Predicting NHL Concussions

Peter Tanner
Email: prtanner@gmail.com

Abstract

With an epidemic of concussions in recent years, including 1,700 man-games lost during the 2011-2012 season, concussions amongst NHL players have become a serious issue. As part of efforts to reduce concussions in the NHL, this study explores which factors are important in predicting which players are most likely to sustain a concussion. By combining injury data with over 120 attributes of each NHL player, logistic regression analysis was carried out as part of the study. Results suggest that shorter players, younger players, defensemen, and players who get more penalty minutes have an increased likelihood of sustaining concussions. Amongst players with a history of concussions, the severity of previous concussions is a significant factor in sustaining another concussion while the recency of their last concussion is not. This study also examines players' performance after returning from concussion and finds that there are statistically significant decreases in many key metrics compared to before their concussion. Players who adapt their style of play to be less aggressive after returning from a concussion are less likely to sustain another concussion than those who do not adapt.

1. Introduction

During the second period of the 2011 Winter Classic, a 23 year old Sidney Crosby collided with David Steckel of the Washington Capitals. Crosby momentarily left the game and was back on the ice by the third period. Four days later, Victor Hedman hit the 5'11" Crosby in a game in Tampa Bay. Crosby would then sit out eleven months with post-concussion syndrome. Two weeks after returning from the injury Crosby took a blow from David Krejci of the Boston Bruins and would be sidelined again for three months.

Were Crosby's concussions random? Or were there attributes about the Pittsburgh superstar that made him more susceptible? How much did Crosby's first concussion put him at risk for another one? And when players like Sidney Crosby come back from concussions are they still the same type of player?

In this paper, we attempt to achieve three objectives:

1.) Understand the factors that influence the likelihood that a player will sustain a concussion, including previous concussion history
2.) Identify players who are at the highest risk of sustaining concussions
3.) Understand any impacts of concussions on players' performance once they return from injury and how any changes in their style impacts the probability of a future concussion.

2. Building the Logistic Regression Model

Game level player data was utilized for the 2009-2010 NHL season through the lockout shortened 2013 season. Only regular season games were analyzed; pre season and post season games were excluded. The game level player data was obtained from the Hockey Summary Project. Players who did not have any ice time in a given game were excluded, including back-up goalies. Data for the 2009-2010 season was used only to create variables describing players' attributes, for example 'Average number of penalty minutes per game last season.' Overall, there are 121,362 man games observed in the dataset for the 2010-2011 through 2013 seasons. These observations formed the main dataset for the analysis.

The concussion data was amalgamated from several sources, including The Concussion Blog, NHL injured reserve reports, media reports, and scratched player data from all NHL regular season game box scores.

Players reported as having a ‘concussion’, ‘post-concussion syndrome’, or ‘head injury’ who missed one or more games due to the injury were labeled as having a concussion. Players who missed games with the official reason being a generic ‘Upper Body Injury’ were not labeled as having a concussion unless that player’s injury was reported as being a concussion by another source. The occurrence of the concussion taking place was attributed to have happened in the player’s last game played before missing a game due to the injury. There were 276 reported concussions in total from the 2010-2011 to 2013 seasons used as ‘positive’ outcomes in the model. This represents a rate of one concussion for every 440 man games or one for every 11 NHL contests.

As sustaining a concussion is driven by ice time (you cannot get a concussion while sitting on the bench), the model predicts the probability of a concussion per unit of ice time. Also, since having a history of concussions is a dominant factor in having future concussions, we built separate models for players with a history of concussions and for those without a recorded concussion. Players with a previous concussion had a 78% higher probability of sustaining another concussion than players with no history of concussions. Variables that were tested in the model include players’ physical attributes, performance in the previous season and so far in the given season, the player’s concussion history, and attributes about the opposing team.

Model For Players With No History of Concussions

Table 1 below shows all variables in the model. Only significant variables were left in the final model of the 129 variables explored. There were four key factors in the model:

1.) Height in inches: Being taller reduces the probability of a concussion. This is hypothesized to be driven by taller players’ heads literally being ‘above the fray’ and instead take the impact of a collision lower down on their body.

2.) Age: Younger players are more likely to suffer a concussion, particularly those 20 and under. This is inline with a study [1] that found that younger athletes are more impacted by concussions.

3.) Position: Goalies are more than five times less likely to suffer a concussion than other players. Defensemen are at a higher risk than forwards. There is not a statistically significant difference between the three forward positions.

4.) Player’s penalty minutes in the previous season: Players who take a lot of penalties are more likely to suffer concussions. It should be noted that the type of penalties the player receives do not provide significant incremental explanatory value once overall PIM is in the model. This includes PIM from fighting, checking penalties (boarding, charging, etc), or stick infractions. All of these variables are strongly correlated, but overall penalty minutes was the most predictive. This lack of statistical significance for PIM from fights may be due to very thin data on fighters. Over 70% of concussions come from players that average less than two fights per season.

There were also several other notable variables that weren’t statistically significant:

1.) Star factor: Players who put up higher goals, assists, or points per game were not more likely to sustain a concussion than marginal players. This suggests that there is not significant ‘head hunting’ of star players.

2.) Goon factor: The opposing team’s average penalty minutes per game, acting as a proxy for a ‘goon factor’, was not significant

3.) Weight: After factoring in height, weight did not add incremental value in the model. Weight to height ratio was also not significant. We did not find that a short skinny player is more likely to sustain a concussion than a short stocky player.

Table 1: Logistic Model For Players With No History Of Concussions

Variable	Coefficient	Standard Error	Wald Statistic	DoF	Significance
Constant (Intercept)	$\beta_0 = -5.7247$	2.9208	3.8414	1	p = 0.0500
Player's Height (Inches)	$\beta_{Height} = -0.0701$	0.0346	4.1073	1	p = 0.0427
Player's Average Penalty Minutes Per Game Last Season	$\beta_{PIM} = 1.0489$	0.2612	16.1256	1	p <.0001
Player's Age	$\beta_{Age} = -0.9400$	0.4523	4.3189	1	p = 0.0377
Plays Center	$\beta_{Center} = 0.1473$	0.1428	1.0639	1	p = 0.3023
Plays Left Wing	$\beta_{LeftWing} = 0.1713$	0.1648	1.0815	1	p = 0.2984
Plays Defense	$\beta_{Defense} = 0.2976$	0.1256	5.6167	1	p = 0.0178
Plays Goalie	$\beta_{Goalie} = -1.1077$	0.2970	13.9098	1	p = 0.0002

Model For Players With A History of Concussions

Looking specifically at players who were identified as having a previous concussion, the number of games missed due to previous concussions is significant while the time since the player's last concussion is not. Overall, the severity of the previous concussions is a much bigger driver of reoccurrence than the recency of those concussions. This leads us to hypothesize that there may be permanent effects of concussions that do not dampen with time and leave the player prone to additional concussions indefinitely.

Table 2 below shows all variables in the model. The model for players with a history of concussions has non-significant variables for explanatory purposes. Similar to the model for players without a concussion history, the player's PIMs in the previous season is significant. The number of prior concussions is not significant once the number of games missed due to the prior concussions is factored in.

Table 2: Logistic Model For Players With A History Of Concussions

Variable	Coefficient	Standard Error	Wald Statistic	DoF	Significance
Constant (Intercept)	$\beta_0 = -12.6647$	0.7765	266.0170	1	p <.0001
More Than 1 Year Since Last Concussion	$\beta_{Time} = -0.3369$	0.2964	1.2921	1	p = 0.2557
More Than 20 Games Missed Due To Concussions	$\beta_{Games} = 1.0228$	0.3496	8.5595	1	p = 0.0034
Number Of Prior Concussions	$\beta_{NumPrior} = -0.7797$	0.6114	1.6262	1	p = 0.2022
Player's Average Penalty Minutes Per Game Last Season	$\beta_{PIM} = 1.3896$	0.5326	6.8075	1	p = 0.0091

Model Predictions On The Holdout Sample

The 2013-2014 NHL season up until the Olympic break was used as a holdout sample to validate the models. Overall the combined model has a Somers' D of 0.1327 in the possible range of -1.0 to 1.0 and a Gini of 0.1247. Figures A.3 and A.4 in the appendix display graphs of the model on the holdout sample. So far in the 2013-2014 NHL season players are getting concussions at a rate of 0.454% per 60 minutes of ice time

compared to an average of 0.733% in the previous three seasons. This has resulted in the model over estimating the total number of concussions in the 2013-2014 season, but still rank ordering the relative risk of concussions amongst players well. The lower rate of concussions may suggest that rule changes or other factors are having an effect on concussion risk, though it is still too early to draw conclusions. Figure A.5 shows the concussion rate by season.

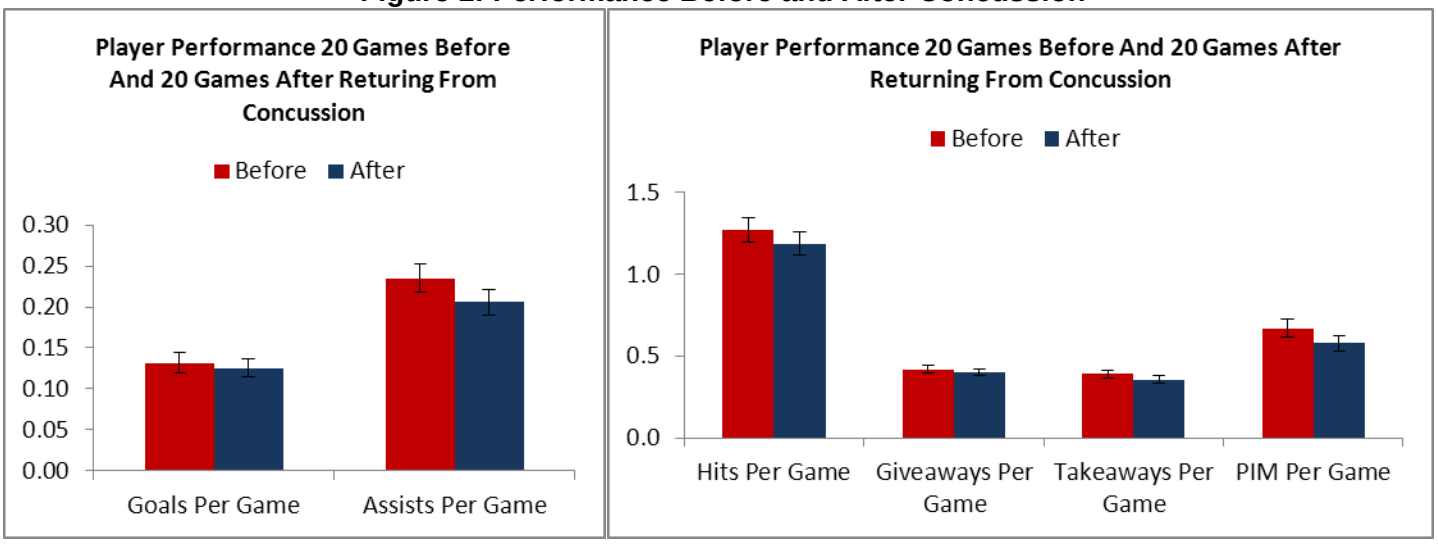
3. Player Performance Post Concussion

One of the key concerns amongst coaches is the impact of concussions on performance after a player returns from a concussion. This study examined key individual player statistics for the 20 games before the concussion and the 20 games after returning from a concussion. A performance drop was observed in most major statistics, including statistically significant decreases in penalty minutes, takeaways, hits, and assists. This is even after a 3% boost to post concussion metrics was factored in to account for the 3% lower average ice time players receive post-concussion than before. The decrease in penalty minutes and hits may be of most interest, suggesting that players adjust their style of play post-concussion to be less physically aggressive.

Table 3: Performance Before and After Concussion

Stat	20 Games Before Concussion	20 Games After Concussion	Change	Significance P(T<=t)
Goals Per Game	0.1316	0.1255	-4.59%	0.3091
Assists Per Game	0.2353	0.2060	-12.45%	0.0144
Hits Per Game	1.2712	1.1890	-6.46%	0.0223
Giveaways Per Game	0.4188	0.4016	-4.11%	0.2334
Takeaways Per Game	0.3903	0.3565	-8.65%	0.0364
Penalty Minutes Per Game	0.6690	0.5791	-13.45%	0.0245

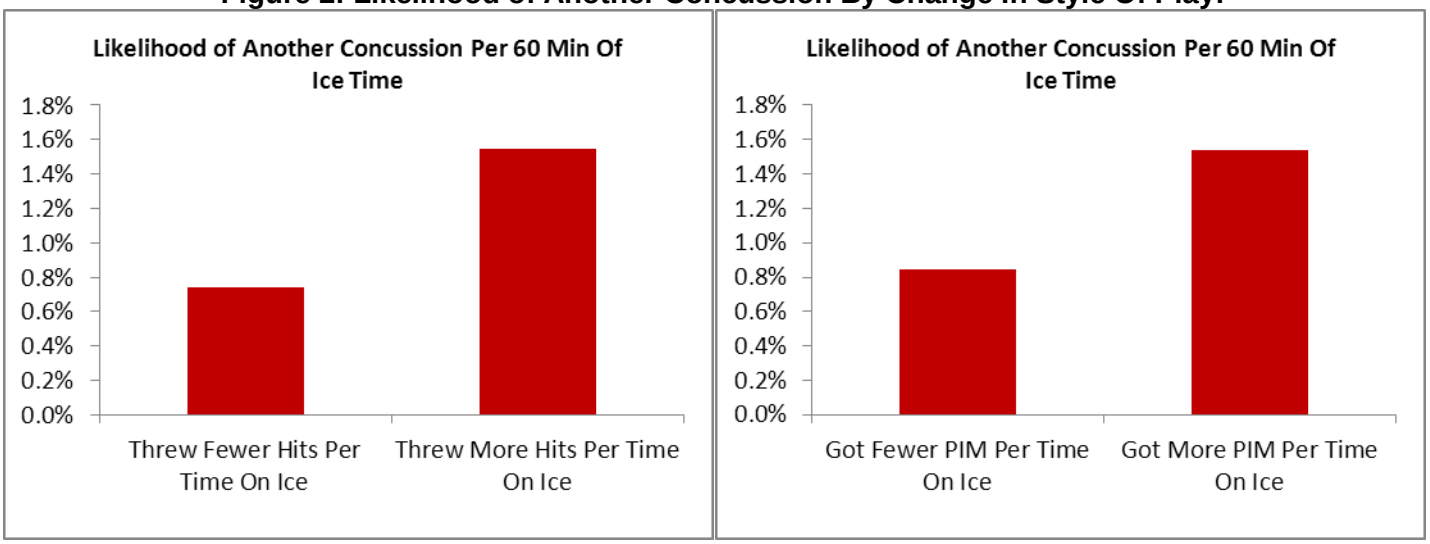
Figure 1: Performance Before and After Concussion



Notably, players who toned down their style of play in the 20 games after returning from a concussion are less likely to sustain another concussion in the longer term (through the lockout shortened 2013 season). Players who threw less hits than before were 52% less likely (p=0.017) to sustain another concussion than those who threw more hits than before. Players who decreased their penalty minutes were 45% less likely (p=0.0537) to go down with a concussion again than players whose penalty minutes increased. These performance variables

were not included in the logistic model as they may be 'forward looking' and another concussion can occur before a player has played 20 additional games.

Figure 2: Likelihood of Another Concussion By Change In Style Of Play:



Players returning from other types of injuries do not have the same decrease in output as those as concussed players. (See appendix figure A.2) In fact, players have a statistically insignificant minor boost in many key statistics. This may be due to players 'playing hurt' before these other injury types or the rest acquired while missing games.

4. Discussion

There is little individual players can act upon for some of the risk factors this study found for concussions. There is nothing you can do to change your age, height, or concussions you've already sustained. Players who receive a high amount of penalty minutes are at a higher risk of sustaining a concussion. This variable may be acting as a proxy for several things, such as a general aggressive style of play, head injuries sustained in fights, and being at higher risk for retribution from opponents.

The finding that players do indeed change their play after returning from a concussion is likely a smart decision for self-preservation, even if it hurts them in some performance metrics. The data shows that players who throw fewer hits and get fewer penalty minutes after returning from their concussion are less likely to sustain another concussion. Crosby threw 19% fewer hits in the 20 games after returning from his second concussion than the 20 games before his first concussion. More analysis should be done to quantify if these types of changes in play are permanent or if players regress back to their original style. In Crosby's case, his hits in 2013 are 25% lower than before his first concussion, but the Penguins are more than happy to have him healthy and on the ice.

The study's finding that the probability of another concussion does not decrease with the time since the player's last concussion is worrisome and suggest permanent impacts from suffering a concussion. More analysis over a longer time period of players' careers should be done to understand this better.

Table 4 shows the players the model identified as most likely to get a concussion in the 2013-2014 season. The top of the list is dominated by players with previous concussions and contains many enforcers. Sidney Crosby, while not in the top 25, is still in the top 15 percentile of players with a 0.32% chance of a concussion in each game.

The main flaw in this analysis is the general poor data quality for concussions. Concussions have been reported by teams as everything from an 'upper body injury' to 'flu like symptoms'. Determining exactly when a concussion took place can also be challenging, for example with Sidney Crosby's first concussion. With this

study we only acquired data back to the 2009-2010 season, so we were not aware of previous concussions earlier in players' careers, perhaps even before they entered the NHL. We encourage all hockey leagues to keep better records of concussions and reduce the culture of trying to mask potential concussions.

5. Conclusions

Concussions are a serious and still deeply mysterious injury which this study hopefully sheds more light on. The logistic regression shows that height, age, position, and a player's penalty minutes are significant factors in predicting concussions. Factors such as the player's weight, the player's points (head hunting stars), and the opposing team's average PIM (goon factor), are not significant. Players who have had a previous concussion are 78% more likely to sustain another concussion. The increased risk of another concussion does not dampen significantly with time, suggesting that concussions may have permanent impacts on players. Players who adapt their style of play after their concussion to be less aggressive are less likely to sustain another concussion. These findings stress the need for the NHL to continue to make rule changes in order to decrease future concussions.

6. Acknowledgements

Thank you to Dustin Fink at The Concussion Blog (http://theconcussionblog.com/) for supplying a list of players and dates of concussions.

7. References

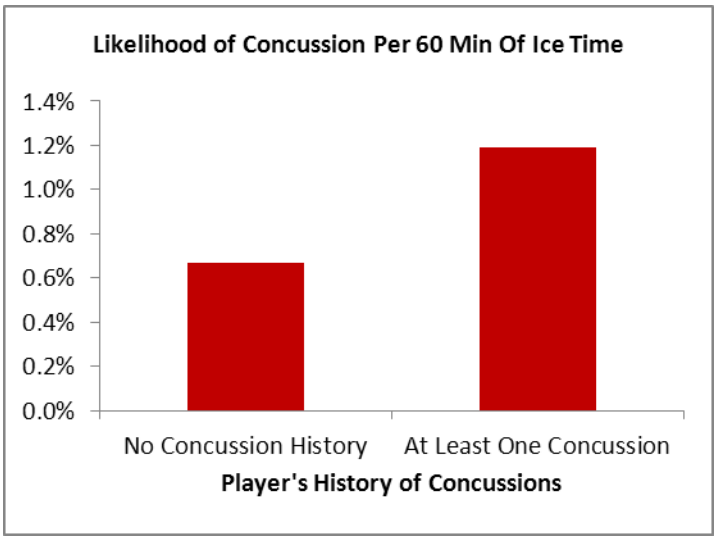
[1] Covassin T, Elbin RJ, Harris W, Parker T, Kontos A. The role of age and sex in symptoms, neurocognitive performance, and postural stability in athletes after concussion. Am J Sports Med. 2012 Jun;40(6):1303-12. doi:

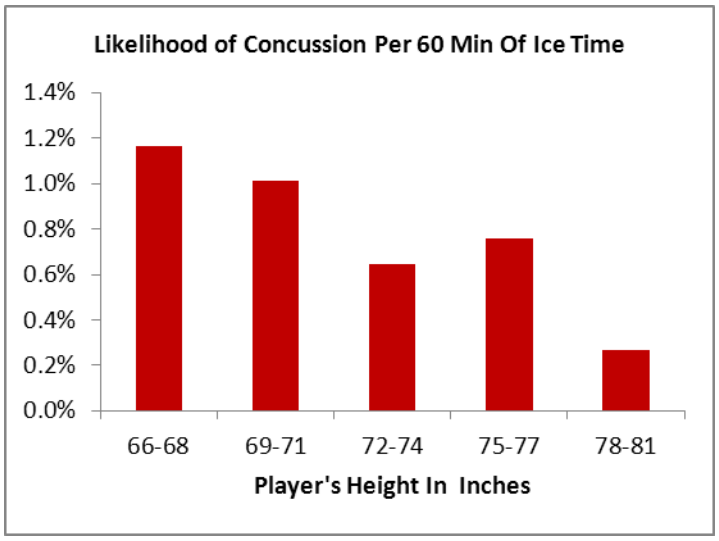
Table 4: 2013-2014 Season Concussion Predictions Of Players At Highest Risk

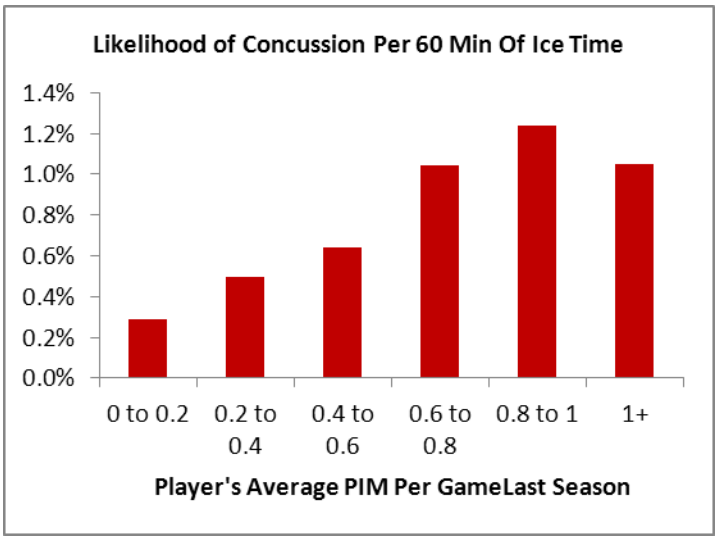
Rank	***Player***	***Probability of A Concussion In A Given Game in the 2013-2014 season***
1	*Robert Bortuzzo*	*1.34%*
2	*Wayne Simmonds*	*1.25%*
3	*Evander Kane*	*1.08%*
4	*Colton Orr*	*1.01%*
5	*Derek Dorsett*	*0.95%*
6	*David Perron*	*0.92%*
7	*Brendan Smith*	*0.90%*
8	*B.J. Crombeen*	*0.88%*
9	*Radko Gudas*	*0.79%*
10	*Rene Bourque*	*0.72%*
11	*Joni Pitkanen*	*0.71%*
12	*Steve Downie*	*0.69%*
13	*Alex Petrovic*	*0.68%*
14	*Bryce Salvador*	*0.67%*
15	*Corey Perry*	*0.67%*
16	*Jared Cowen*	*0.66%*
17	*P.K. Subban*	*0.64%*
18	*Eric Selleck*	*0.62%*
19	*Matt Greene*	*0.62%*
20	*Jonathan Toews*	*0.61%*
21	*Mike Weber*	*0.60%*
22	*Brandon Dubinsky*	*0.57%*
23	*Alexander Semin*	*0.56%*
24	*Patrick Kaleta*	*0.56%*
25	*Steve Ott*	*0.56%*

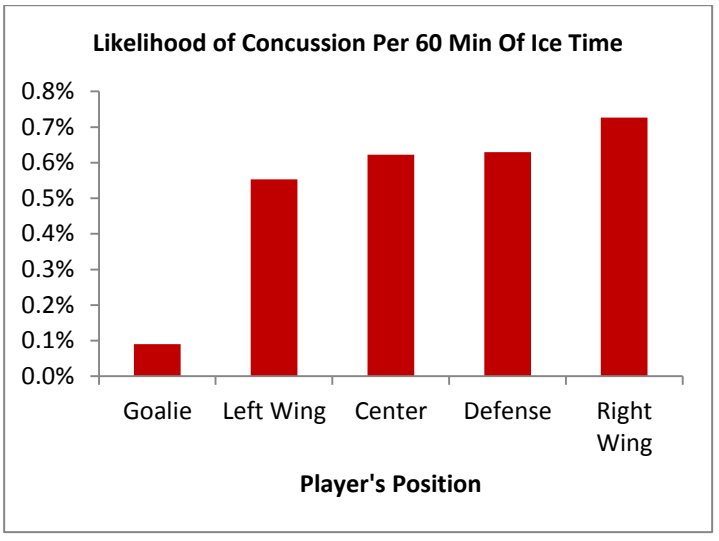
8. Appendices

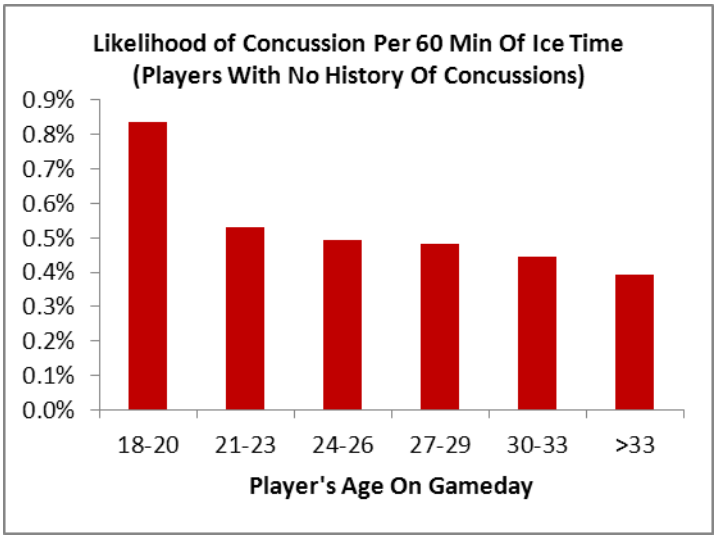
Figure A.1: Graphs Of Concussion Likelihood By Notable Variables

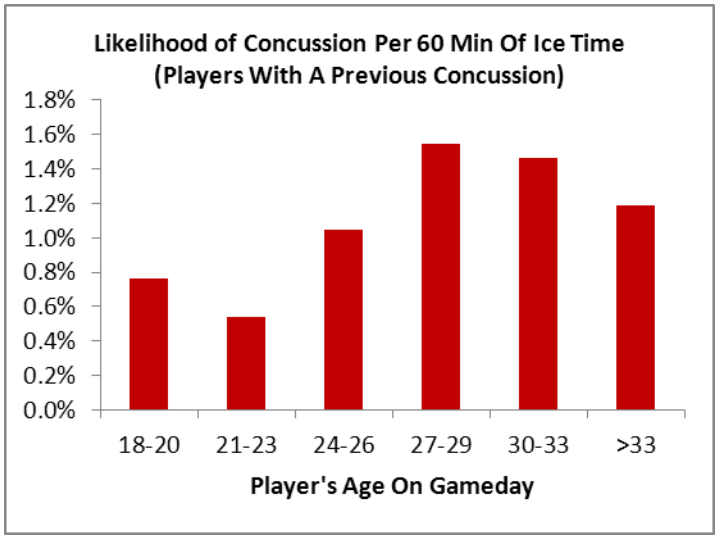

Likelihood of Concussion Per 60 Min Of Ice Time
1.4%
1.2%
1.0%
0.8%
0.6%
0.4%
0.2%
0.0%
No Concussion History
At Least One Concussion
Player's History of Concussions

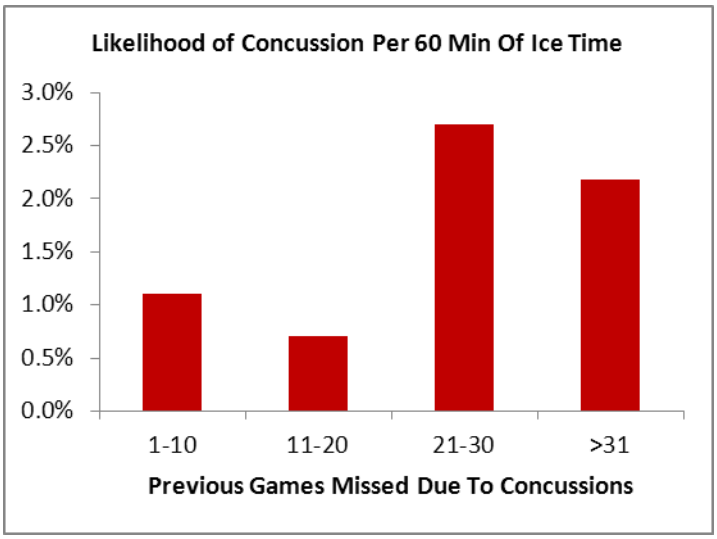

Likelihood of Concussion Per 60 Min Of Ice Time
1.4%
1.2%
1.0%
0.8%
0.6%
0.4%
0.2%
0.0%
66-68
69-71
72-74
75-77
78-81
Player's Height In Inches

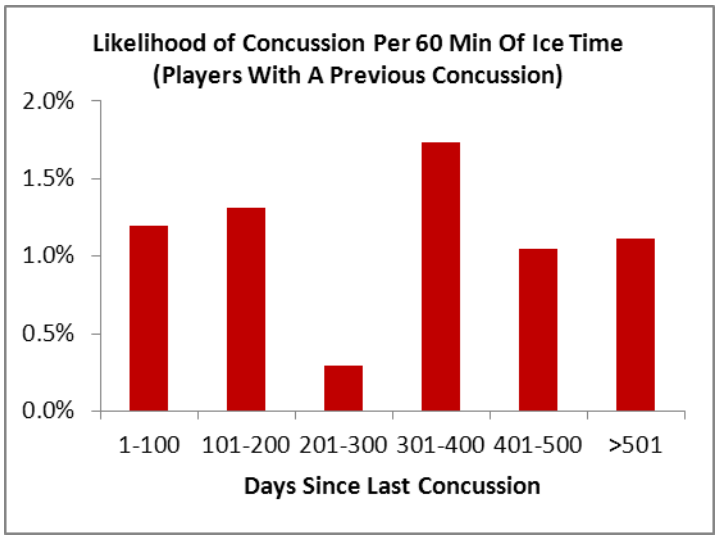

Likelihood of Concussion Per 60 Min Of Ice Time
1.4%
1.2%
1.0%
0.8%
0.6%
0.4%
0.2%
0.0%
0 to 0.2
0.2 to 0.4
0.4 to 0.6
0.6 to 0.8
0.8 to 1
1+
Player's Average PIM Per GameLast Season

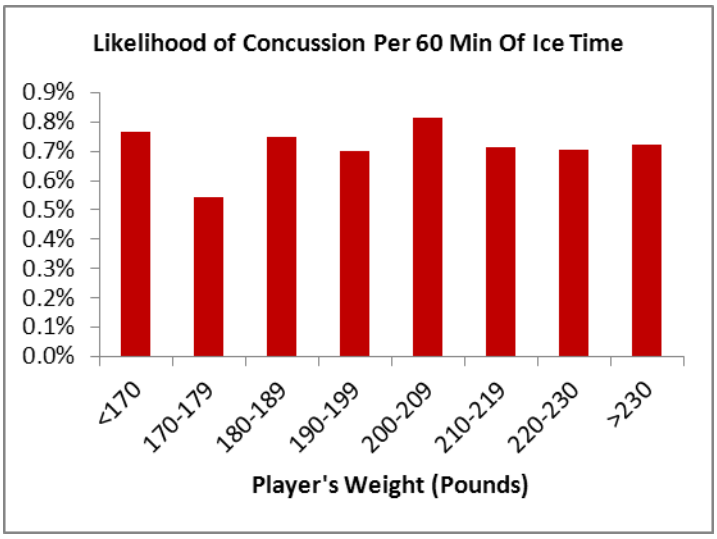

Likelihood of Concussion Per 60 Min Of Ice Time
0.8%
0.7%
0.6%
0.5%
0.4%
0.3%
0.2%
0.1%
0.0%
Goalie
Left Wing
Center
Defense
Right Wing
Player's Position


Likelihood of Concussion Per 60 Min Of Ice Time
(Players With No History Of Concussions)
0.9%
0.8%
0.7%
0.6%
0.5%
0.4%
0.3%
0.2%
0.1%
0.0%
18-20
21-23
24-26
27-29
30-33
>33
Player's Age On Gameday


Likelihood of Concussion Per 60 Min Of Ice Time
(Players With A Previous Concussion)
1.8%
1.6%
1.4%
1.2%
1.0%
0.8%
0.6%
0.4%
0.2%
0.0%
18-20
21-23
24-26
27-29
30-33
>33
Player's Age On Gameday







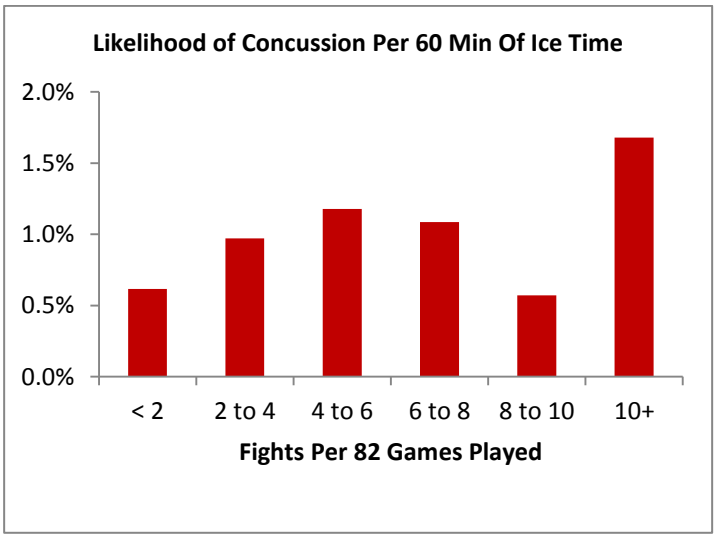
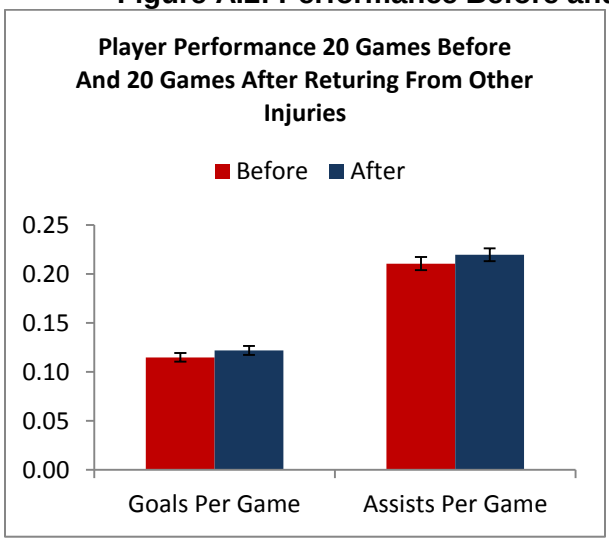


Figure A.2: Performance Before and After Returning From For Non Concussion Injuries



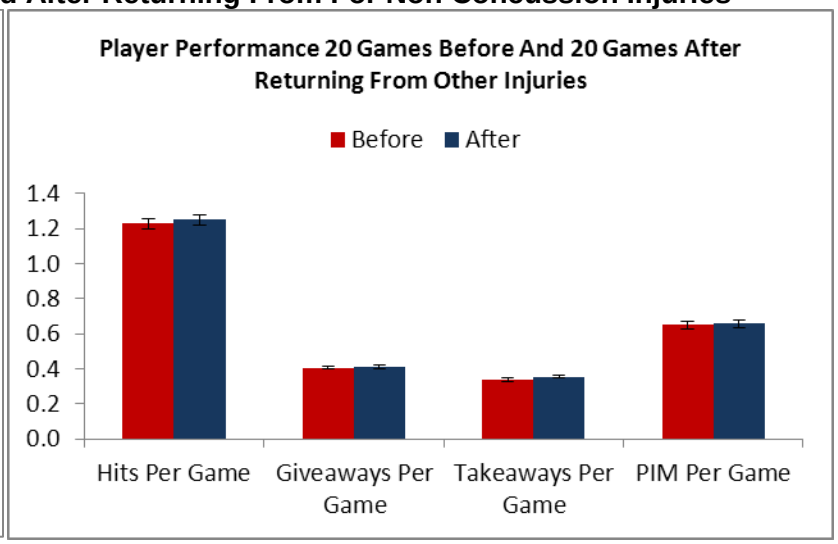


Figure A.3: Predicted Vs. Actual Rate Of Concussions Per 60 Minutes Of Ice Time For The 2013-2014 NHL Season Up To The Olympic Break

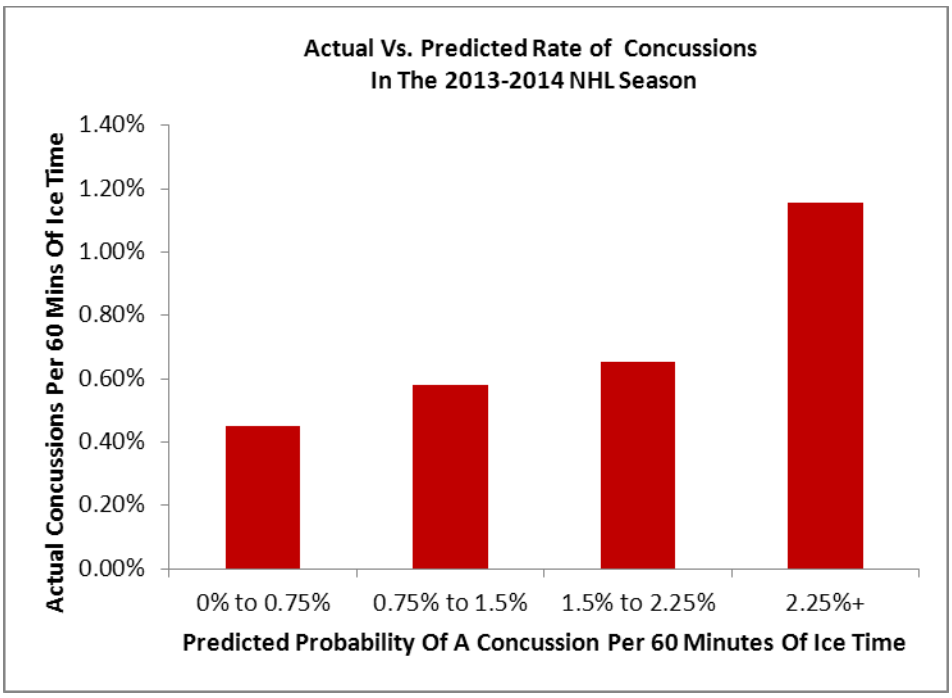


Figure A.4: Predicted Vs. Actual Rate Of Concussions Per Game Played For The 2013-2014 NHL Season Up To The Olympic Break

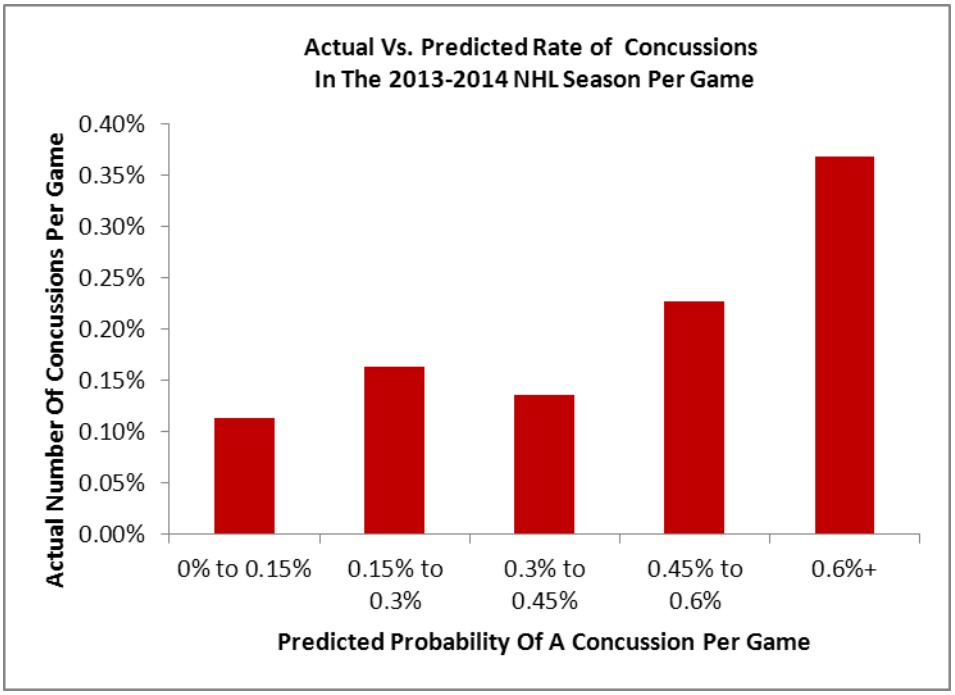


Figure A.5: Rate of Concussions By NHL Season