Table of Contents

Overview

The purpose of the Risk portion of the NFL Digital Athlete is to identify athletes that are at increased risk for hamstring injury due to workload and identify specific interventions that can be applied in the near term to reduce their risk. The risk of injury for any specific athlete on any specific day of training camp is low, but there are very large differences in the risk levels between athletes on a given day, even within the same position. To simplify the expression of risk for the NFL population, the risk probabilities are expressed in terms of **normalized risk factors**, which reflect the difference in risk between a specific athlete and the typical player both within the NFL population as a whole and the population of NFL athletes at their specific position. The definitions of the two normalized risk values are:

- **NFL Risk Factor:** An NFL athlete at the 50th percentile for daily risk of injury across the NFL preseason period 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 the player with an NFL Risk Factor of 1.0. An athlete with an NFL Risk Factor of 2.0 has double 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 5 times the risk of a typical player.

- **Positional Risk Factor:** An NFL CB at the 50th percentile for daily risk of injury across the NFL preseason period has a Positional Risk Factor of 1.0. A CB with an NFL Risk Factor of 1.5 has a 50% higher risk of injury than the player with an NFL Risk Factor of 1.0. An athlete with an NFL Risk Factor of 2.0 has double 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 5 times the risk of a typical player.

Due to the large difference in risk for injury between positions, positional risk factors can be very different than 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 is at about 70% higher risk for hamstring injuries than the average NFL player. Conversely, a ST player with a Positional Risk Factor of 1.0 has an NFL Risk Factor of only 0.5, reflecting that they have a greatly reduced risk of hamstring injury.

NFL Risk Factors provide the ability to quickly identify the athletes most at risk across your team (largely supported with the Athlete Roster Table workflow described below), regardless of position. Positional Risk Factors provide the ability to quickly identify the athletes on your roster at each position that are at increased risk relative to their peers at the same position (largely supported with the Positional Risk Workflow), facilitating discussions with positional coaches.
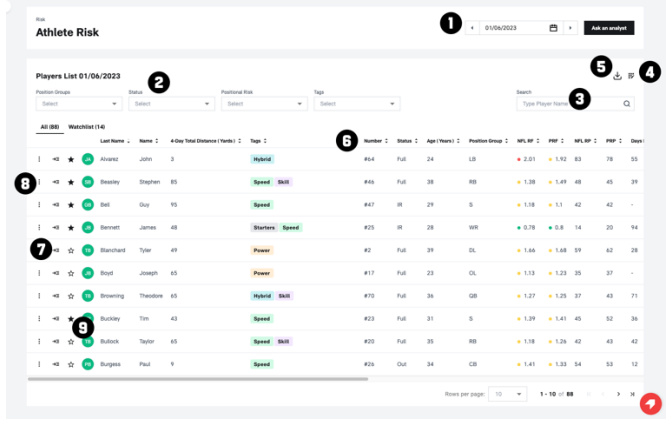
An athlete's risk factors for a given day are calculated to express the athlete's risk <u>tomorrow</u>, given the training load they have experienced during the season <u>up to and including today.</u>

Athlete Roster Table Workflow

The Athlete Roster Table workflow is designed to quickly review the status of all athletes on the roster for a particular day. Every athlete recorded on your active roster (the official NFL roster as populated in the NFL databases) will appear in the table. In the situation that the NFL's roster database was not updated in a timely fashion for the day, the previous day's roster will be used to populate this table until your team updates the status (thus there are cases where athletes that have participated may not appear). The platform checks for new performance data every 10 minutes. Once new data is found, it takes approximately 15 minutes to ingest and process the data into the application.

The diagram below describes a workflow for quickly identifying athletes most at risk and recommended interventions for reducing their risk.

1. Select a date to review records for athletes on your team. The default date is today.

2. If needed, use filters to narrow your results.

3. If needed, search for a specific athlete by name.

4. Click here to customize the columns shown in the table. Your customization will be stored in preferences and reloaded when you return.

5. Export the current table view to CSV.

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

7. Clicking on the "Details" link or the athlete's image will open the Athlete Details page for the selected athlete.

8. Click on the vertical dots to add or edit tags for an athlete.

9. Click the star to add or remove an athlete to your Watchlist.

Guide continues on next page…

Positional Risk Workflow

The Team Risk page provides the opportunity to review the risk history for all athletes at a given position. This facilitates quickly identifying athletes at each position that are at increased risk as well as identifying athletes whose risk is trending up.

1. Select a position to view. All athletes on the roster at that position are displayed.

2. You may filter within a position by tags assigned to players in that position.

3. Clicking on a given day in the chart updates the bar plot at the bottom (item 4) to reflect the relative risk between athletes on that specific date. The default date when the page loads is the last date in the global filters week selector.

4. Positional daily risk is shown for each player in the selected position.

5. Clicking on a player's risk bar will go to the Athlete Details page for the date reflected in the chart.

6. You may compare multiple positions to each other for risk.

It is important to note that a positional risk value of 1.0 denotes the average risk value over the entire Preseason. As the graphic below illustrates, early in training camp the risk is much higher at most positions (note the changes in risk after acclimation, during the first week of contact integration, and right around the first preseason game). For this reason, the positional risk plots provide the average positional risk value for the NFL position cohort for each day (this is calculated for each team based on days since vet report date to account for varying team schedules). Athletes above this line are at higher-than-expected risk for the time period. Those that are below are at lower risk than average for their position at that time period.
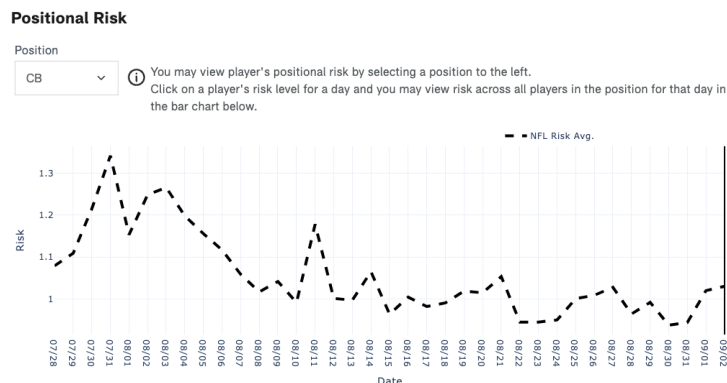
Athlete Details Page

The Athlete Details page provides a detailed overview of a given athlete's load and risk history. Specifically, it provides the following information:

- The daily risk profile for an athlete for the selected period.
- Detailed explanation for the load features that contribute to the athlete's risk value for the day, providing opportunities for intervention to reduce risk.
- Load history for the athlete for the selected period.
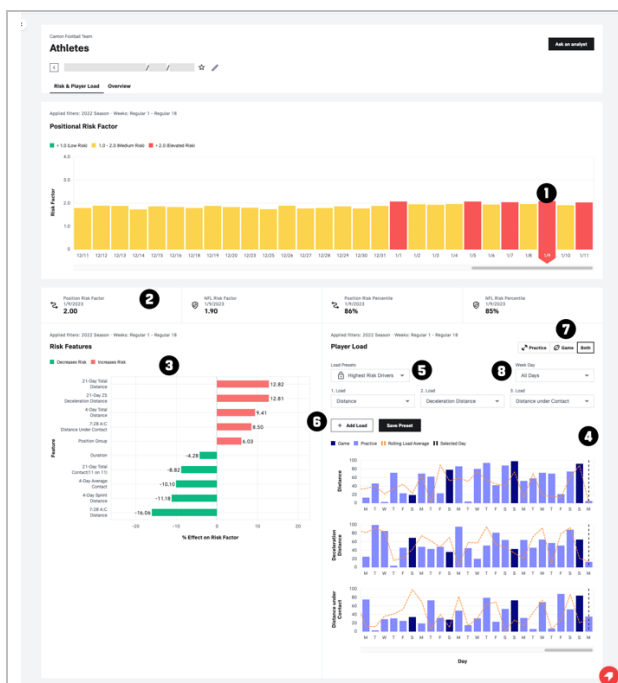
1. The athlete's history is displayed up to the date selected in the global filters. If a date is selected for which there is no time history, available dates will be plotted (and the risk level details plot will be blank since there is no data for that date). Clicking on a bar in the Positional Risk Factor history chart will update the data of the Risk Features Plot (item 3) and Player Load Plots (item 4)+++++.

2. The Risk Level Details chart provides the position risk factor, the NFL risk factor, position risk percentile, and the NFL risk percentile for the athlete-day.

3. The Risk Features chart provides opportunities for interventions by plotting the % importance that each model feature contributes to the difference between this athlete and the average NFL athlete. See the user guide section **Explanation of Risk Mitigation Plots** for more details.

4. Player load statistics for the selected date range are also included. Up to five different player load features can be plotted simultaneously. By default, the system will display the 3 load items that are contributing most to the player's risk.

5. You may choose from several preset load values or choose from personal presets here.

6. You may alter or add load types to the Player Load section and then save those items as presets.

7. Use this selector to view game, practice, or both loads displayed in the charts below.

8. The Week Day selector allows you to view only load data for a particular day of the week. This is most useful during regular season when practice schedules are more normalized.

Guide continues on next page…

Risk Mitigation Model Overview

The section describes the structure of the machine learning model, the feature selection process used to populate the model, a brief visual summary of the process by which probabilities are assigned, and a guide on interpreting the risk mitigation (daily risk details) plots.

Daily risk estimates are calculated each day using a machine learning model that predicts an athlete's probability of injury <u>tomorrow</u>, given the training load they have experienced during the season up to and including today excluding walkthrough and rehab sessions and starting from the 1st day of preseason early report. 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 that summarizes how the model assigns probabilities for athlete risk each day. These probabilities are converted into normalized risk factors for ease of interpretation as discussed in the **Application Overview** section above.

Model Feature Selection

Model Feature Selection Summary (Selected Feature Types Indicated in Gray)

Type	Base Feature	Temporal Window				
		Fixed	**Cumulative**	**Daily**	**4 Day Window**	**Acute:Chronic**
Non-Modifiable	Position	■				
Non-Modifiable	Age	■				
Schedule	Training Days		■			
Schedule	Sequential Days		■			
Duration	Duration			■		
Intensity	Density (y/m)			■		
Speed	Speed Zones				■	
Contact	Contact				■	
Acceleration	Acceleration Zones				■	
Distance	Distance				■	■
Deceleration	Deceleration 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 a set of load features that teams routinely track on a daily basis. 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 of a type listed in the Base Feature column of the table above. Note that within a base feature description, there were potentially multiple metrics considered. For example, for Speed Zones, teams may track Zone 4 (High Speed Running), Zone 5 (Sprints), or the combination of both zones. They may also track either sprint efforts or sprint distance. All these various 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 as depicted in the table above. The Daily features represent the acute load for the day's training period. The 4-Day Window features included the sum and average (when training) for each metric over the immediately preceding 4-day training block. The Acute:Chronic features represent the ratio of the athlete's acute load to the athlete's average daily load when training (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. Definitions for each of the selected features is available in the **Feature and Load Statistic Definitions** section below.
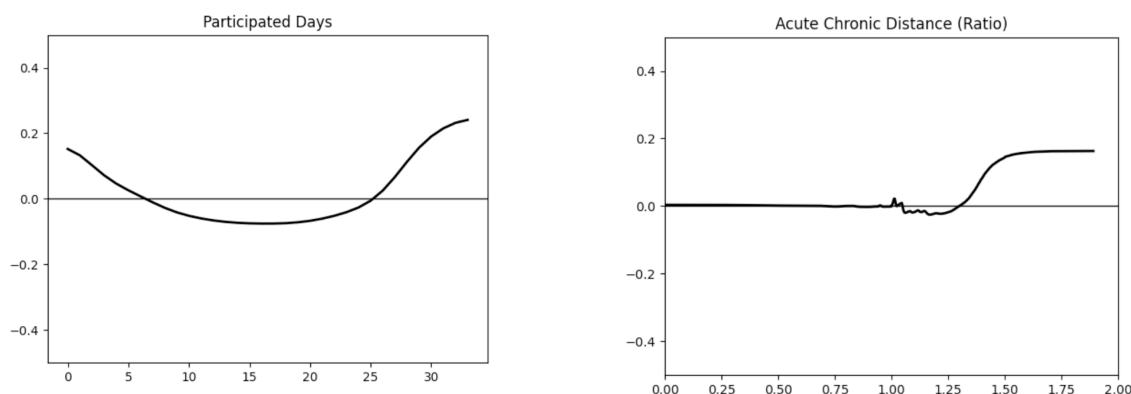
Feature Effects Plots

This section provides a visual explanation for how the machine learning model assigns risk. The selected machine learning model is designed to perform well for decisions involving features that have complex interactions. It is important to note that model does not consider the features for a given athlete by independently adding risk feature by feature as in a regression model, but rather assigns risk by simultaneously considering all of the features for a given athlete using a large ensemble of decision trees. However, it is possible to use visualizations of model assignments to understand the decision-making process.

Feature Effects Plot for Age for NFL Cohort
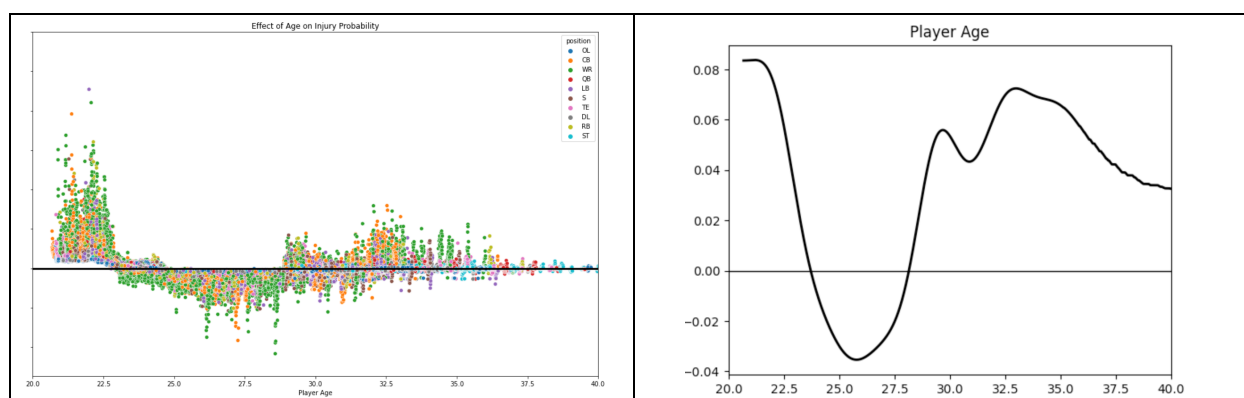
Consider the example above, in which the panel on the left plots the effect assigned to every athlete in the NFL based on their age over multiple seasons (each point on the graph represents an athlete-day). Note that there is variation in the assigned risk due to interaction with other features. The panel on the right depicts the smoothed average value at each observed age (the y axis depicts the change in log-odds). When this line is above 0, it indicates that the model (on average) assigns higher risk to the athlete, and when below 0 the model (on average) assigns reduced risk at that age. The ability of the model to handle non-linear effects is clearly illustrated above. The interpretation of the effect from these plots is that athletes under the age of 24 (i.e., rookies etc.) are at the highest risk. Athletes with experience between the ages of 24 and 28 are at reduced risk. Athletes above the age of 28 are also at increased risk.

Average Feature Effects for the NFL Cohort for Selected Features

Two more example plots of feature effects are provided above, plotted on the same scale. At high values, these features contribute more to a player's risk value than player age (note the y-axis scale difference). It can be seen that the maximum effect of these features is similar, and the overall non-linear pattern is evident. The interpretation of these plots is that athletes are at increased risk during the 1st 5 days of their participation (for most athletes this encompasses the acclimation period). Once an athlete has accumulated more than 25 training/game days of load in the preseason, their risk again begins to climb (possibly indicating a fatigue effect). The pattern for Acute:Chronic Distance is somewhat different. Risk remains unchanged for Acute:Chronic Distance values up to 1 (the slightly non-linearity around the value of 1 occurs because this is a very common value for athletes so other features are driving the risk scores for athletes around this value). Athlete-days with A:C Distance ratios between 1 and 1.25 are at slightly reduced and then risk begins to climb as an athlete's A:C Distance ratio increases.

The chart below depicts the average effect of all of the model features over the domain of values in the NFL cohort plotted on the same scale. Features at the top of the chart are more influential in affecting risk with position and acceleration features providing the most influence. 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. Although it is difficult to see when plotted at the same scale as other features, each day of the Days Since Off feature contributes a small amount of increased risk.

*Average Feature Effects for the NFL Cohort*

9
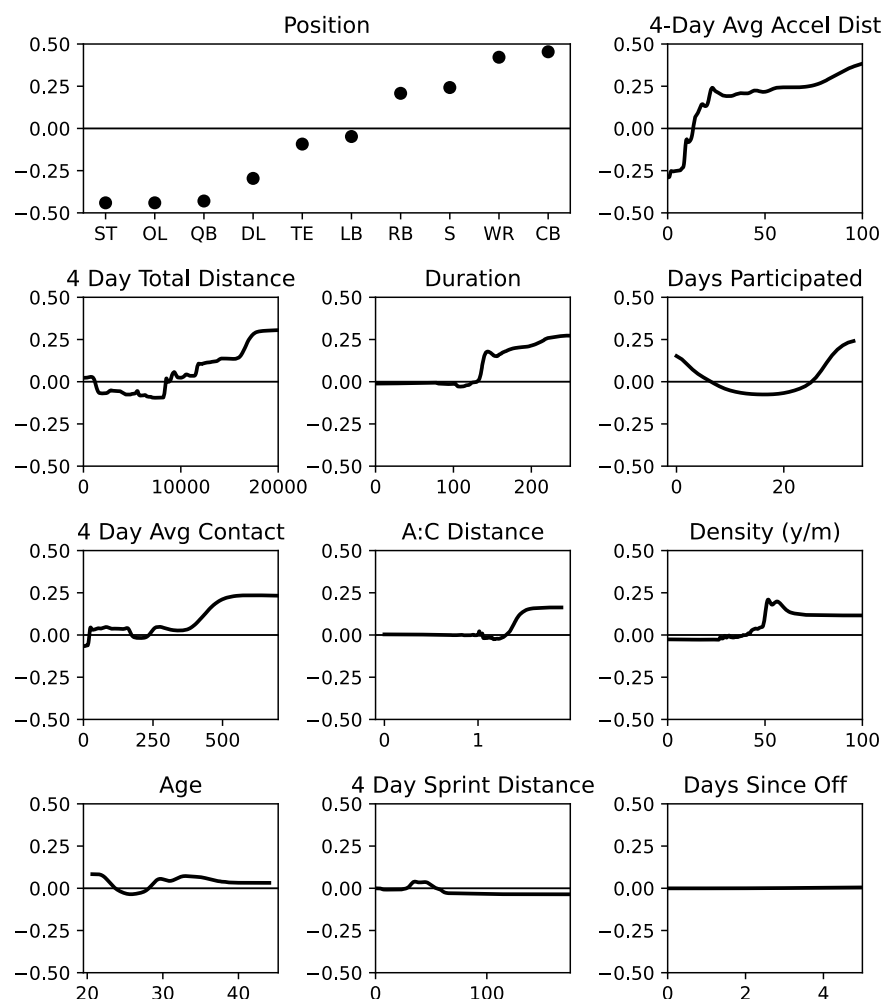
Explanation of Risk Mitigation Plots

The values in the risk mitigation plots (i.e., Risk Level Details) represent the risk effects discussed in the last section applied to a specific athlete-day. The risk effects are plotted as a % of the overall difference between this athlete-day and the 50th percentile NFL athlete (an athlete with an NFL Risk Factor of 1.0). Bars in red indicate contribution to increased risk. Bars in green indicate contributions towards reduced risk. For an athlete with an NFL Risk Factor of exactly 1.0 (note that Risk Factors are rounded to 1 decimal place), the sum of the red and green bars will add up to 0, representing an athlete that is exactly average for the NFL cohort. It is important to note 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 has a high-risk factor due to large values for density (yards per minute), 4-day total distance, and 4-day average acceleration. 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.

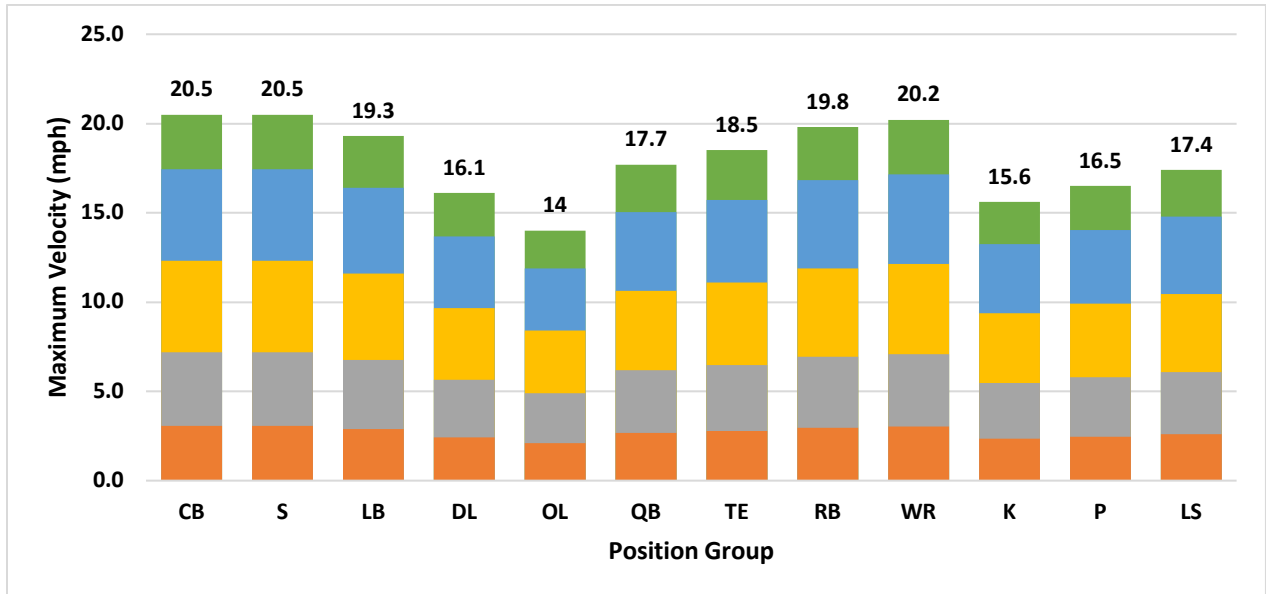*Example Risk Mitigation Plot for an NFL Athlete at Increased Risk*
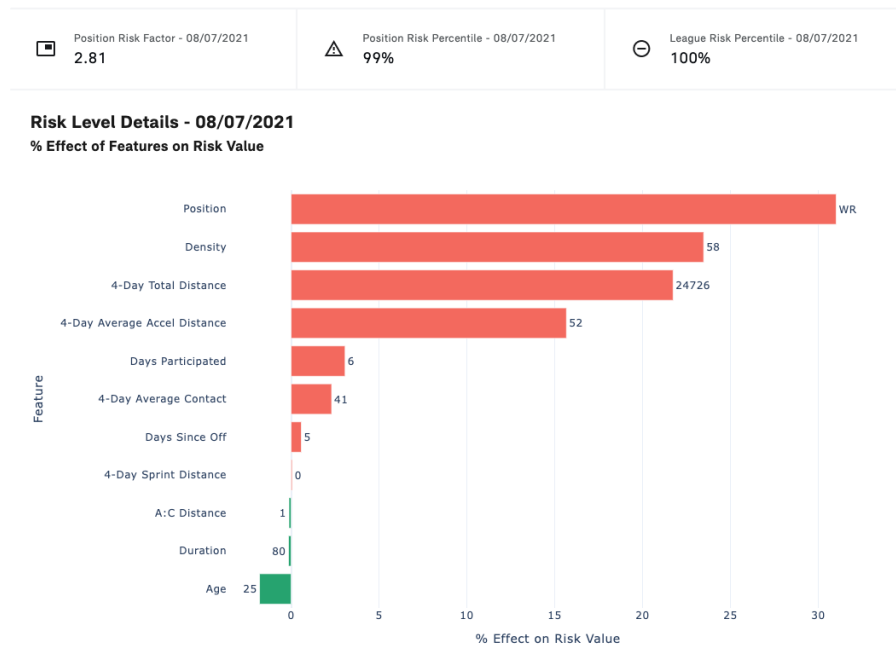
Feature and Load Statistic Definitions

The following table documents definitions for all of the load statistics and model features contained in the application. The model column identifies those features used in the risk model. The load plot column denotes load statistics available for daily plotting in the Athlete Details page. All features and load statistics are available for display in the Athlete Roster table.

Table	Load Plot	Model	Feature Name	Definition
✓		✓	Position	Athlete's assigned position in NFL database
✓		✓	Age	Athlete age in years
✓		✓	Days Participated	The cumulative number of days of participation by athlete in current season.
✓		✓	Days Since Off	The cumulative number of days of participation in a row.
✓		✓	4-day Total Distance	Total distance travelled over the prior 4 days.
✓		✓	4 Day Sprint Distance	Sum of sprint distance over last 4 days.
✓		✓	Duration	Total duration of tracking session
✓		✓	Density	Total distance divided by total duration
✓		✓	4-Day Average Contact	The average contact load over the last 4 days.
✓		✓	4-Day Average Accel Distance	Average distance travelled while accelerating between 3.5 and 5.5 m/s/s calculated over the prior 4 days
✓		✓	A:C Total Distance	The ratio of today's load to the players average daily load
✓	✓		Distance	Total distance travelled during a tracking session
✓	✓		Max Speed	Maximum velocity achieved during a tracking session
✓	✓		High-Speed Distance	Distance travelled at speeds > 60% positional maximum velocity
✓	✓		Sprint Distance	Distance travelled at speeds > 85% positional maximum velocity
✓	✓		Sprint Efforts	Count of times a player reaches speeds above 85% positional maximum velocity and sustains that speed for 0.5s
✓	✓		Contact	Total contact during a tracking session
✓	✓		Distance under Contact	Distance travelled while under contact
✓	✓		Time in Contact	Total time spent while under contact
✓	✓		Contact (11on11)	The contact load from 11on 11 periods (only functions if periods tagged)
✓	✓		Distance under Contact (11on11)	Distance travelled during 11on11 periods while under contact
✓	✓		Time in Contact (11on11)	Total time spent in 11on11 periods while under contact
✓	✓		Z4 Acceleration Distance	Distance travelled while accelerating between 3.5 and 5.5 m/s/s
✓	✓		Z5 Acceleration Distance	Distance travelled while accelerating above 5.5 m/s/s
✓	✓		Acceleration Distance	Distance travelled while accelerating above 3.5 m/s/s
✓	✓		Deceleration Distance	Distance travelled while decelerating below -3.5 m/s/s
✓	✓		Acceleration Efforts	Count of times a player accelerates above 3.5 m/s/s and maintains that acceleration for 0.3s
✓	✓		Deceleration Efforts	Count of times a player decelerates below -3.5 m/s/s and maintains that deceleration for 0.3s
✓	✓		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 travelled 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.

Positional Velocity Bands

Guide continues on next page…

Position

21 Day Total Distance

21 Day Sum of Max Speeds

21 Day Total Duration

21 Day Contact

Acute to Chronic 7:28 Day Distance

Acute to Chronic 7:28 Yards in Contact

Acute to Chronic 7:28 Total Duration

Acute to Chronic 7:28 High Speed Distance

Acute to Chronic 7:28 Contact