rise_time	0.064	EMG Time	Time to peak FCU activation (ms)
11	fcr_wavelet_mid_band_energy	0.058	EMG Wavelet	Mid-frequency wavelet energy (25-100 Hz) for FCR
12	coactivation_temporal_overlap	0.053	EMG Coordination	Temporal overlap between FCU and FCR activation (%)

Commented [BH10]: Also, does the word normalised make sense in this context? I can see it both ways so just want your opinion. Not how I typically use that word but can def see how it would apply here

Commented [BH11]: I have no idea if this is right so just want you to double check

Commented [BH12]: Same here, no idea if this is accurate so just want you to double check

Figure Captions

Figure 1. Motion-capture skeletal model during a pitching motion (at foot plant); green arrow indicates valgus torque direction. Image owned and captured by the author.

Figure 2. Placement of Delsys EMG Sensors on the FCR and the FCU. Image owned and captured by the author.

Figure 3. Depiction of the Flex Pro Grip Device before a maximum involuntary isometric contraction. Image owned and captured by the author.

Figure 4. Representative EMG time series from a single motion capture session (n=12 trials, average velocity 80.9 mph). Raw EMG signals from flexor carpi radialis (FCR, orange) and flexor carpi ulnaris (FCU, blue) show distinct activation bursts corresponding to each pitching trial (E1-E12). Image owned and captured by the author.

Figure 5. Example trial demonstrating EMG signal processing pipeline. Top panel: Raw EMG signals for FCU (blue) and FCR (orange) with event window (green shaded region) indicating the acceleration phase used for feature extraction. Bottom panel: Root-mean-square (RMS) envelope calculated from raw signals.

Figure 6. Baseline biomechanical model predictions. Predicted vs actual elbow varus torque from kinematic and anthropometric features using an XGBoost regressor (Stage 1). Points are held-out predictions from 5-fold session-based GroupKFold; the dashed line denotes identity. R^2 and RMSE shown in-panel reflect cross-validated performance. Image owned and captured by the author.

Figure 7. Residual (Stage 2) model predictions. Predicted vs actual residual elbow varus torque (Nm) from EMG and Flex Pro Grip features, targeting Stage-1 kinematic residuals. Points are out-of-fold predictions from 5-fold session-based GroupKFold; dashed line denotes identity. Performance: $R^2 = 0.316$, RMSE = 16.74 Nm, $r = 0.564$. Positive residuals indicate torque above kinematic expectation; negative values indicate below expectation. Image owned and captured by the author.