Introduction

In competitive sports, injuries can significantly affect both team outcomes and individual performances. As physical demands and training loads continue to increase, medical staff and coaches will rely on measurable data to prevent injuries before they happen. This would raise an important question: Could these key indicators be used to determine players' risk of injury? The reason it matters is we are able to identify trends that are predictive. It is crucial for informing injury prevention which can be used to modify training interventions before an injury may occur. It can optimize performance to ensure training volume and intensity is appropriately balanced to minimize injury risk while having maximized athletic performance. There is additionally an application into health where evaluating and application of findings have the potential to be applied in fields like rehabilitation and physical therapy.

Methods

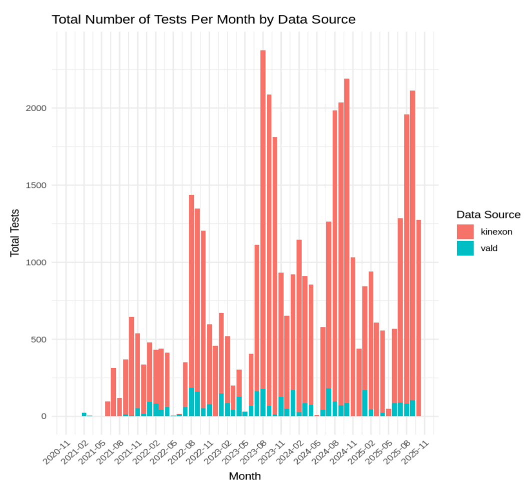
The metrics selected for this review were viewed as the most important for injury assessment performance optimization. The metric of total distance (TD) is a measure that tracks the entire distance a player travels during a competition or training session, typically using wearable GPS or LPS units (Junior et al., 2021). The metric of Max Force, or Maximal Strength, measures the greatest force a muscle group can generate, typically measured during maximal voluntary isometric contraction (Ashworth et al., 2025). The metric of torque measures the rotational force produced by a muscle group acting around a joint (Ashworth et al., 2025). The metric of Accumulated Acceleration Load (AAL) is a measure of external workload that quantifies the accumulated intensity of a player's movements over time, typically tracked using accelerometers within wearable microtechnology units (Koyama et al., 2024).

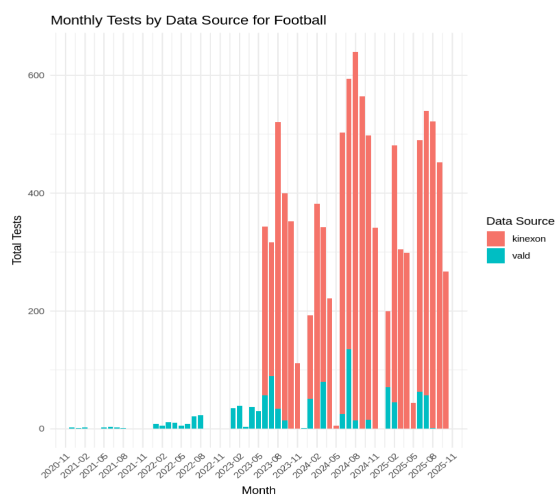
The approach used to filter the data and ensure the data's quality was to analyze any missing data from the current data. For each sport/team we calculated what percentage of the athletes have at least 5 measurements for the selected metrics ('leftMaxForce', 'rightMaxForce', 'leftTorque', 'rightTorque', 'accel_load_accum', 'distance_total'). Based on the data we found, while a significant amount of athletes have sufficient measurements for the selected metrics, a considerable amount of athletes remain with insufficient data. This may impact the strength of our analysis and validity of the conclusions gathered. We determined that any players with team names in ('Unknown','Player Not Found','Graduated (No longer enrolled)') would be excluded from our analysis. We designated these players as inactive to ensure data integrity and accuracy. To address this we propose the following: The first is data imputation, for athletes that have missing measurements, exploring data imputation techniques that can estimate missing values based on value trends, retaining more athletes in analysis. The second is a focused analysis, considering to narrow the research question to focus on athletes with sufficient data, ensuring findings are based on reliable measurements. The third is additional data collection, looking into collecting additional data for athletes with insufficient measurements to enhance the datasets completeness. The fourth is sensitivity analysis, to conduct sensitivity analyses to understand how the presence of missing data might affect interpretations or results.The approach used for data cleaning was first to define the list of metrics and player names. Using metric_list = ['leftMaxForce', 'rightMaxForce', 'leftTorque', 'rightTorque', 'accel_load_accum', 'distance_total'] and using playername_list = ['PLAYER_755', 'PLAYER_690', 'PLAYER_1128']. Next call the function to get data in "long format" (one row per metric per timestamp) to do "wide format" get_data_in_wide_format_by_athlete_and_metric(metric_list, playername_list,"wide") and get_data_in_wide_format_by_athlete_and_metric(metric_list, "all","wide").

The statistical methods used in the review was a mix between descriptive and inferential statistics. The descriptive statistics were used to find the means, medians and ranges for the most recorded metric. The inferential statistics used were t-tests, correlation analyzes, and regression models to explore relationships between max force/torque and accumulated acceleration/distance.

Results

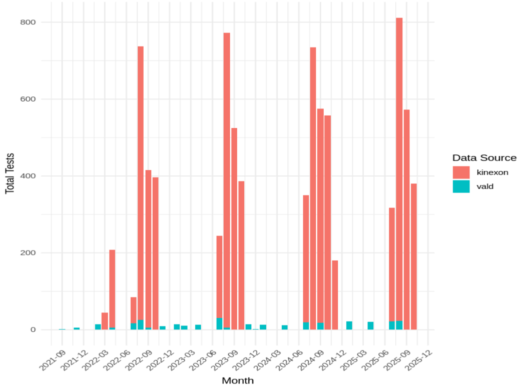
In our findings the symmetrical metrics were recorded on a separate device (Vald) from non-symmetrical metrics for acceleration and distance using (kinexon). For the testing behaviors among the teams, the output for kinexon data, "total_distance" and "accel_load", are more frequently tested versus Vald data "Torque" and "Max_force". There is also a weak correlation found between left and right max force and torque versus accumulated acceleration or distance for each player and team.



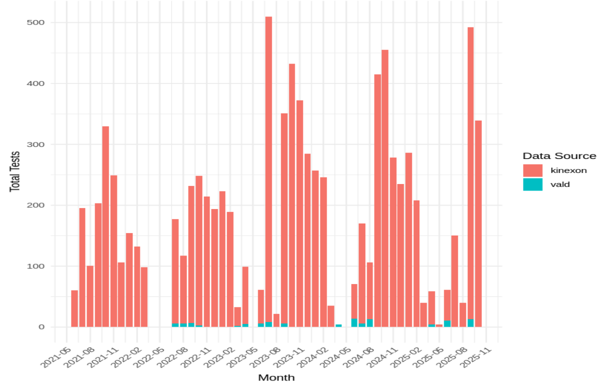


Total Number of Tests Per Month (Kinexon and Vald)

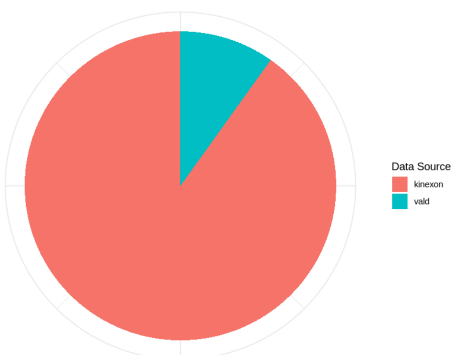
Monthly Tests by Data Source for Women's Soccer



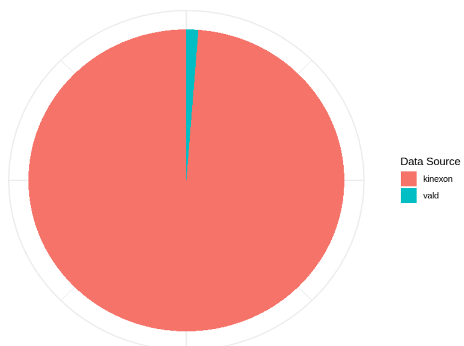
Monthly Tests by Data Source for Mens Basketball



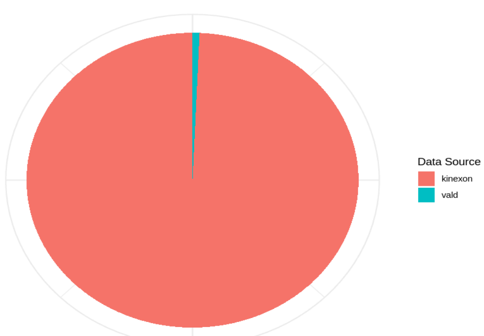
Percentage of Source Testing for Football



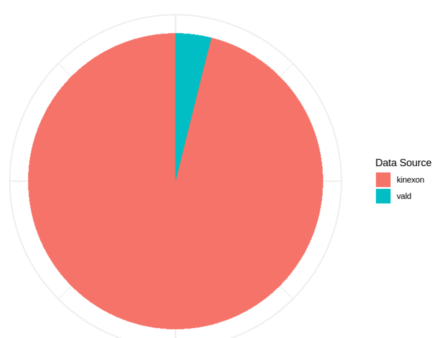
Percentage of Source Testing for Men's Basketball

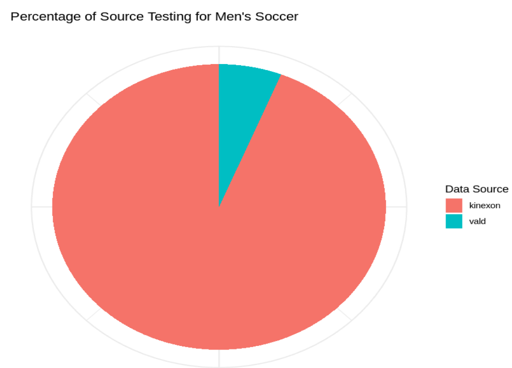


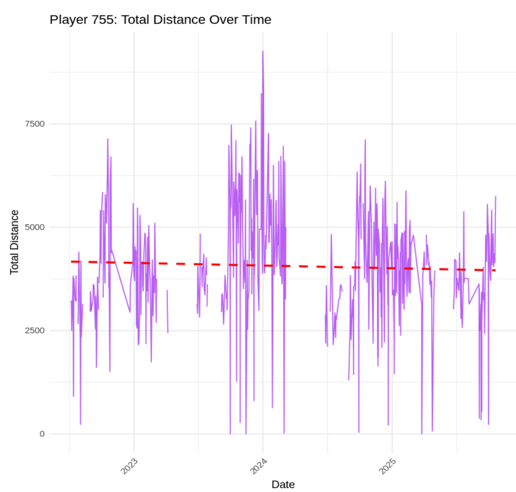
Percentage of Source Testing for Women's Basketball

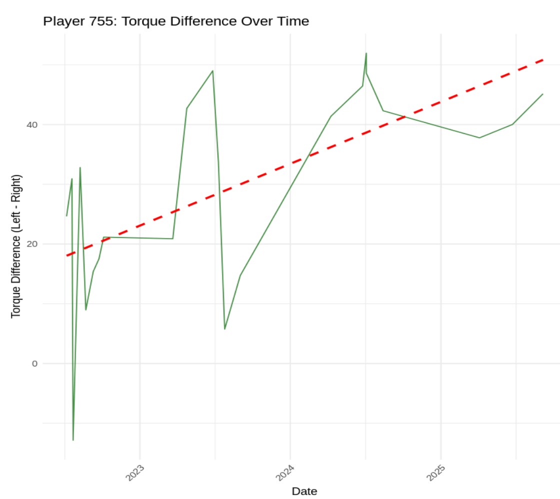


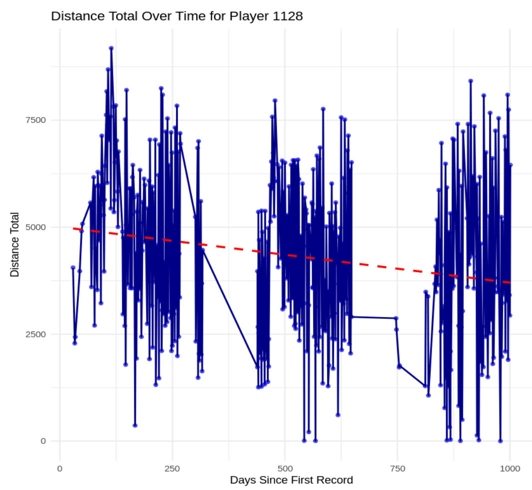
Percentage of Source Testing for Women's Soccer

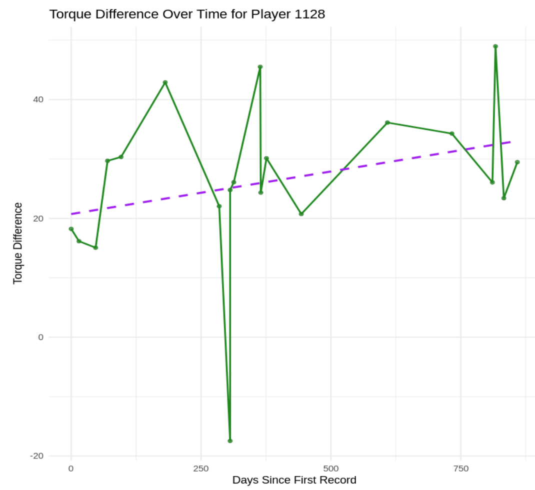


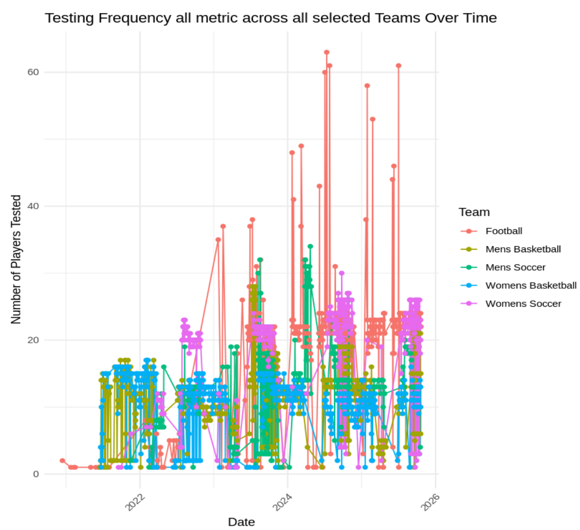

Percentage of Source Testing for Men's Soccer
Data Source
kinexon
vald

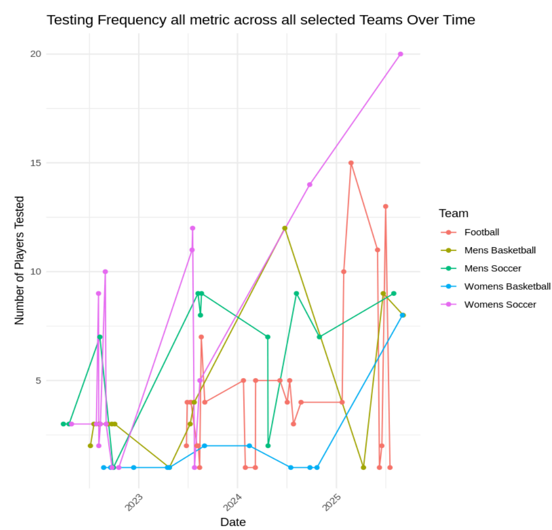

Player 755: Total Distance Over Time
Total Distance
7500
5000
2500
0
2023
2024
2025
Date


Player 755: Torque Difference Over Time
Torque Difference (Left - Right)
40
20
0
2023
2024
2025
Date









In the t-test results for torque and max force for p-values. accel_load_accum vs. torque_asymmetry_category: P-value = 0.8672, distance_total vs. torque_asymmetry_category: P-value = 0.3264, accel_load_accum vs. max_force_asymmetry_category: P-value = 0.9979, distance_total vs. max_force_asymmetry_category: P-value = 0.4202. P-value is greater than 0.05. In the linear regression results for p-values and r-square values. accel_load_accum vs. avg_torque_asymmetry: P-value = 0.1191; R-squared = 0.003805, distance_total vs. avg_torque_asymmetry: P-value = 0.0404; R-squared = 0.008872, accel_load_accum vs. avg_max_force_asymmetry: P-value = 0.05434; R-squared = 0.007393, distance_total vs. avg_max_force_asymmetry: P-value = 0.05434; R-squared = 0.007393. P-value is greater than 0.05 and R-square lower than 0.3.

Discussion

Our findings compared to our literature shows that the statistical findings suggest that both torque_asymmetry and max_force_asymmetry have a very limited, if any, linear relationship with accel_load_accum and distance_total in this dataset. Most p-values for t-tests and linear regressions were not statistically significant, and the adjusted R-squared values were extremely low ranging from (0.002773 to 0.008872). Overall Average Asymmetrical Torque (avg_asymmetrical_torque): 1.936%. Overall Average Asymmetrical Max Force (avg_asymmetrical_max_force): 1.961%. Our literature review indicated athletes with 10-15% asymmetrical differences are at a higher risk of injury. However the asymmetrical average from our dataset were all less than 2%, this suggests the majority of our participants displayed mild asymmetric symptoms compared to those flagged in our literature review. It also indicates asymmetries would negatively impact an athlete's ability to cover distance efficiently or generate acceleration effectively. The data also suggests possible factors that contribute to our weak correlation is limited data matched players metrics records by date Vald vs Kinexon. Our asymmetrical averages were all less than 2%, our literature review identified groups with 10% or more as high risk, perhaps the majority of the candidates in our dataset may not have symptoms severe enough to affect acceleration and distance performance. This study does not account for athletes' natural abilities to adapt and overcome asymmetrical assessments.

The gaps we identified in our findings were that it created a personalized injury-risk model that will help players understand their own unique load patterns. To help better understand force/torque monitoring to manage load-management to protect player fatigue and improve endurance training. Some practical implications for coaches/trainers is to identify players with high standard deviation of symmetrical metrics to determine players risk of injury. Monitor active players inactivity to determine whether this is due to injury or impaired abilities from fatigue. Create a regularized screening and testing schedule: One set of bilateral & one non-bilateral simultaneously. Suggest weekly intervals: More data to calculate ACWR: 3-6 weeks of testing.

Limitations & Future Directions

The challenges that were encountered with the data's limitations were there was missing or incomplete data. In the data there was a duplication of team names and entries with missing values which reduced the number of usable values. These variations created duplicates and required normalization to ensure accurate analysis and grouping. This additionally required imputation or removal of incomplete entries, potentially affecting robustness of the results. Invalid team names that were listed in the data did not correspond to the active teams requiring filtering and manual coding. This can lead to upstream data-entry or integration problems. There was difficulty involving aligning symmetrical metrics with accumulated acceleration and total distance values. This issue complicated integration of multiple performance metrics into a unified analysis. In some cases athlete identification codes were missing due to factors like illness, injury, absences, or failure. This prevented linking athlete metrics across sessions, limiting longitudinal analysis and reducing completeness of performance trends.

The recommendations for research in the future would be to expand the data collection for injury risk assessment. Additional data is needed to determine current performance metrics that can reliably serve as predictors for risk of injury. A larger more comprehensive dataset would improve statistical power and allow more robust modeling on injury related outcomes. In future research it should closely examine specific metrics such as maximum force and acceleration load; the variables may reflect stresses that contribute to injury susceptibility, but the relationship would require a targeted analysis. To better understand factors linked to injury severity and occurrence, future work should analyze injury absences related to total distance, acceleration load, and asymmetry percentages. How these metrics interact may reveal patterns that identify athletes at elevated risks beyond which injury likelihood increases. In further research they integrate weekly workload measures that measure symmetry based variables like force or torque asymmetry. These factors may help determine imbalances or rapid changes in workload and calculate injury trends.

References

1. Ashworth, B., Hank, M., Khaiyat, O., Coyles, G., Fallon Verbruggen, F., Zemkova, E., Zahalka, F., & Maly, T. (2025). Early Rate of Force Development and Maximal Strength at Different Positions of the Athletic Shoulder Test in Baseball Players. *Sports, 13*(9), 300.
2. García, F., Schelling, X., Castellano, J., Martín-García, A., Pla, F., & Vázquez-Guerrero, J. (2022). Comparison of the most demanding scenarios during different in-season training sessions and official matches in professional basketball players. *Biology of Sport, 39*(2), 237–244.
3. Junior, M. N. S. d. O., Veneroso, C. E., Ramos, G. P., Johnson, K. E., Guilkey, J. P., Sena, A. F. d. C., Cabido, C. E. T., & Cholewa, J. M. (2021). Distance and intensity profiles in Division I women's soccer matches across a competitive season. *Sports, 9*(5), 63.
4. Koyama, T., Nishikawa, J., Yaguchi, K., Irino, T., & Rikukawa, A. (2024). A comparison of the physical demands generated by playing different opponents in basketball friendly matches. *Biology of Sport, 41*(2), 527–533.
5. Nakatani, M., Murata, K., Kanehisa, H., & Takai, Y. (2021). Force-velocity relationship profile of elbow flexors in male gymnasts. *PeerJ, 9*, e10907.
6. Prudholme, D. C., Coburn, J. W., Lynn, S. K., & Lockie, R. G. (2022). Relationships between sprint, acceleration, and deceleration metrics with training load in Division I collegiate women's soccer players. *Journal of Human Kinetics, 85*(1), 53–62.
7. Ruiz-Rios, M., Setuain, I., Cadore, E. L., Izquierdo, M., & Garcia-Tabar, I. (2024). Physical conditioning and functional injury-screening profile of elite female soccer players: a systematic review. *International Journal of Sports Physiology and Performance, 19*(11), 1355–1366.
8. Şahbat, Y., Kütük, E., Çat, G., Ünsalan, O., Kart, H., Topkar, O. M., Baysal, Ö., & Erol, B. (2022). An unusual injury pattern: arm wrestling injury, treatment modalities, clinical outcomes, and return to sport. *Acta Orthopaedica et Traumatologica Turcica, 56*(1), 1–6.
9. Schaefer, L. V., Carnarius, F., Dech, S., & Bittmann, F. N. (2023). Repeated measurements of Adaptive Force: Maximal holding capacity differs from other maximal

strength parameters and preliminary characteristics for non-professional strength vs. endurance athletes. *Frontiers in Physiology, 14*, 1020954.

10. Spiering, B. A., Clark, B. C., Schoenfeld, B. J., Foulis, S. A., & Pasiakos, S. M. (2023). Maximizing Strength: The Stimuli and Mediators of Strength Gains and Their Application to Training and Rehabilitation. *Journal of Strength and Conditioning Research, 37*(4), 919–929.
11. Tavares, N., Vilas-Boas, J. P., & Castro, M. A. (2025). Effect of Preventive Exercise Programs for Swimmer's Shoulder Injury on Rotator Cuff Torque and Balance in Competitive Swimmers: A Randomized Controlled Trial. *Healthcare, 13*(5), 538.
12. Xiao, M., Nguyen, J. N., Hwang, C. E., & Abrams, G. D. (2021). Increased lower extremity injury risk associated with player load and distance in collegiate women's soccer. *The Orthopaedic Journal of Sports Medicine, 9*(10), 23259671211048248.