2025-2026 NFLPA Data Analytics Case Competition The Cloudscrapers

Benjamin Harris-Myers
Sarah Akiki

Word Count: 1952

Executive Summary

The National Football League Collective Bargaining Agreement (CBA), while an extensive piece of legislation, contains little detailed provisions regarding the connections between travel, game environment, and their impact on player health, injury risk, or post-injury workload. To be clear, the CBA only discusses team travel within the context of it being a reimbursable expense, an omission that seems increasingly untenable considering the league's growing travel demands and wide variation in environmental conditions across venues.

Using publicly available data from Pro Football Reference and NFL.com's injury report, this study examines the relationship between travel burden, environmental factors, prior injury history, and subsequent player workload during the 2016–2025 seasons, corresponding to the playing era of the NFL as determined by our group. We construct a cumulative workload metric combining snap counts and various on-field production stats following documented injuries, and model this outcome as a function of travel distance, time-zone changes, altitude, climate, and player production stats. Our central hypothesis is that more difficult travel conditions and dramatic environmental transitions are positively correlated with elevated injury risk and increased cumulative workload following injury, reflecting heightened physiological stress and incomplete recovery. While our hypothesis was not entirely proven, our findings do suggest a relationship between environmental stressors and injury rate and post-injury performance.

By empirically linking travel and environmental variability to post-injury workload outcomes, this study identifies travel as a structural blind spot in the current CBA and provides evidence relevant to future negotiations concerning player health, workload and roster management.

Introduction

The NFLPA released in 2020 a new CBA, which protects and advocates for its members' rights (NFL & NLPA, 2020). The improvements center around larger share of revenue, expanded rosters, increases in minimum salaries, improved benefits for both active and retired members, and enhanced health and safety protocols. Nevertheless, the NFLPA are always looking to update and improve their members' needs and rights from an ethical perspective. For this reason, we are tasked to look at how the cumulative effects of workload can impact key metrics such as player performance, injury risk, earnings or team wins and losses. After thoroughly analyzing this effect, we will be providing the exact policies and CBA provisions already put in place that are contributing to and reinforcing the insights. Lastly, we will be suggesting three priorities that the NFLPA should focus on to better support and advocate for the interests of players, which could include types of injury, injury rates or even environmental and travel factors. For our analysis, we will scrape data from the *Pro Football Reference* and *NFL.com* websites, which will allow us to study our key metrics and make improved recommendations. Additionally, the League Office's Injury Report Policy standardizes public reporting of practice participation (e.g., Did Not Participate, Limited Participation, Full Participation), which underpins the injury-report data structure used in this study (NFL, 2016)**.**

Institution and Policy Context

The CBA is generally put in place to protect and advocate for players' rights. With that in mind, we will review some of the most relevant policies that are directly tied to our area of study. Firstly, Articles 21 through 24 detail workout requirements and limits. For example, article 21 reviews offseason workouts, what players are allowed or not allowed to do as well as their rights during that period. There are three phases with differing football activities where players are only allowed to engage in a certain number of workouts per week, unless they are injured. Article 22 covers the rules for mandatory and voluntary minicamps programs, how many players a club can have, when they can happen, how long they can be, expenses as well as injury protection. Article 23 discusses rules for training camp and workload, which goes into detailed practice conduct limits and filming requirements. More details about the in-season practice limits can be found in article 24, which goes into maximum on-field time and prohibited drills. What is not covered in this section is how practices should or shouldn't change with respect to the play environment, something we look to address with our findings.

Moving onto travel requirements, article 36 explains the reimbursement that players get when they're relocated or traded, as well as travel costs and transportation requirements. As for the medical care and treatment rules, article 39 goes in depth to set the framework for care standards, physician credentials and access to medical info and/or action plan programs. Unfortunately, that is essentially all that is covered in terms of travel in the current CBA. With the results from our models, we provide solutions to fix this current blindspot for the NFLPA.

Methods

To begin, our study focuses on the years from 2016 to 2025. This was determined because 2016 was the first year the League Office released its formal Injury Report Policy, outlining the three practice-participation categories, and 2018 marks the broader shift toward mobile quarterbacks, meaning the data would be representative of the modern NFL.

For this analysis, we scraped data from the *Pro Football Reference* website in order to compare snaps before and after using weather and elevation or travel and injury.

From the pipeline output, we created a full dataset loaded with 307,636 rows after initial preparation. There are two feature groups, the pre-injury predictor features and the post-injury performance outcomes. The pre-injury predictor features capture environmental exposure, playing surface/roof context, workload, and travel burden, specifically, average weather, average humidity, most common surface, most common roof, average snaps, travel magnitude, elevation, direction and international travel. The post-injury performance outcomes capture passing, rushing, receiving, defense, special teams, and derived rate metrics. We also added average snap count to represent post-injury production for each player. Before modeling, we ran a low-level analysis in order to get the correlational coefficients between the different predictor variables. This insight would be used to identify which features to look out for and utilize in the modeling phase.

First, we employed two modeling facets for our pre-injury predictions. For each of the models we used a 80-20% train and test split. We chose to use a Random Forest Classifier due to its capacity for accommodating nonlinearities and complex interactions typical for such structured data found in football data, and we used a learned XGBoost Classifier model because of its competitiveness with its ability to consider the importance of multiple features at once. We standardized all features prior to model training. Second, for forecasting post-injury performance, we used Random Forest Regression models for each stipulated performance metric with categorical injury features encoded for injury location, practice status, and player position. Third, for a holistic descriptive analysis, we used K-Means Clustering with pre-injury feature standardization for a discrete "risk profile" feature and compared mean post-injury performance for each category.

Results and Recommendations

To begin, we present our findings of the initial low-level analysis. At first glance, we did not find any significant correlation between many of our travel and environmental variables and injury risk. However, we did find a positive correlation between previous snap counts and injury risk.

In total, of the predictor features we chose, average_snaps_before and average_weather_before had the highest correlations of 0.075 and 0.014, respectively. Given the large sample size ($N \approx 307{,}636$ observations), the correlation between average_snaps_before and injury risk was statistically significant ($p < 0.05$), despite a modest effect size. By contrast, most travel-related variables did not reach conventional significance thresholds ($p > 0.05$), suggesting limited independent linear association with injury occurrence.

Overall, this investigation pointed us in a few directions. Firstly, the biggest predictor for an injury is likely snap_count. In other words, the more snaps that a player has under their belt, the higher chance of injury. As we will touch upon later, this was supported by our models.

Secondly, the biggest environmental predictor for an injury was average_weather_before, an aggregated statistic that measured the average temperature of games before an injury was recorded, over a certain number of games per player. This finding isn't too surprising — as it is well documented that colder environments lead to higher injury rates as muscles, tendons, and joints stiffen, reducing flexibility and slowing reaction times (Paulsen)— but what was unexpected is that it was the only environmental feature that had a high correlation, going against the focus of our thesis. After modeling, we did discover other relationships with environmental features, but we assume our hypothesis to still likely be incorrect.

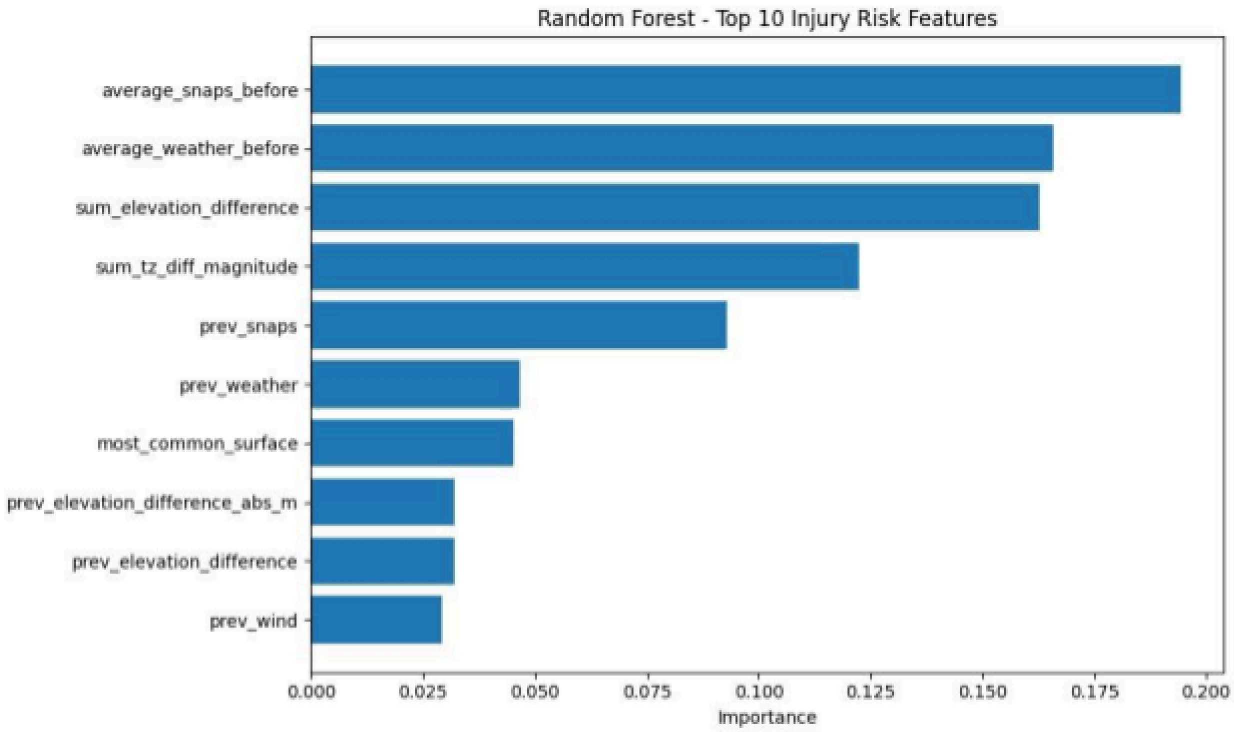
Finally, the injury rate per position and injury location tables are below. This is an area that could use more exploration in combination with the production statistic analysis we will touch on later:

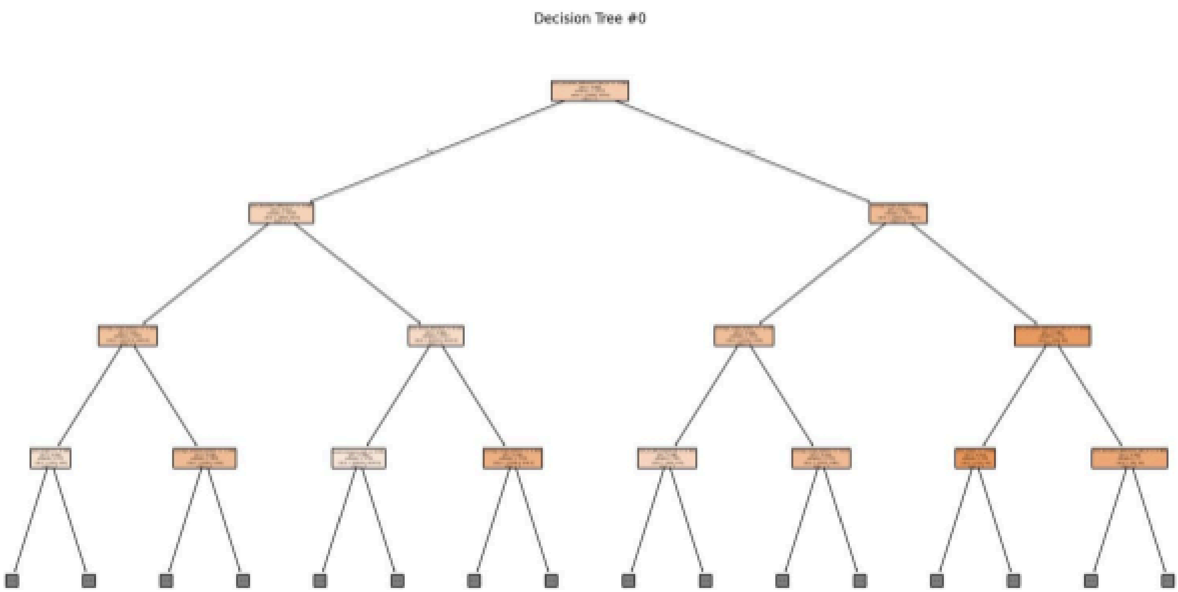
Position	**Injury Rate (%)**	**Number of Players**	**Share of Total Players (%)**
CB	56.1	6,643	15.8
K	54.5	352	0.8
S	53.9	4,420	10.5
FB	53.8	173	0.4
WR	51.5	7,267	17.3
RB	51.3	4,158	9.9
LB	50.4	7,258	17.3
T	50.0	4,769	11.4
TE	48.8	3,483	8.3

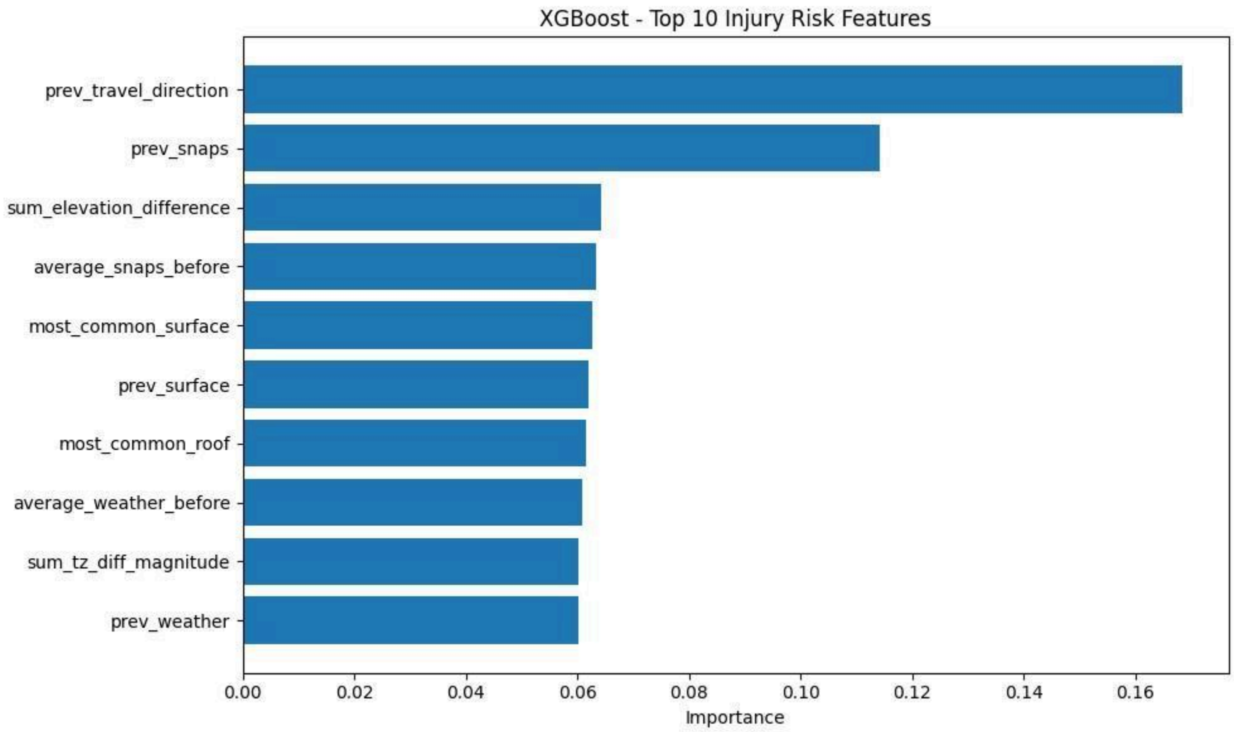
G	46.7	3,463	8.2

Injury Location	**Number of Injuries**	**Share of Total Injuries (%)**
Knee	4,234	22.1
Ankle	3,726	19.4
Hamstring	2,923	15.2
Shoulder	1,651	8.6
Concussion	1,600	8.3
Groin	1,139	5.9
Foot	1,131	5.9
Calf	974	5.1
Back	930	4.8
Illness	891	4.6

Moving on to our modeling results, our primary main finding is that snaps_played had the highest and second highest significance in our RandomForest and XGBoost models, respectively. Both outcomes are presented in the graphs below:


Random Forest - Top 10 Injury Risk Features
average_snaps_before
average_weather_before
sum_elevation_difference
sum_tz_diff_magnitude
prev_snaps
prev_weather
most_common_surface
prev_elevation_difference_abs_m
prev_elevation_difference
prev_wind
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
0.200
Importance

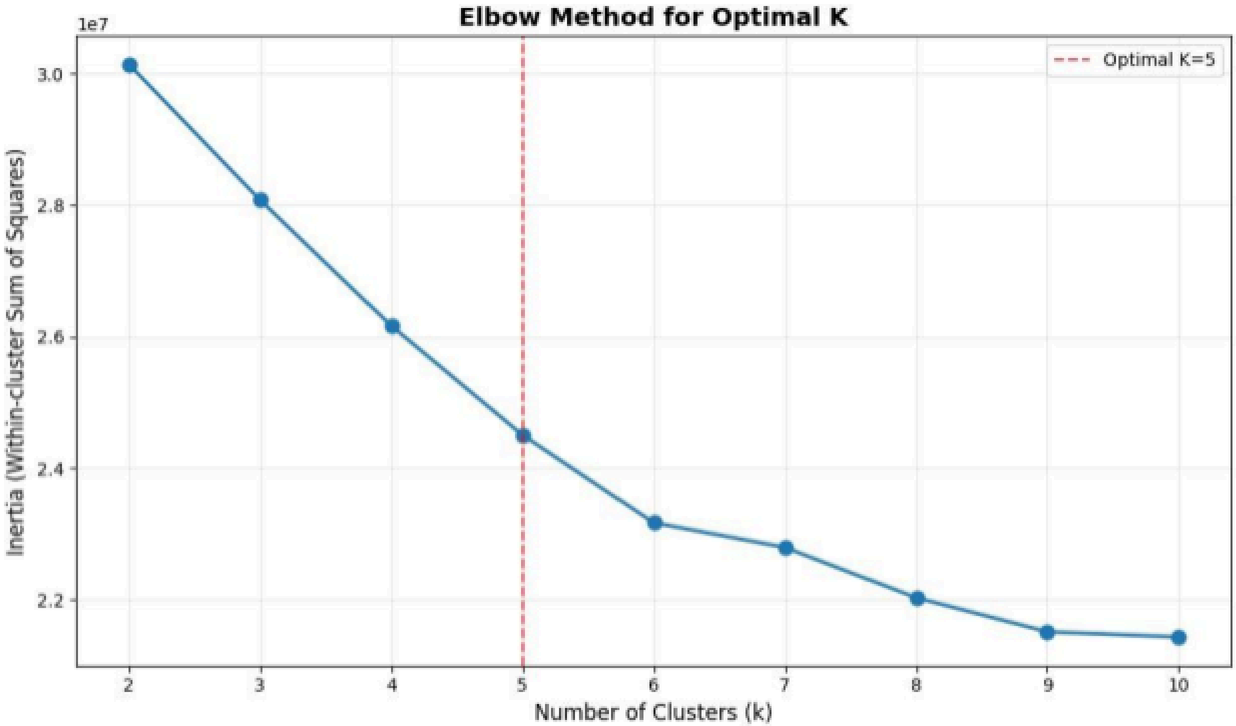

Decision Tree #0

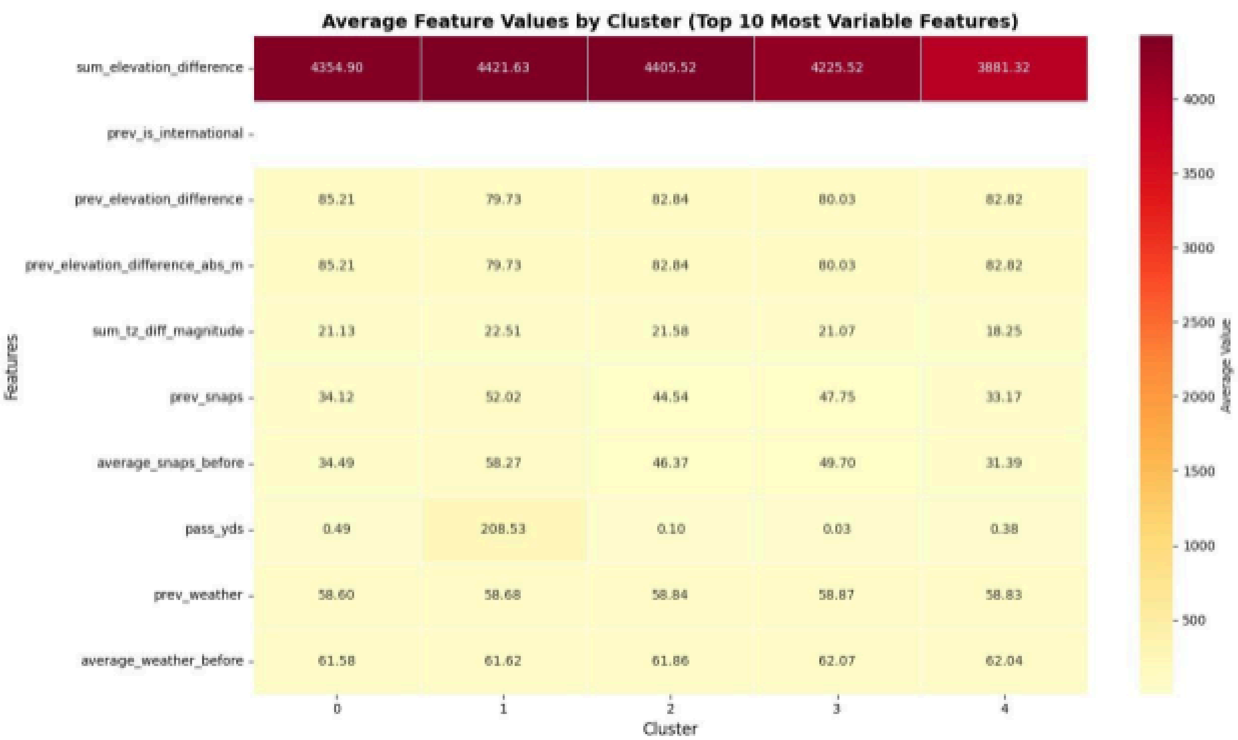


Average_weather_before did show up as significant in our Random Forest model, backing our initial analysis, but did not show up as all too important in our XGBoost model. Interestingly, sum_elevation_difference is the only other feature to score in the top three of significance for both models. This suggests that dramatic shifts in elevation are also a key predictor of injury, and the CBA should do more to combat elevation-related stress.

One actionable measure that the NFLPA can employ in response to these issues is to increase the squad size of an NFL roster, as well as encourage more rotation of players by teams. Ultimately, if playing more snaps leads to a higher number of injuries, we should encourage teams to play their players less per game, and one way of accomplishing that is through allowing teams to have more players on a roster, allowing for more rotation. Furthermore, an extended roster would prevent the most common issues with elevation/altitude (our second most significant feature overall). According to a 2015 study out of SportsHealth Journal the most essential component of injury and illness prevention at high altitude is individualized and appropriate acclimatization. (Khodaee, Morteza, et al.) Extending the roster size will inherently allow for players to rest more often when adapting to dramatic shifts in altitude.

Finally, after running our K Means model to track different profiles of injuries, sum_elevation_difference turned out to be the most valuable feature when defining a centroid when K=5 by a significant margin. K was determined to be 5 using the elbow method.


Elbow Method for Optimal K
1e7
3.0
2.8
2.6
2.4
2.2
Inertia (Within-cluster Sum of Squares)
Optimal K=5
2
3
4
5
6
7
8
9
10
Number of Clusters (k)


Average Feature Values by Cluster (Top 10 Most Variable Features)
sum_elevation_difference
4354.90
4421.63
4405.52
4225.52
3881.32
prev_is_international
prev_elevation_difference
85.21
79.73
82.84
80.03
82.82
prev_elevation_difference_abs_m
85.21
79.73
82.84
80.03
82.82
sum_tz_diff_magnitude
21.13
22.51
21.58
21.07
18.25
prev_snaps
34.12
52.02
44.54
47.75
33.17
average_snaps_before
34.49
58.27
46.37
49.70
31.39
pass_yds
0.49
208.53
0.10
0.03
0.38
prev_weather
58.60
58.68
58.84
58.87
58.83
average_weather_before
61.58
61.62
61.86
62.07
62.04
Features
0
1
2
3
4
Cluster
4000
3500
3000
2500
2000
1500
1000
500
Average Value

A second way for the NFLPA to respond to the issues of snap count and elevation is to increase the time of practices, specifically the time dedicated for warm ups and non-contact play. The current CBA heavily regulates the time players can practice and even mandates specific days off. We are not suggesting that the CBA cut back entirely in this area, as shortened practice time has other reasons for existing, including the prevention of injuries. Instead, we suggest that teams should have more time tacked on to the beginning or ends of practices so that they are able to properly warm up and cool down before and after intense exercise (another way to ease the adaptation to higher altitudes and intense snap workloads). This includes both practices during the weeks and practices before games (both of which are regulated by the CBA).

Beyond injury occurrence, this study examined how injuries relate to post-injury performance outcomes. Of the 111 post-injury performance features analyzed, 26 exhibited relatively strong correlations with injury location, including passing efficiency (e.g., pass_rating, pass_air_yards), defensive production (e.g., def_int_yds, def_cmp), ball security (e.g., fumbles_forced), and special teams involvement (e.g., kick and punt return metrics). Most remaining features were primarily correlated with player position rather than injury characteristics. This distinction suggests that certain injury locations may impair specific functional demands, warranting more targeted recovery protocols and workload management. Future research should identify which injury locations most strongly affect these performance domains and advocate for enhanced medical oversight accordingly. (see Appendix 1)

Conclusion

This report explores travel and stadium environment variables and their link with injury risk and post-injury workload outcomes. We built a dataset that highlights the relationship between pre-injury exposure (travel distance, time-zone change, altitude, and weather) to workload and performance proxies. With the help of descriptive statistics and predictive modeling, we pinpointed the most relevant factors that relate to injuries and post-injury usage.

Most notably, cumulative workload, proxied by snap count, emerged as the strongest and most consistent predictor of injury risk across both Random Forest and XGBoost models. This reinforces evidence that sustained on-field exposure, as opposed to isolated conditions as our hypothesis suggests, is the primary driver of injury risk. From a policy perspective, this supports advocating for expanded roster sizes and greater player rotation, enabling teams to distribute snaps more evenly and reduce overuse-related injuries. Additionally, elevation change consistently ranked among the most important non-workload predictors in both modeling and clustering analyses. This suggests that abrupt altitude shifts introduce physiological stress that is not currently addressed by the CBA. Given established medical research on acclimatization, the NFLPA could reasonably push for additional rest, medical accommodations, or roster flexibility when teams face significant elevation changes.

In conclusion, the NFL constantly places players in situations that affect fatigue, recovery and readiness, whether it’s from long-distance travel, rapid time-zone shifts, or wide variation in altitude and climate. Sports medicine research has long documented that time-zone transitions can produce travel fatigue/jet lag via circadian misalignment and sleep disruption, mechanisms that might affect recovery and readiness in athletes (Waterhouse et al., 2007). Nevertheless, the Collective Bargaining Agreement (CBA) treats travel exclusively as reimbursable logistics rather than as a stressor. Therefore, greater attention needs to be given to travel and environmental factors, especially as the league expands internationally and competitive advantage is given to teams who better control recovery and preparation.

Citations:

NFL & NFLPA. *March 15, 2020 NFL–NFLPA Collective Bargaining Agreement: Final Executed Copy.* 2020. NFLPA Blob Storage

National Football League. *2016 Personnel (Injury) Report Policy.* June 7, 2016. NFL Football Operations

Khodaee, Morteza, et al. "Athletes at High Altitude." *Sports Health*, U.S. National Library of Medicine, 2016, pmc.ncbi.nlm.nih.gov/articles/PMC4789936/.

Paulsen, Emily. "Exercising in Cold Weather: What You Need to Know to Prevent Injury." Mission Health, 2024, www.missionhealth.org/healthy-living/blog/exercising-in-cold-weather-what-you-need-to-know-to-prevent-injury.

Waterhouse, J., Reilly, T., Atkinson, G., & Edwards, B. (2007). Jet lag: trends and coping strategies. *The Lancet, 369*(9567), 1117–1129. https://doi.org/10.1016/S0140-6736(07)60529-7

Appendix:

1. Post-Injury Production Statistics and Feature Importance

Target	Feature 1	Importance 1	Feature 2	Importance 2	Feature 3	Importance 3
pass_cmp_x	position	0.875583	injury_location	0.100803	practice_status	0.023615
pass_att	position	0.878101	injury_location	0.098919	practice_status	0.022979
pass_yds	position	0.870451	injury_location	0.104551	practice_status	0.024998
pass_td	position	0.846026	injury_location	0.124745	practice_status	0.029229
pass_int	position	0.840288	injury_location	0.131443	practice_status	0.028269
pass_sacked_x	position	0.867715	injury_location	0.108568	practice_status	0.023716
pass_sacked_yds	position	0.87374	injury_location	0.103643	practice_status	0.022617
pass_long	position	0.874491	injury_location	0.102292	practice_status	0.023216
pass_rating	injury_location	0.748167	practice_status	0.198994	position	0.052838
rush_att_x	position	0.899937	injury_location	0.065532	practice_status	0.034531
rush_yds_x	position	0.889639	injury_location	0.076878	practice_status	0.033482
rush_td_x	position	0.824739	injury_location	0.130835	practice_status	0.044426
rush_long	position	0.897249	injury_location	0.066717	practice_status	0.036034

targets_x	position	0.879867	injury_location	0.096316	practice_status	0.023817
rec_x	position	0.847765	injury_location	0.121001	practice_status	0.031235
rec_yds_x	position	0.874408	injury_location	0.099563	practice_status	0.026029
rec_td_x	position	0.749185	injury_location	0.179687	practice_status	0.071128
rec_long	position	0.865882	injury_location	0.097438	practice_status	0.03668
fumbles	position	0.819389	injury_location	0.150581	practice_status	0.030029
fumbles_lost	position	0.741242	injury_location	0.214322	practice_status	0.044436
def_int_x	position	0.781459	injury_location	0.174646	practice_status	0.043896
def_int_yds	injury_location	0.486825	position	0.421778	practice_status	0.091397
def_int_td	injury_location	0.74298	position	0.189086	practice_status	0.067934
def_int_long	injury_location	0.492825	position	0.41723	practice_status	0.089945
pass_defended	position	0.893355	injury_location	0.077129	practice_status	0.029516
sacks_x	position	0.747113	practice_status	0.140463	injury_location	0.112423
tackles_combined_x	position	0.754571	practice_status	0.16477	injury_location	0.080658
tackles_solo	position	0.868311	injury_location	0.077271	practice_status	0.054419
tackles_assists	position	0.737718	practice_status	0.160732	injury_location	0.101549
tackles_loss	position	0.735299	practice_status	0.170829	injury_location	0.093873
qb_hits	position	0.781211	practice_status	0.13635	injury_location	0.082439
fumbles_rec	position	0.763815	injury_location	0.189651	practice_status	0.046534
fumbles_rec_yds	position	0.446079	injury_location	0.42436	practice_status	0.129561
fumbles_rec_td	injury_location	0.6137	practice_status	0.194892	position	0.191408
fumbles_forced	injury_location	0.499138	position	0.376864	practice_status	0.123998
xpm	injury_location	0.685582	position	0.192434	practice_status	0.121984
xpa	injury_location	0.659309	position	0.234876	practice_status	0.105815
fgm	injury_location	0.645056	position	0.223388	practice_status	0.131556
fga	injury_location	0.653747	practice_status	0.197456	position	0.148797
punt	position	0.762565	practice_status	0.158971	injury_location	0.078464
punt_yds	position	0.762981	practice_status	0.160081	injury_location	0.076937
punt_yds_per_punt	position	0.458128	injury_location	0.315107	practice_status	0.226765
punt_long	position	0.828341	injury_location	0.110089	practice_status	0.06157
kick_ret	position	0.464974	injury_location	0.434509	practice_status	0.100517
kick_ret_yds	position	0.517067	injury_location	0.371049	practice_status	0.111884
kick_ret_yds_per_ret	injury_location	0.650411	position	0.180802	practice_status	0.168787

kick_ret_td	injury_location	0.686896	position	0.208107	practice_status	0.104997
kick_ret_long	position	0.495426	injury_location	0.401792	practice_status	0.102782
punt_ret	position	0.682079	injury_location	0.26087	practice_status	0.057051
punt_ret_yds	position	0.615084	injury_location	0.321519	practice_status	0.063397
punt_ret_yds_per_ret	injury_location	0.666232	practice_status	0.201283	position	0.132485
punt_ret_td	injury_location	0.663493	practice_status	0.188464	position	0.148043
punt_ret_long	position	0.612175	injury_location	0.319337	practice_status	0.068488
pass_first_down	position	0.876104	injury_location	0.096852	practice_status	0.027043
pass_first_down_pct	position	0.8811	injury_location	0.092833	practice_status	0.026066
pass_target_yds	position	0.86021	injury_location	0.110159	practice_status	0.029632
pass_tgt_yds_per_att	position	0.883635	injury_location	0.089551	practice_status	0.026814

pass_air_yds	injury_location	0.659443	position	0.232741	practice_status	0.107816
pass_air_yds_per_cmp	position	0.877674	injury_location	0.097794	practice_status	0.024532
pass_yac	position	0.88435	injury_location	0.088351	practice_status	0.027299
pass_drops	position	0.887778	injury_location	0.087878	practice_status	0.024344
pass_drop_pct	position	0.87615	injury_location	0.097294	practice_status	0.026555
pass_poor_throws	position	0.877791	injury_location	0.097235	practice_status	0.024974
pass_poor_throw_pct	position	0.876864	injury_location	0.09683	practice_status	0.026307
pass_sacked_y	position	0.893699	injury_location	0.082795	practice_status	0.023505
pass_hurried	position	0.841746	injury_location	0.127296	practice_status	0.030958
pass_hits	injury_location	0.619028	position	0.200749	practice_status	0.180222
pass_pressured	position	0.839286	injury_location	0.133561	practice_status	0.027153
pass_pressured_pct	position	0.797295	injury_location	0.147184	practice_status	0.055521
rush_scrambles	injury_location	0.554762	position	0.232353	practice_status	0.212885
rush_scrambles_yds_per_att	position	0.867506	injury_location	0.097312	practice_status	0.035182
rush_first_down	position	0.484312	injury_location	0.410211	practice_status	0.105477
rush_yds_before_contact	position	0.858016	injury_location	0.120489	practice_status	0.021495
rush_yds_bc_per_rush	position	0.775429	injury_location	0.186325	practice_status	0.038246
rush_yac	position	0.850947	injury_location	0.118765	practice_status	0.030289
rush_yac_per_rush	position	0.889085	injury_location	0.088142	practice_status	0.022773
rush_broken_tackles	position	0.868606	injury_location	0.106241	practice_status	0.025154
rush_broken_tackles_per_rush	position	0.624938	injury_location	0.309895	practice_status	0.065167

rec_first_down	position	0.624569	injury_location	0.287215	practice_status	0.088216
rec_air_yds	position	0.885442	injury_location	0.091384	practice_status	0.023173
rec_air_yds_per_rec	position	0.569875	injury_location	0.335815	practice_status	0.09431
rec_yac	position	0.526772	injury_location	0.354898	practice_status	0.11833
rec_yac_per_rec	position	0.516818	injury_location	0.388923	practice_status	0.094259
rec_adot	injury_location	0.535692	position	0.310161	practice_status	0.154147
rec_broken_tackles	position	0.761923	injury_location	0.186638	practice_status	0.05144
rec_broken_tackles_per_rec	injury_location	0.681415	practice_status	0.213278	position	0.105308
rec_drops	position	0.932242	injury_location	0.05669	practice_status	0.011068
rec_drop_pct	position	0.907494	injury_location	0.064605	practice_status	0.027901
rec_pass_rating	position	0.868951	injury_location	0.103672	practice_status	0.027376
def_targets	position	0.663518	injury_location	0.250997	practice_status	0.085485
def_cmp	position	0.416451	injury_location	0.413147	practice_status	0.170402
def_cmp_perc	position	0.720329	injury_location	0.209928	practice_status	0.069743
def_cmp_yds	injury_location	0.630077	position	0.197107	practice_status	0.172816
def_yds_per_cmp	position	0.88623	injury_location	0.092843	practice_status	0.020927
def_yds_per_target	position	0.905327	injury_location	0.071686	practice_status	0.022986
def_cmp_td	position	0.465024	injury_location	0.405962	practice_status	0.129014
def_pass_rating	position	0.845637	injury_location	0.105329	practice_status	0.049034
def_tgt_yds_per_att	position	0.832132	injury_location	0.101961	practice_status	0.065906
def_air_yds	position	0.824692	injury_location	0.096869	practice_status	0.07844
def_yac	position	0.87279	injury_location	0.069584	practice_status	0.057625
blitzes	injury_location	0.489121	position	0.380193	practice_status	0.130686
qb_knockdown	position	0.752757	injury_location	0.199495	practice_status	0.047749