PLAYER HEALTH & SAFETY

Risk Mitigation Model Guide
NFL Digital Athlete

Updated July 2024

Table of Contents

Application Overview ***3***

Athlete Risk & Load Table 4

Positional Risk 5

Athlete Details 7

Risk Mitigation Model Overview ***8***

Injury Types 8

Model Prediction Mode 9

Model Feature Selection 9

Risk Mitigation Plots 10

Feature Effects Plots 11

Feature and Load Statistic Definitions ***13***

Positional Velocity Bands 14

Acceleration/Deceleration Bands 15

Appendix ***16***

Preseason Feature Importance and Feature Effects Plots 16

Regular Season Feature Importance and Feature Effects Plots 19

Application Overview

The purpose of the risk portion of the Digital Athlete is to identify athletes who are at increased risk for hamstring or adductor strain due to workload and to identify specific interventions that can be applied in the near term to reduce their risk.

The app operates in two different "modes" depending on the time within the season. During the **Preseason**, an athlete's risk for a given day is calculated to express the athlete's risk of hamstring or adductor strain tomorrow, given the training and game load he has experienced during the season up to and including today. Once the **Regular Season** starts, an athlete's risk for a given day is calculated to express the athlete's risk of hamstring or adductor strain for the following game, given their load up to and including the current day.

The risk of injury for a specific athlete on any given day of the season is low, but there are substantial differences in the risk levels between athletes on a single day, even within the same position. To simplify the expression of risk for the NFL population, the risk values are normalized into **risk factors**, or how much more or less at risk a specific player is relative to the NFL population as a whole and the population of NFL athletes within his same position. The definitions of the two normalized risk values are:

- **NFL Risk Factor (NFL RF):** An NFL athlete at the 50th percentile for daily risk of injury across the NFL has an NFL Risk Factor of 1.0. An athlete with an NFL Risk Factor of 1.5 has a 50% higher risk of injury than a player with an NFL Risk Factor of 1.0. An athlete with an NFL Risk Factor of 2.0 has twice the risk of injury of an athlete with an NFL Risk Factor of 1.0. NFL Risk Factors are capped at 5.0, so a risk factor of 5.0 reflects a player who has at least five times the risk of a typical player.

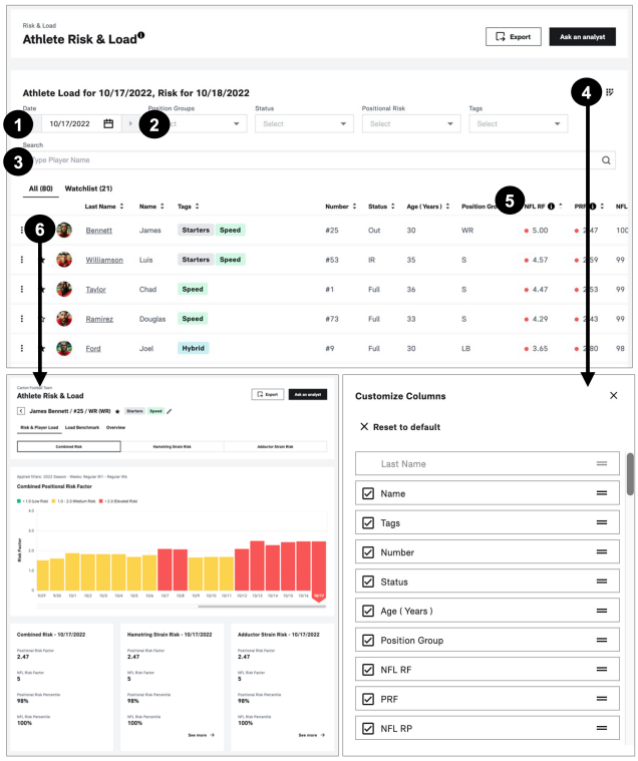
- **Positional Risk Factor (PRF):** A CB at the 50th percentile for daily risk of injury across all cornerbacks has a Positional Risk Factor of 1.0. A CB with a Positional Risk Factor of 1.5 has a 50% higher risk of injury than a CB with a Positional Risk Factor of 1.0. Positional Risk Factors are also capped at 5.0, so a Positional Risk Factor of 5.0 reflects a player who has at least five times the risk of a typical player at their position.

Due to the considerable difference in strain risk between positions, positional risk factors can differ greatly from NFL risk factors. For example, a CB with a Positional Risk Factor of 1.0 has an NFL Risk Factor of approximately 1.7, reflecting that the average cornerback has about 70% more risk for hamstring or adductor strains than the average NFL player. Conversely, an OL player with a Positional Risk Factor of 1.0 has an NFL Risk Factor of 0.4, reflecting that offensive linemen have a reduced risk of hamstring or adductor strain.

NFL Risk Factors provide the ability to quickly identify the most at-risk athletes across your team, regardless of position. Positional Risk Factors provide the ability to promptly identify the athletes on your roster at each position who are at increased risk relative to their peers at the same position, facilitating discussions with positional coaches.

Athlete Risk & Load Table

The Athlete Risk & Load Table allows you to quickly review the status of all athletes on your active roster for a particular day. The diagram below describes a workflow for rapidly identifying athletes most at risk and pinpointing recommended interventions for reducing their risk. It should be noted that daily load values will be shown as 0 until the performance data collected by your club is populated into the database. This should occur within one hour of syncing data through your performance vendor.



1. Select a date to review records for athletes on your team. If you select a date that is not populated in the database (a date in the future or a date in the off-season), no table will be displayed. If the date is unable to be selected, expand the weeks range filter to include the desired date. The default date is today.

2. If needed, you may filter by position, status, positional risk factor, or other player tags.

3. Alternatively, you may search for a specific athlete by name.

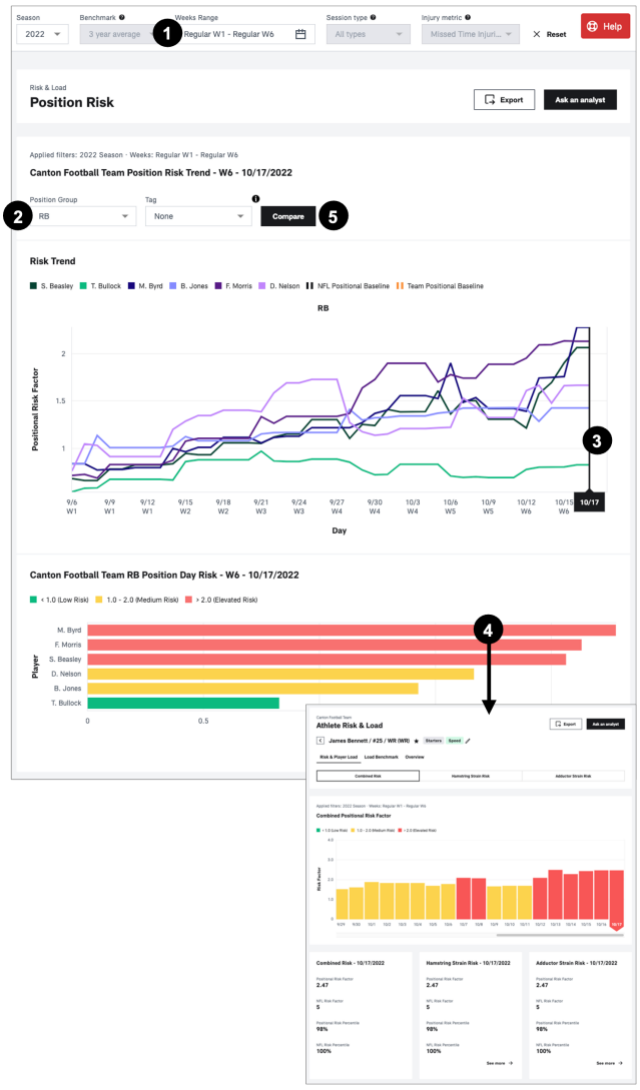
4. Columns may be customized based on your user preference. By default, NFL and positional risk factors and percentiles will show combined risk. You may add values for hamstring strain or adductor strain specifically to the table.

5. Sort the athletes by any column. The default sort is in decreasing order of NFL Risk Factor.

6. Click on a player's icon or last name to view that player's risk details for the day (see the **Athlete Details** section).

Positional Risk

The Positional Risk page provides the opportunity to review the risk history of all athletes in the selected position. This facilitates quickly identifying athletes within each position who are or have been at increased risk of adductor or hamstring strain, and those whose risk is trending up.



1. The weeks range selector at the top controls the temporal window.

2. Select a position to view. If needed, narrow down players using the Tag filter.

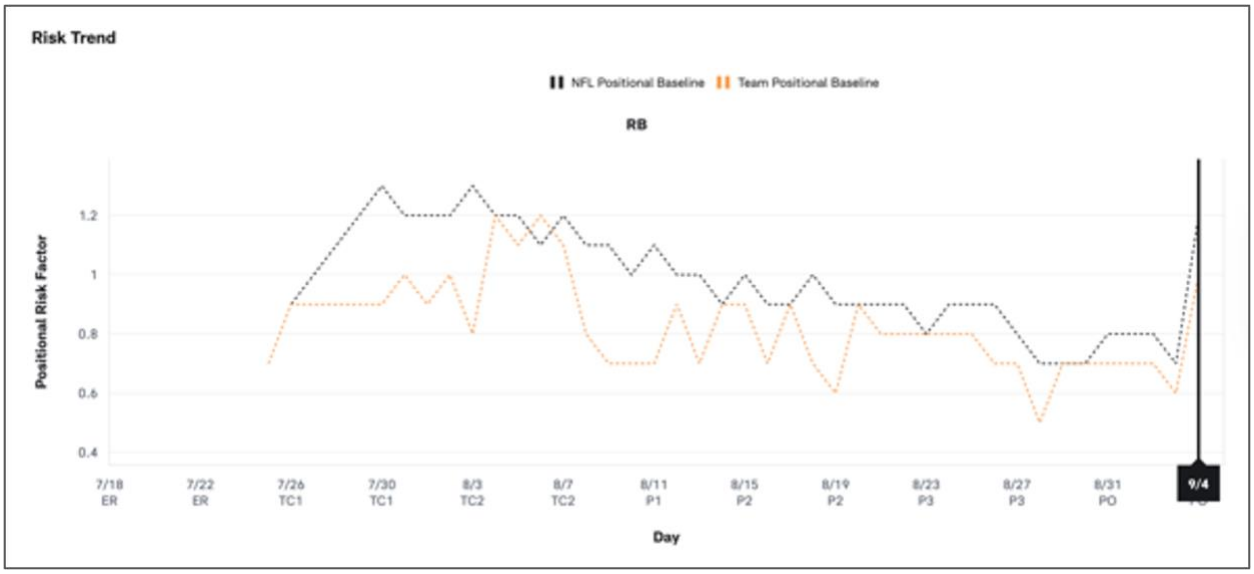
3. Hover over the chart to view data for each athlete by day. Clicking on a date updates the bar plot at the bottom to reflect positional risk factors for the selected date.

4. Click on a player's bar to view that player's risk details for the day (see the **Athlete Details** section).

5. You may compare risk between position groups using the "Compare" button.

A positional risk value of 1.0 denotes the median risk value for a position group over the entire Preseason or Regular Season. However, risk can also vary based on the time during the season. For example, most position groups have a higher risk of strain injury during training camp than the rest of the preseason (Figure 1). The positional risk plots provide the average positional risk value for the NFL position cohort for each day, calculated based on days since veteran report to account for varying schedules. Athletes above this line are at higher-than-expected risk for the time period, while those below are at lower than median risk for their position at that time period. These plots also show your club's average risk value to indicate whether your players are more or less at risk than their positional peers across the league.

Figure 1. Average positional risk of running backs during Preseason



Athlete Details

The Athlete Details page provides an overview of a given player's load and risk history. Specifically, it shows:

- How an athlete's risk is changing over time
- Risk factors and percentiles for hamstring strain, adductor strain, or a combined risk of both strains
- Load features that contribute to the athlete's risk of strain for a day, providing opportunities for intervention to reduce risk
- Load history for the athlete for the selected period, defaulting to those metrics most increasing his risk

1. The athlete's combined risk history is displayed for the selected Weeks Range.

2. Clicking on a bar in the Positional Risk Factor chart will update the risk factors and percentiles for that specific day. Values are shown for combined risk of hamstring or adductor strain, and for hamstring strain risk and adductor strain risk separately.

3. Clicking Hamstring Strain Risk or Adductor Strain Risk navigates to a player's risk history for that specific injury type. The Positional Risk Factor history chart will be updated to reflect risk of a single injury.

4. By selecting a bar in the Positional Risk Factor history chart, the player's risk feature contribution plot for the selected date will open. This plot shows the percent contribution each model feature contributes to the player's risk of injury that day, providing opportunities for possible intervention. See the **Risk Mitigation Plots** section for more details

5. Player load histories for the selected range are shown. Metrics default to those most increasing a player's risk but may be customized.

Risk Mitigation Model Overview

This section describes the structure of the machine learning model, explains the feature selection process used to populate the model, gives a brief visual summary of the process by which probabilities are assigned, and offers a guide on interpreting the risk mitigation plots.

Daily risk estimates are calculated each day using a machine learning model that predicts an athlete's probability of injury tomorrow (during preseason) or for the following game (during the regular season), using information on the athlete's training and competitive loads during the season up to and including today. An athlete's load history starts from the 1st day of preseason early report or the athlete's first day of participation at your club. The selected machine learning model is an extension of a decision tree model that employs numerous decision trees to handle the highly non-linear and complex task of assigning probabilities to these rare events. The **Feature Effects Plots** section below provides a visual explanation summarizing how the model assigns probabilities for athlete risk each day. As discussed in the **Application Overview** section above, these probabilities are converted into normalized risk factors for ease of interpretation.

Injury Types

Within Digital Athlete, players are identified as being at increased risk for hamstring or adductor strain due to their workload. While a single risk value is expressed, this risk consists of the results of two separate models: a model predicting an athlete's risk of a hamstring strain and a separate model predicting an athlete's risk of an adductor strain. Both models contain the same features; however, given the different injury mechanisms, how the models use these features differs between hamstring and adductor strains. The plots in the **Appendix** illustrate the relative importance of the features in the preseason and regular season models when applied to the entire NFL population. Features at the top of the plot contribute more to the risk scores of athletes than those at the bottom. While 4-day average acceleration distance is the most important load feature for predicting hamstring strains in the preseason, it is only the fourth most important feature for predicting preseason adductor strains.

To simplify reporting in Digital Athlete, the predictions from both models are combined into a single value to express the risk of either injury. Model predictions were combined assuming values were independent:

$$P_{ham\ or\ add} = 1 - (1 - P_{ham})(1 - P_{add})$$

Most instances in the application show this combined prediction; however, specific predictions for hamstring or adductor strain are available.

Model Prediction Mode

As mentioned above, the application operates in two different "modes" depending on the time within the season:

- **Preseason mode:** The app operates in this mode starting after the 4-day on-field acclimation period through the final week of preseason games. While in this mode, the application predicts an athlete's risk for a given day and is calculated to express the athlete's risk of hamstring or adductor strain **tomorrow,** given the training and game load he has experienced during the season **up to and including today**. Due to the consistent loading of players throughout the preseason, this target is suitable for managing an athlete's load.

- **Regular Season mode:** Once the regular season starts, the app will switch to operating in this mode, and predictions will be made to express the athlete's risk of hamstring or adductor strain for **the following game**, given their load up to and including the current day. This change is necessary due to the cyclical nature of the regular season.

Model Feature Selection

Table 1. Model Feature Selection Summary. Feature types appearing in the models are indicated in gray

Type	Base Feature	Temporal Window				
		Fixed	Cumulative	Daily	4-Day	Acute:Chronic
Non-Modifiable	Position					
Non-Modifiable	Age					
Schedule	Training Days					
Schedule	Sequential Days					
Duration	Duration					
Intensity	Density (yd/min)					
Distance	Distance					
Speed	Speed Zones					
Contact	Contact					
Acceleration	Accel/Decel Zones					

To select the features used in the predictive model, we first consulted a focus group of NFL athletic trainers, sports scientists, and strength & conditioning coaches to identify load features that teams routinely track. Critically, we asked them to identify load features they tracked daily for which, when flagged for adding risk, there was a known intervention that could be employed to reduce the athlete's risk. Those features were all defined as a type listed in the Base Feature column in Table 1 above. Note that within a base feature description, multiple metrics were considered. For example, for Speed Zones, teams conveyed that they used high-speed running (Zone 4, 60% max speed), sprinting (Zone 5, 85% max speed), or a combination of both. Some clubs tracked sprint efforts, while others focused on sprint distance. All variations of these metrics were considered for inclusion in the model.

Once a comprehensive list of base features was assembled, we applied those features over various types of temporal windows, illustrated in in Table 1. Daily features represent the acute load for the day's training session. The 4-day window features included each metric's sum and average (when training) over the immediately preceding four days. Acute:chronic features represent the ratio of the athlete's acute load to the athlete's average daily training load, computed using an exponentially decaying average so that changes in athlete capacity over time due to adaption were considered. We then ran a machine learning feature selection process to identify an ideal combination of features for daily intervention. The selected feature types are depicted in gray in the chart above, and definitions for each of the selected features are available below in the **Feature and Load Statistic Definitions** section and the **Digital Athlete Glossary** on the Digital Athlete application.

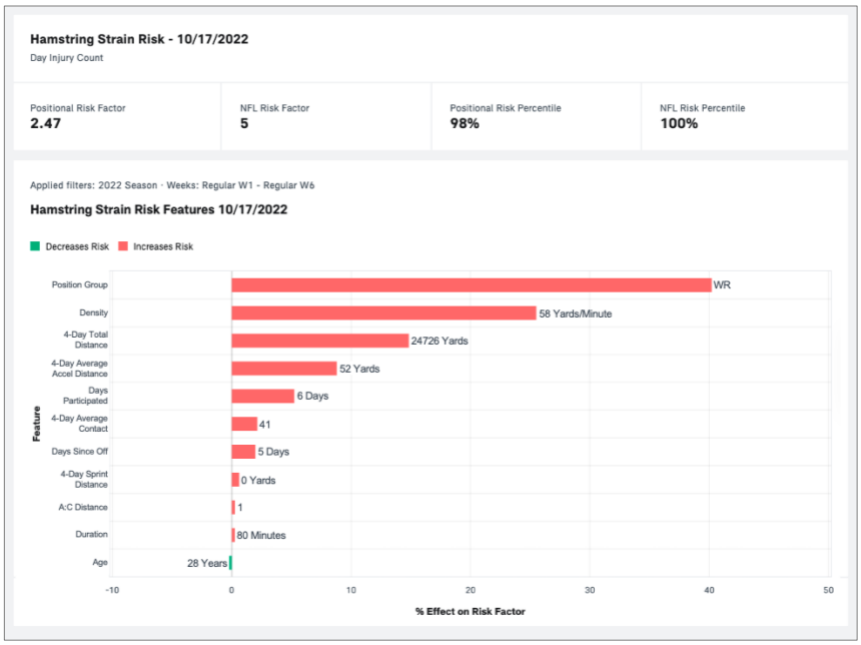
When in regular season mode, the model leverages additional longer-range features over a 21-day window. These features provide a view into whether athletes are being worked above or below their typical loading.

Risk Mitigation Plots

The values in the Risk Features plot represent the risk effects, as discussed below, for a specific athlete day. The risk effects are plotted as a percentage of the overall difference between this athlete day and the 50th percentile NFL athlete, who has an NFL Risk Factor of 1.0. Bars in red are features contributing to increased risk, while green bars indicate reduced risk. For an athlete with an NFL Risk Factor of exactly 1.0, the sum of the red and green bars will add up to 0. It should be noted that the bars represent the relative contribution to that player's deviation from average risk. Therefore, a large value does not necessarily indicate that the feature significantly impacts the player's overall risk.

Large effects that add risk provide opportunities for intervention. The athlete depicted below (Figure 2) has a high-risk factor due to large values for density, 4-day total distance, and 4-day average acceleration distance. This indicates that the model believes this athlete is in an overloaded state and could benefit from reduced overall distance and acceleration in the next few practices.

Figure 2. Example Risk Mitigation Plot for an NFL Athlete at Increased Risk

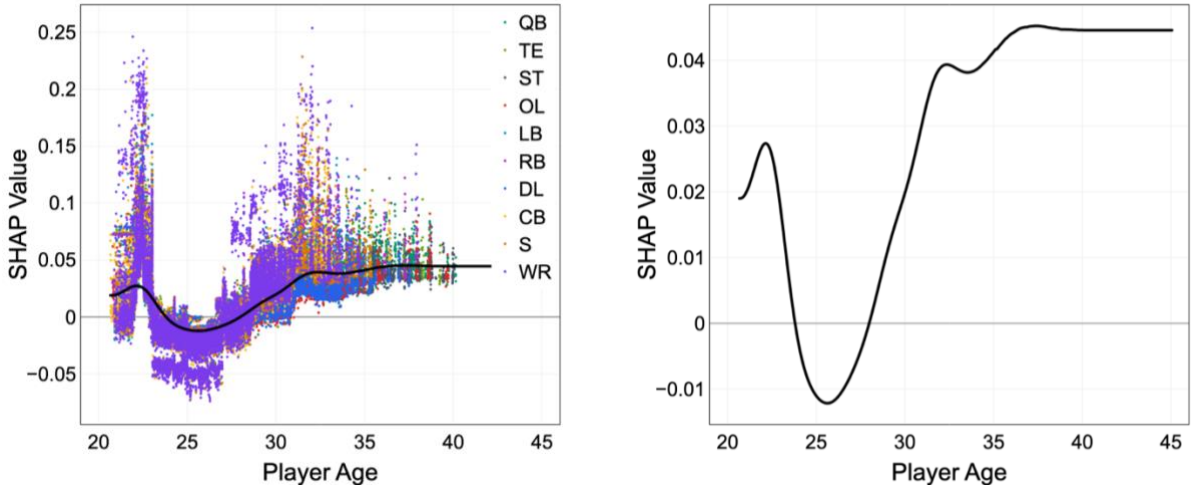


Feature Effects Plots

After exploring the Risk Features plots to identify possible risk-reducing interventions, the next question is how to modify the player's load. The selected machine learning model is designed to perform well for decisions involving features that have complex interactions. It does not independently add risk on a feature-by-feature basis as a regression model does, but instead it assigns risk by simultaneously considering all features for a given athlete using a large ensemble of decision trees. It is still possible, however, to visualize the model assignments to understand the decision-making process and to identify interventions to influence a player's risk.

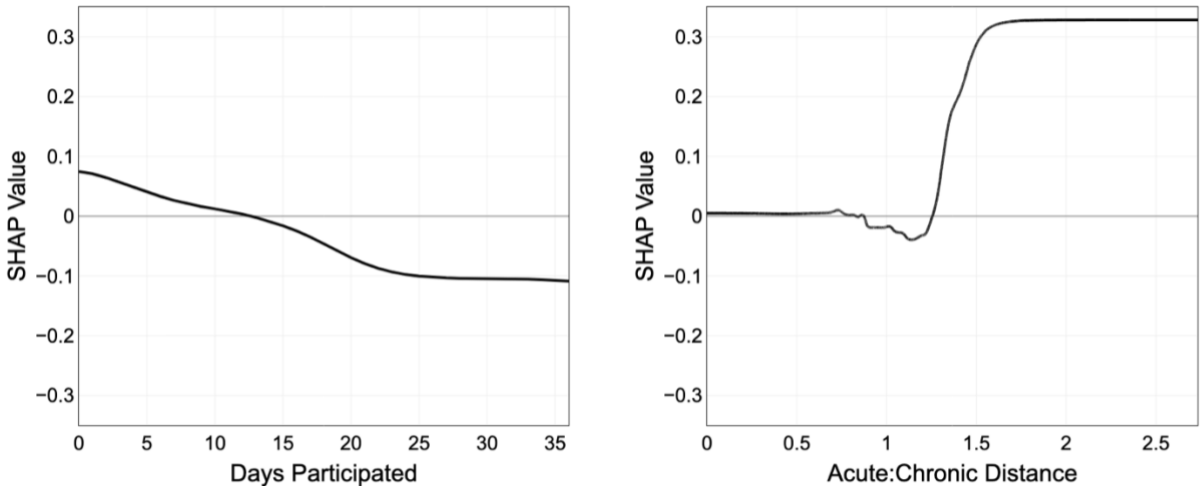
Consider Figure 3 below, in which the plot on the left shows the effect on risk assigned to every athlete in the NFL based on their age over multiple seasons. Each point on the graph represents an athlete's day, with the player's age on the x-axis and the y-axis depicting the change in log-odds. For a single age, there is variation in the assigned risk due to interaction with other features. The graph on the right depicts the smoothed average value at each observed age. When this line is above 0, it indicates that, on average, the model assigns higher risk to the athlete; when below 0, the model (on average) assigns reduced risk at that age. The interpretation of the effect from these plots is that athletes between the ages of 24 and 28 are at reduced risk, while athletes over the age of 24 and under the age of 24 are at the highest risk.

Figure 3. Feature Effects Plot for Age for NFL Cohort



Two more examples of feature effects are provided below (Figure 4). At high values, these features contribute more to a player's risk value than the player's age (note the y-axis scale difference on Figure 4 below compared to Figure 3 above). The left plot shows that athletes are at increased risk during the first ten days of their participation, with risk decreasing steadily by the day. For most athletes, this encompasses the acclimation period and continues into the first few padded practices. The pattern for Acute:Chronic Distance, on the right, is somewhat different. Risk remains relatively unchanged for Acute:Chronic Distance ratios up to 0.9. Athlete days with A:C Distance ratios between 0.9 and 1.25 are slightly reduced, and then risk begins to climb as an athlete's A:C Distance ratio increases.

Figure 4. Average Feature Effects for the NFL Cohort for Selected Features



The **Appendix** contains charts depicting the average effect of all features used in each model. Features at the top of the chart are more influential in affecting risk. The features at the bottom of the chart may contribute little to an athlete's risk on average but remain in the model because they improve the model's performance by contributing meaningfully in certain situations to improve overall model performance.

Feature and Load Statistic Definitions

The following table defines all the load statistics and model features contained in the application. The model column identifies those features used in the risk models (either Preseason or Regular Season). The Load Plot column denotes load statistics available for daily plotting on the Athlete Details page. All features and load statistics are available for display in the Athlete Roster table.

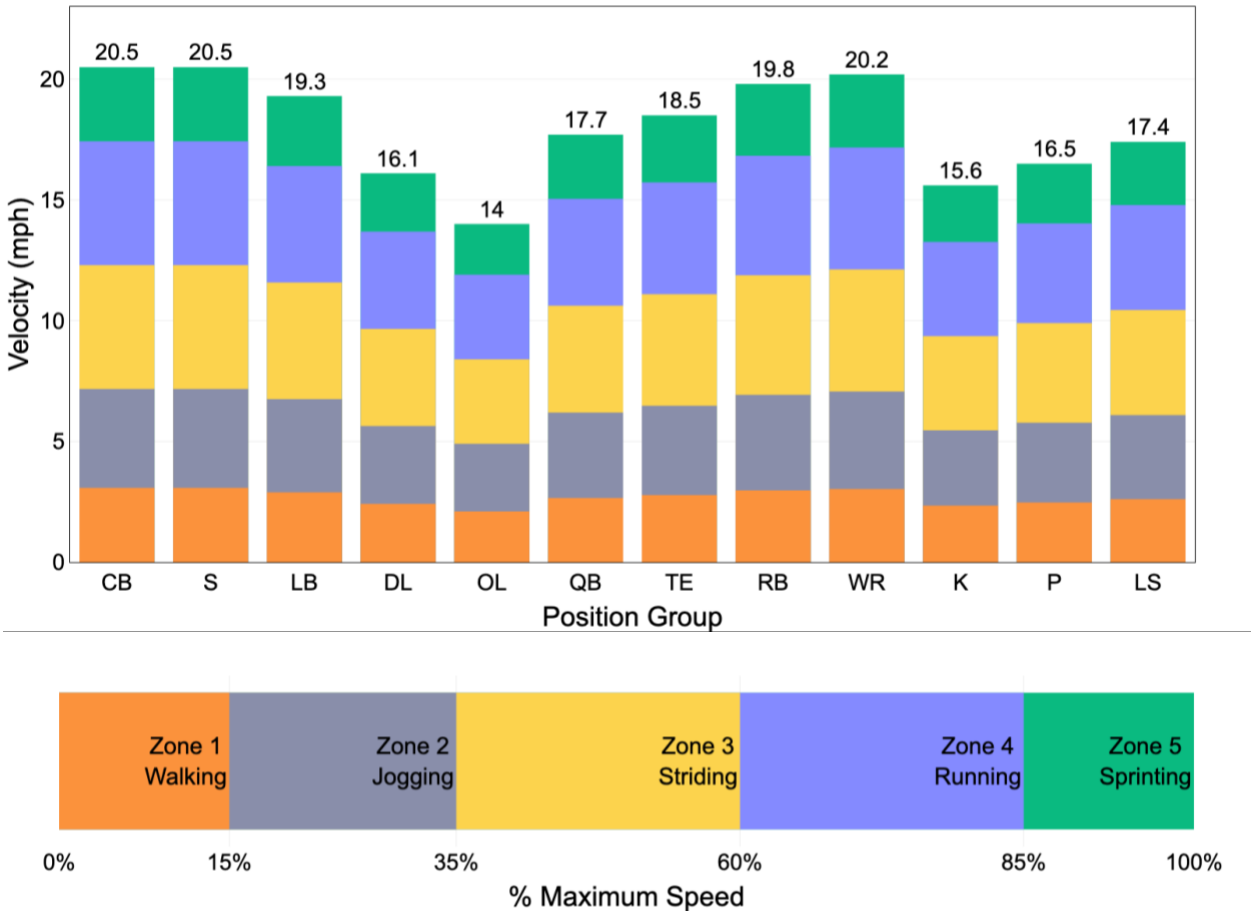
Table 2. Digital Athlete Load Statistics and Model Features

Table	Load Benchmark	Load Plot	Pre Model	Reg Model	Feature Name	Definition
					Position	Athlete's assigned position in NFL database
					Player Age	Athlete age in years
					Days Participated	Cumulative number of days of participation by athlete in current season
					Days Since Off	Number of days an athlete has participated in a row
					Days until Next Game	Number of days until the team's next game
					Duration	Sum of duration of tracking session
					21-Day Duration	Sum of duration of tracking sessions over 21-day window
					7:28 A:C Duration	Ratio of the average daily duration over the last 7 days over the average daily duration over the last 28 days
					Distance	Total distance traveled during a tracking session
					Acute:Chronic Distance	Ratio of today's load to the players average daily load
					4-Day Distance	Sum of distance traveled over the prior 4 days
					21-Day Distance	Sum of distance traveled over the prior 21 days
					7:28 A:C Distance	Ratio of the average daily distance over last 7 days to the average daily distance over last 28 days
					Density	Total distance divided by total duration
					Max Speed	Maximum velocity achieved during a session
					Sprint Distance	Distance traveled at speeds > 85% positional maximum velocity
					Sprint Efforts	Count of times a player reaches speeds above 85% positional maximum velocity (Zone 5) and sustains that speed for 0.5s
					4-Day Sprint Distance	Sum of sprint distance over last 4 days, defined as Zone 5 (>85% positional maximum speed).
					21-Day Sprint Distance	Sum of distance in Zone 5 (>85% positional maximum speed) over previous 21 days
					Running Distance	Distance traveled at speeds > 60% positional maximum speed
					21-Day Running Distance	Sum of distance in Zone 4 and Zone 5 (>60% positional maximum speed) over previous 21 days
					7:28 A:C Running Distance	Ratio of the average daily high speed distance over last 7 days to the average daily high speed distance over last 28 days, where high-speed is defined as zones 4 and 5 or above 60% positional maximum speed
					Acceleration Efforts	Count of times a player accelerates above 3.5 m/s/s and maintains that acceleration for 0.3s
					Acceleration Distance	Distance traveled while accelerating above 3.5 m/s/s
					Z4 Acceleration Distance	Distance traveled while accelerating between 3.5 and 5.5 m/s/s (acceleration zone 4)
					4-Day Average Accel Distance	Average distance traveled while accelerating between 3.5 and 5.5 m/s/s calculated over the prior 4 days on days the athlete participated in games/practices
					Deceleration Efforts	Count of times a player decelerates below -3.5 m/s/s and maintains that deceleration for 0.3s
					Deceleration Distance	Distance traveled while decelerating below -3.5 m/s/s
					Z5 Acceleration Distance	Distance traveled while accelerating above 5.5 m/s/s (acceleration zone 5)
					21-Day Z5 Deceleration Distance	Total distance traveled while decelerating below -5.5 m/s/s (deceleration zone 5) over the past 21 days.
					Lateral Accel Efforts	Count of times a player accelerates to change direction above 3.5 m/s/s and maintains that acceleration for 0.3s. This is the component of total acceleration used to change directions and not change speed
					Lateral Accel Distance	Distance traveled while a player accelerates to change direction above 3.5 m/s/s. This is the component of total acceleration used to change directions and not change speed
					Contact	Total contact during a tracking session, derived through a machine learning algorithm of the player tracking data
					4-Day Average Contact	Average contact load over the last 4 days on days the athlete participated in games/practices
					7:28 A:C Contact	Ratio of the average daily contact over previous 7 days to the average daily contact over previous 28 days
					Distance Under Contact	Distance traveled while under contact
					7:28 A:C Distance Under Contact	Ratio of the average daily distance under contact over previous 7 days to the average daily value over previous 28 days
					Time in Contact	Total time spent while under contact
					Contact (11on11)	The contact load from 11on11 or team periods (only functions if periods tagged)
					21 Day Total Contact (11 on 11)	Total contact that occurred during 11on11 or team periods over the prior 21-day window
					Distance under Contact (11on11)	Distance traveled during 11on11 periods while under contact
					Time in Contact (11on11)	Total time spent in 11on11 periods while under contact

Positional Velocity Bands

Sprinting and running are defined using relative positional maximum velocities as opposed to absolute thresholds. Zone 5, or sprinting, is defined as 85% of the positional maximum speed. High-speed running, zone 4, has a lower threshold of 60% of the positional maximum speed. Values for each of the zones are plotted below. Speeds above 100% of the positional maximum speed are included in zone 5. More information on the development of these bands is available from Sanchez et al. 2023: Positional Movement Demands during NFL Football Games.

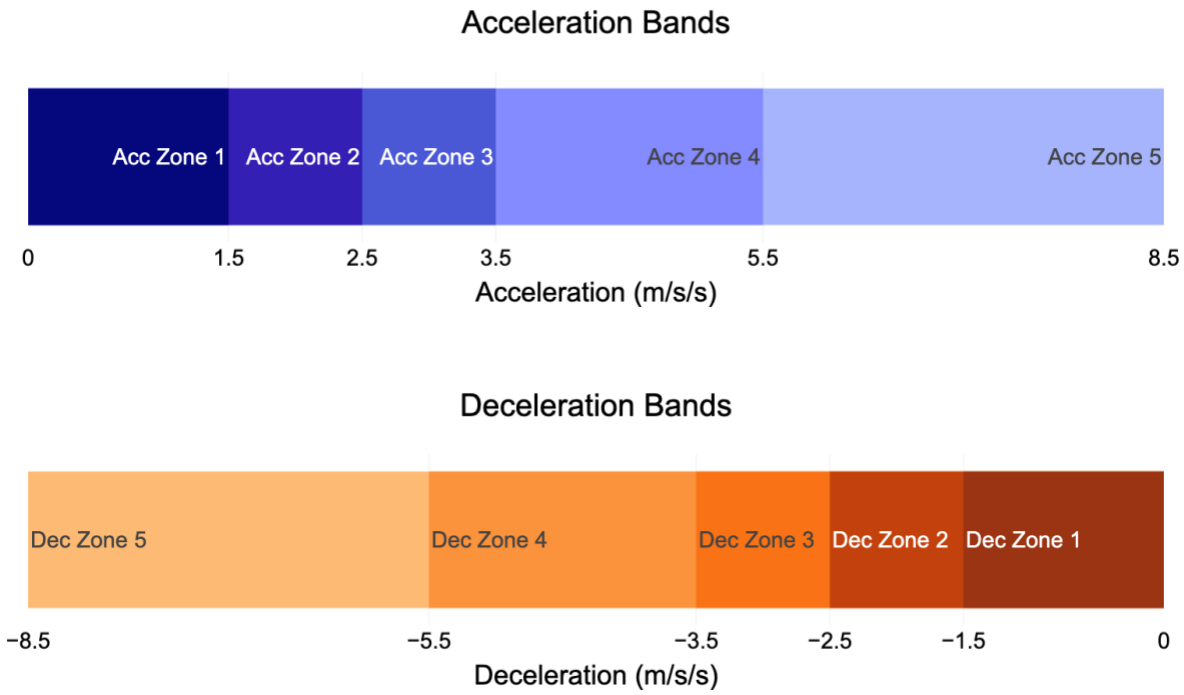
Figure 5. Positional Velocity Bands



Acceleration/Deceleration Bands

Accelerations and decelerations were also defined relative to bands, as shown below.

Figure 6. Acceleration and Deceleration Bands

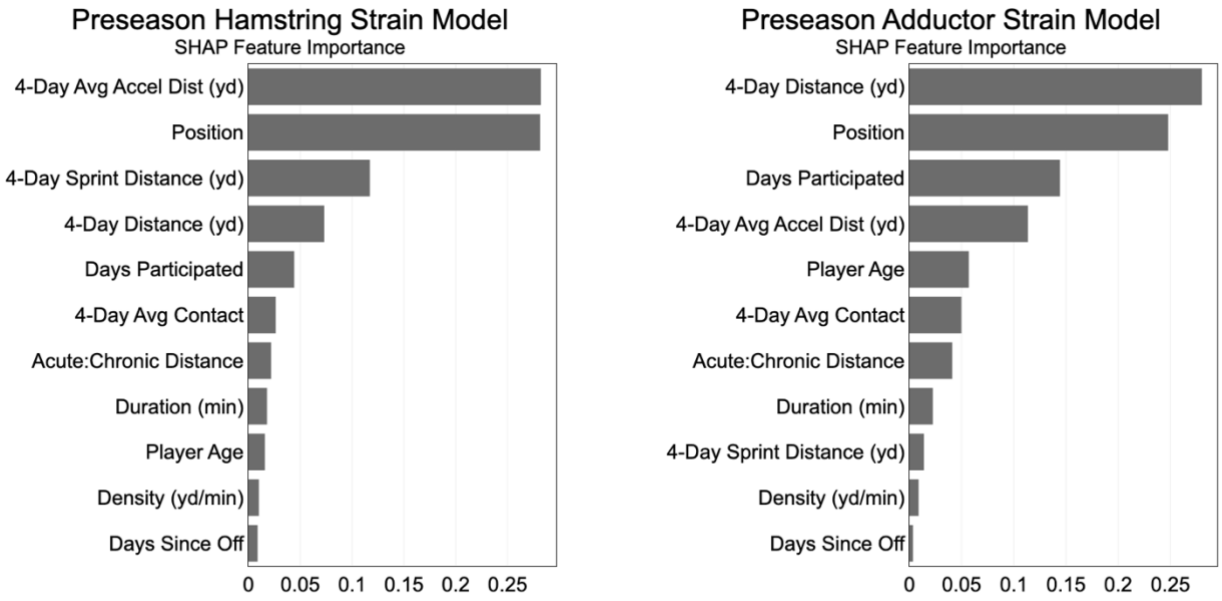


Appendix

Preseason Feature Importance and Feature Effects Plots

The below plots illustrate the relative importance of the features in the preseason hamstring strain risk mitigation model and the preseason adductor strain risk mitigation model when applied to the entire NFL population (Figure 7). Features at the top of the plot contribute more to the risk scores of athletes than those at the bottom.

Figure 7. Preseason Risk Mitigation Model Feature Importance



The following pages illustrate the average effect of each risk feature across the NFL population. Features at the top of the chart have more influence on risk. The features at the bottom of the chart may contribute little to an athlete's risk on average but remain in the model because they improve the model's performance by contributing meaningfully in certain situations that improve overall model performance.

Figure 8. Preseason Hamstring Strain Risk Feature Effects

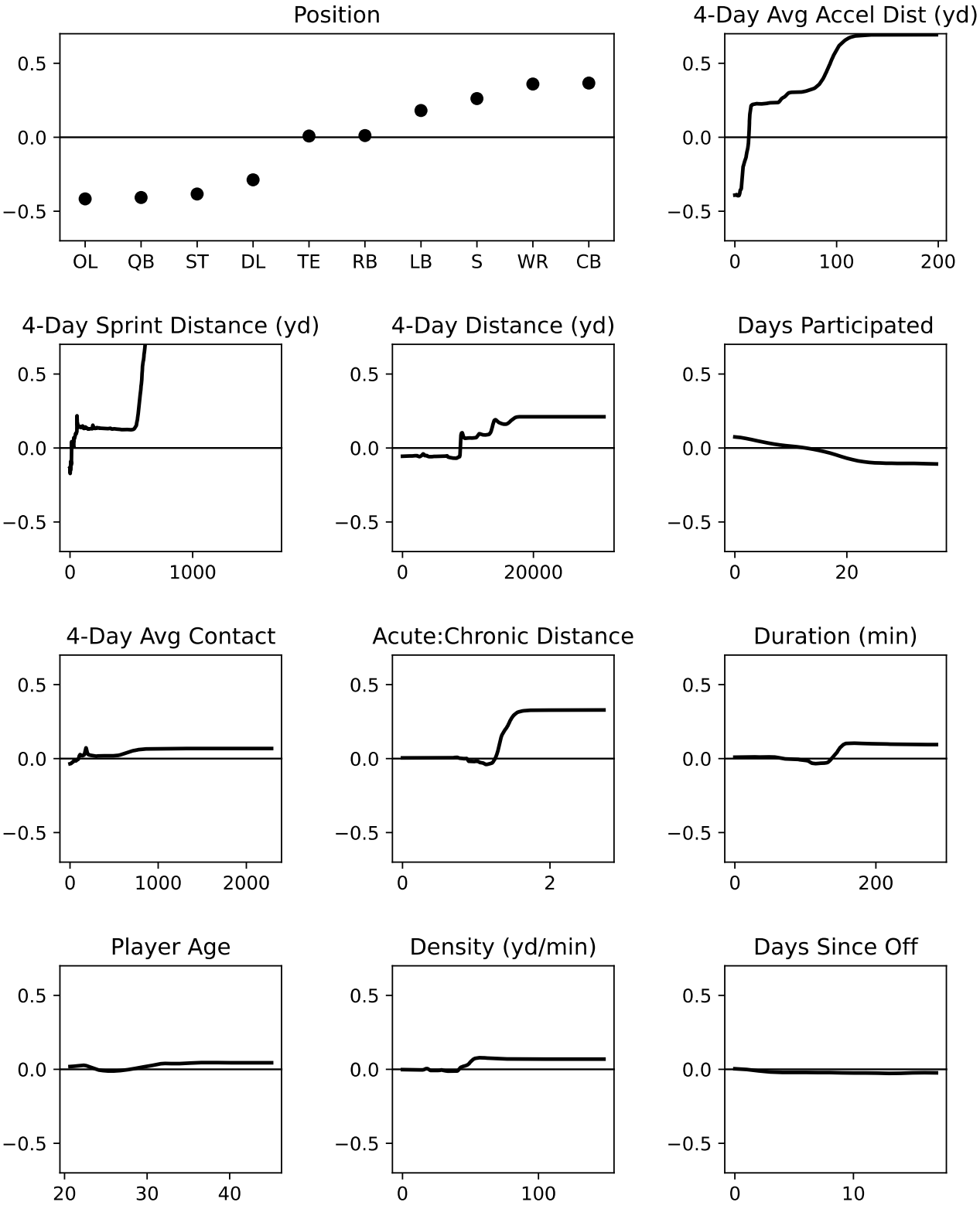
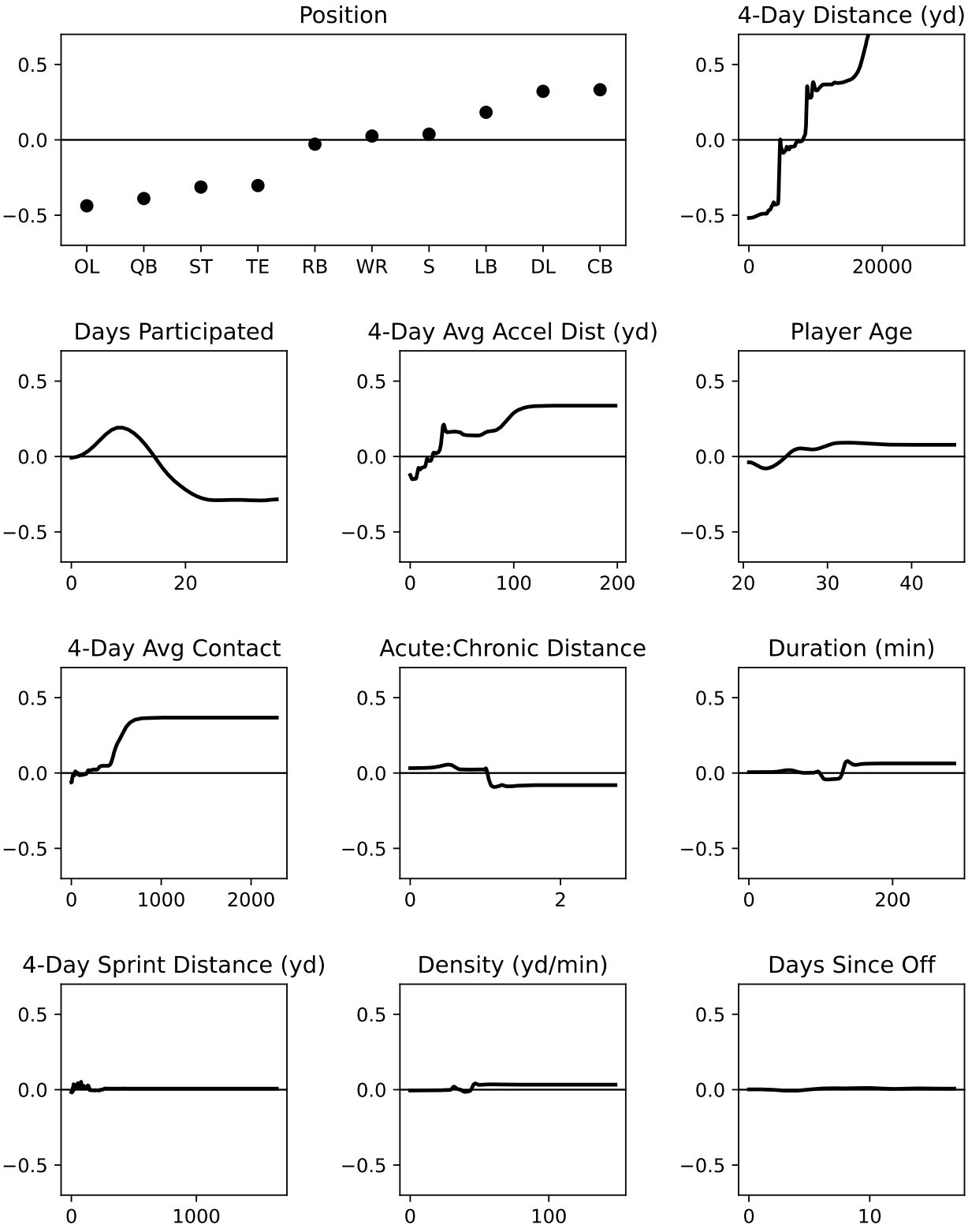


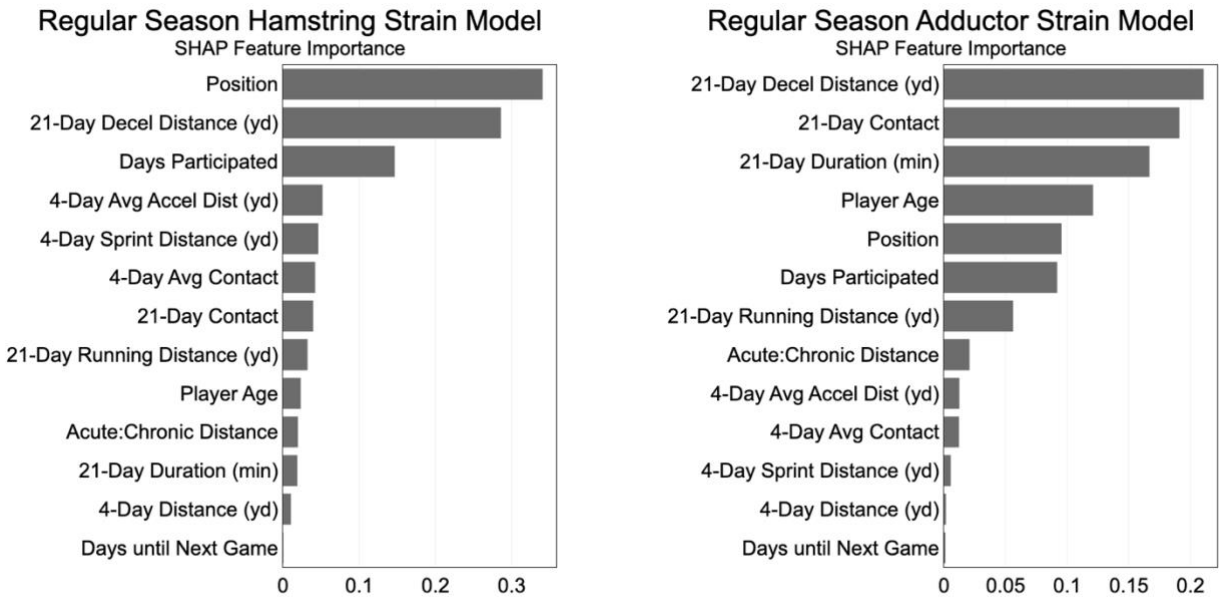
Figure 9. Preseason Adductor Strain Risk Feature Effects



Regular Season Feature Importance and Feature Effects Plots

Figure 10 again illustrates the relative importance of the features, this time for the regular season hamstring and adductor strain risk mitigation models. Features at the top of the plot contribute more to the risk scores of athletes than those at the bottom.

Figure 10. Regular Season Risk Mitigation Model Feature Importance



The plots on the following pages illustrate the average effects of each feature for each value across the NFL population, first for the hamstring strain model followed by the adductor strain model. Note that the position features effects plot appears first despite it being the fifth most important feature for the adductor strain model.

Figure 11. Regular Season Hamstring Strain Risk Feature Effects

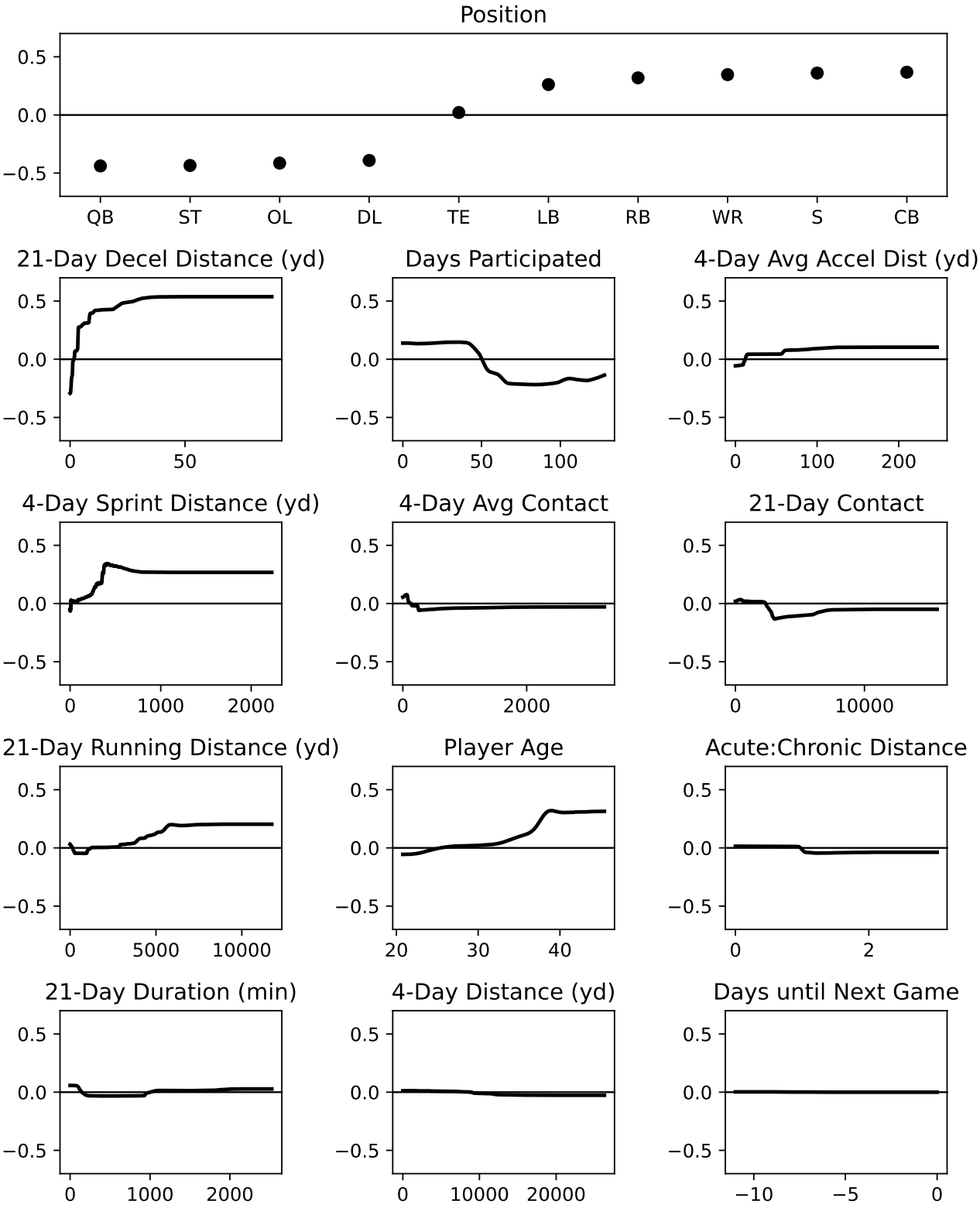


Figure 12. Regular Season Adductor Strain Risk Feature Effects