AN EXAMINATION OF COMPUTERIZED NEUROCOGNITIVE TEST SCORES AS PREDICTORS OF RISK OF CONCUSSION IN COLLEGIATE ATHLETES

by

Laura McKinley Berbert

B.A. Mathematics-Economics, University of Pittsburgh, 2011

Submitted to the Graduate Faculty of

Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Laura M. Berbert

It was defended on

December 8, 2014

and approved by

Thesis Advisor:
Jeanine Buchanich, PhD, Assistant Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Committee Member:
Ada Youk, PhD, Assistant Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Committee Member:
Thomas Songer, PhD, Assistant Professor
Department of Epidemiology
Graduate School of Public Health
University of Pittsburgh

Jeanine M. Buchanich, M.Ed., PhD.

AN EXAMINATION OF COMPUTERIZED NEUROCOGNITIVE TEST SCORES AS PREDICTORS OF RISK OF CONCUSSION IN COLLEGIATE ATHLETES

Laura M. Berbert, M.S.

University of Pittsburgh, 2014

ABSTRACT

Sport-related concussion affects between 1.6 and 3.8 million people annually. Incidence of concussion in the population of collegiate athletes is a public health issue because there are nearly half a million student-athletes at risk every year and physiological effects have the potential to be lifelong and severe in nature. To combat incidence and severity of concussion in college athletes, the CDC and NCAA have partnered to promote education about concussion and implementation of a concussion protocol. At the University of Pittsburgh, this protocol includes a baseline ImPACT test for each incoming student-athlete for the purpose of tracking recovery and facilitating “return to play” decisions following incidence of a concussion. This study explores the utility of the baseline ImPACT test scores in determining risk for concussion. Computerized neurocognitive assessments are developed for post-injury purposes. The analysis in this study uses the ImPACT test scores in a novel way, pre-injury, and examines novel risk factors for concussion, e.g., the composite scores, which were hypothesized to affect incidence through an effect on athletic performance.

Two independent survival analyses were performed: case-wise deletion and multiple imputation by chained equations. Because the data were determined to have MAR missingness, multiple imputation models were adopted in lieu of case-wise deletion models to ensure unbiased

parameter estimates. Two independent models were created: male and female. The model for female athletes included the history indicator and contact expected by sport variables. The model for male athletes included the history indicator, contact expected, processing speed composite score, and height variables. Although these results cannot be generalized with certainty, the results of this study provide an exploratory analysis of the utility of computerized neurocognitive assessments in determining risk for concussion. Four of the six composite scores (visual memory, processing speed, reaction time, and total symptom) were univariably significant (α=0.20) for male athletes. The multivariable model for male athletes included the processing speed composite score. Subsequent studies and analyses, with a focus on securing a representative sample, are required to corroborate the findings of this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank my academic and thesis advisor, Dr. Jeanine Buchanich, for her constant support and advice throughout my academic career in the University of Pittsburgh Graduate School of Public Health. It was important to me to research a question of scientific and medical relevance and a topic of personal interest. I thank Dr. Buchanich for her encouragement when securing such a topic was not always straightforward. Furthermore, I would like to extend my gratitude for her remarks and comments throughout the process of writing my master's thesis.

Additionally, I would like to thank the members of my master's thesis committee, Dr. Ada Youk and Dr. Thomas Songer, for their insight and comments during my research and writing.

I would like to thank my former athletic trainer, Jennifer Csonka, without whom I would not have been able to secure the data set to address my particular research question. Moreover, I would like to thank her for clarifying points of interest in the data throughout this process.

Lastly, I would like to thank my family (Rebecca, Mike, James, Matt, Jessica, and Connor Berbert) for their ever-present and unconditional love, support, and honesty in everything I set out to achieve.

1.0 INTRODUCTION

1.1 BACKGROUND

Concussion is a public health concern that has the potential to affect the quality of life of every family in the nation. Although incidence rates differ by sport, concussion is a possible outcome of any athletic or recreational activity ("Concussion in Sports"). It is estimated that 75% of youth boys, ages 8-17, and 69% of youth girls participate in organized athletics outside of school (Kelley & Carchia, "Hey, Data Data"). Even if children choose to not participate in sports outside of school, most states mandate some level of physical education within their school system (Kuczynski-Brown, "Physical Education"). Therefore, an even larger percentage (≥69%) of American youth ages 5-18 participate in sports and/or recreational activities and are thus potentially exposed to concussion, at varying risk levels.

Likewise, much of the population 18 and older is at some level of risk for concussion. Sport and recreational activities remain a common part of life throughout adulthood. During the 2010-2011 academic year, there were 444,077 student-athletes who competed for a National Collegiate Athletic Association (NCAA) team at the Division I, Division II, or Division III level ("NCAA Participation Rates"). In 2012, nearly half of the American population over the age of five, 141.9 million people, participated in some form of outdoor recreation (*Outdoor Participation*

Report 2013). In the same year, there were 50.2 million health club members ("IHRSA – About the Industry"). Risk of sports-related concussion exists for nearly half of all Americans for much of life, although it varies by factors including type and intensity of recreation.

Currently, concussion is defined as a type of traumatic brain injury (TBI) caused by a bump, blow, or jolt to the head or body. The result is a change in the physiologic function of the brain with respect to memory, behavior, learning, or emotions and is not limited to a singular dysfunction ("Get the Facts about TBI").

This general classification reflects the lack of clarity and uniformity regarding concussion within the medical community, as well as the general population. Similarly, it reflects the individualized nature of concussion, i.e., the inability to develop a definition to be all-inclusive for each individual case. Over time, diagnosis of a concussion has changed from one of specificity to a much broader definition in use today. At one time, loss of consciousness was an integral component. At present, only about 10% of diagnosed concussions include loss of consciousness ("Concussion Signs and Symptoms"). Instead, concussion may be caused by any degree of impact that results in subsequent abnormal brain functioning.

The physiological effects and temporality of concussion often vary by individual. There are as many as 28 different symptoms that may be associated with a concussion. These symptoms are defined within four inclusive categories: physical, cognitive, emotional, and sleep disturbance ("Concussion Signs and Symptoms"). Immediate effects are defined by a change in how the brain functions with respect to memory, behavior, learning, emotions or any combination thereof. Long-term effects include significant life-long impairment of the aforementioned (*Facts about Concussion*). There is uncertainty with regards to long-term effects, specifically the extent to which brain function is affected.

Incidence of concussion in the population of collegiate athletes is a public health issue because there are nearly half a million student-athletes at risk every year and physiological effects have the potential to be lifelong and severe in nature.

The generally accepted estimate of sports-related mild traumatic brain injuries (mTBI) is between 1.6 and 3.8 million annually (*Heads Up: Facts for Physicians*). This figure excludes non-sports related concussions, such as those sustained during vehicular accidents and/or military service, and any severely graded traumatic brain injuries (TBI), sports-induced or otherwise. Nearly a decade ago, the Centers for Disease Control and Prevention (CDC) reported that 300,000 sports-related mild traumatic brain injuries occurred annually (*Heads Up: Concussion in High School Sports*). This estimate was generated under the guidelines that loss of consciousness was a necessary condition for diagnosing a concussion. Using this previously accepted definition and estimate, in combination with the current knowledge that only 10% of diagnosed concussions include a loss of consciousness, an estimate of 3 million concussions annually can be postulated. This estimate falls at the high end of the interval that is generally accepted, between 1.6 and 3.8 million annually.

As evidenced by the drastically different estimates over the span of a decade, the multitude of signs and symptoms that may be associated with concussion, and the broad generalization of the definition of concussion, the diagnosis and treatment of concussion is an ever-changing field.

1.2 NCAA CONCUSSION PROTOCOL

To combat incidence and severity of concussion among collegiate athletes, the CDC and NCAA have partnered to provide information and resources to coaches, student-athletes, and Athletic Department medical staff ("Attention College Sports Fans"). Concussion policy and legislation set forth by the NCAA Executive Committee, as of April 2010, requires affiliated institutions to have a concussion management plan in the event that a student-athlete is suspected of having sustained a concussion. At a minimum, this plan mandates annual education of student-athletes regarding signs and symptoms of concussion, annual acknowledgement of the responsibility to self-report concussion-related injuries, and a process through which student-athletes will be removed from play if a concussion is suspected, disallowed from returning to play on the same day of diagnosis, and medically cleared prior to returning to play following diagnosis (Klossner, *2013-14 NCAA Sports Medicine Handbook*, 64-65).

The major shortcoming of this policy is the emphasis on the existence of a plan versus the application and implementation of a plan. A National Broadcasting Company (NBC) News affiliate reported in October 2014 that nearly 20% of NCAA schools "either don't have the required concussion management plan or have done such a poor job in educating their coaches, medical staff and compliance officers that they are not sure [if] one exists" (Associated Press, "Concussion Study"). The article garners conclusions from a survey-based study published in the *American Journal of Sports Medicine* in October 2014.

The lack of a comprehensive concussion management protocol, with regards to implementation, threatens the well-being of thousands of college athletes whose programs do not have a clearly defined or implemented plan. Concussion is quite frequent in the population of collegiate student-athletes. One study of student-athletes at the conclusion of their intercollegiate athletic careers found that, overall, 49.7% of study respondents reported at least one acknowledged, unreported, or potential concussion. The self-reported rate of concussion was 33.5% (Llewellyn TL, "Concussion Reporting Rates").

1.3 IMPACT TESTS

Immediate Post-Concussion Assessment & Cognitive Testing (ImPACT) was developed in the late 1990s for the purpose of scientifically determining safe "return to play" following the incidence of a concussion. It is the result of collaboration between three clinical experts in the field of neuroscience, Mark Lovell, Ph.D., Joseph Maroon, M.D., and Michael Collins, Ph.D ("About ImPACT: ImPACT Founders"). The test is a neurocognitive assessment tool for trained professionals to help determine safe "return to play". It is neither sufficient as a stand-alone assessment of concussion nor as a tool to aid in diagnosis ("About ImPACT: Overview" *Impact Test*).

The structure of the 25-minute, computerized ImPACT test is consistent across baseline and post-injury formats. Through four sections and six modules, the test measures cognitive function with respect to attention span, working memory, sustained and selective attention time, response variability, non-verbal problem solving, and reaction time ("About ImPACT: The ImPACT Test" *Impact Test*). The structural components of ImPACT are as follows:

Section 1: Demographic Profile and Health History Questionnaire

Section 2: Current Concussion Symptoms and Conditions

Section 3: Baseline and Post-Injury Neurocognitive Tests

- Module 1: Word Discrimination
- Module 2: Design Memory
- Module 3: X's and O's
- Module 4: Symbol Match
- Module 5: Color Match
- Module 6: Three Letter Memory

Section 4: Graphic Display.

The first two sections are self-reported. Section 1 gathers demographic information including age, sport, position, history of concussion, etc. Section 2 gathers information on the current severity of symptoms at the time the test is taken. For each of the 22 symptoms, the test-taker selects the appropriate degree based on a seven-point Likert scale. The first two sections ask identical questions for every test, so that comparisons may be conducted across individuals and across repetitions of testing within the same individual ("The ImPACT Test").

Section 3 is the cognitive assessment portion of ImPACT and consists of the six modules listed above. Each module evaluates and measures multiple different aspects of brain function.

In the first module, word discrimination, the test-taker is presented with a list of 12 words to memorize. The words are listed one at a time and are repeated once to facilitate learning. The test-taker is presented with 24 words, 12 from the prior list and 12 words from the same semantic category. This module assesses verbal memory.

In the second module, design memory, the test-taker is presented with 12 linear designs to memorize. The designs are shown one at a time and are repeated once to facilitate learning. The test-taker is presented with 24 designs, 12 from the prior list and 12 designs that have been rotated in space. This module assesses visual memory.

In the third module, X's and O's, the test-taker must perform a distracter task, e.g., press the left button ("Q") for a blue square and press the right button ("P") for a red circle. The test-taker is presented with a random display of X's and O's, three of which are illuminated in yellow. After 1.5 seconds, the distracter task must be completed. Then, the X's and O's display reappears and the test-taker is prompted to identify the previously illuminated letters. This module assesses visual processing speed and visual working memory.

In the fourth module, symbol match, the test-taker is presented with nine common symbols, which are each correlated with a number from one to nine. A symbol appears below this display and the test-taker must click the appropriate number. Following 27 trials, the nine symbols disappear but the numbers remain. Again, a symbol appears and the test-taker must select the appropriate number through recall. This module assesses visual processing speed.

In the fifth module, color match, the test-taker is presented with one of the following words: red, green, or blue. If the word is displayed in the same colored ink as the word itself, the test-taker is to click the textbox immediately. This module assesses reaction time.

In the sixth module, three letter memory, the test-taker is presented with three letters and then a distracter task. After completing this task, the test-taker must recall the three letters. This module assesses verbal memory.

The results of the third section are combined in such a way as to output five composite scores reflecting the neuropsychological status of the test-taker. The verbal memory composite, visual memory composite, and processing speed composite scores are directly related to performance in these categories. The reaction time composite and impulse control composite scores are inversely related to performance ("The ImPACT Test").

Section 4 consists of the graphically displayed output of each composite score from the test. Additionally, the total symptom composite from section 2 is displayed. This score is merely the summation of values selected by the test-taker for each of the 22 symptoms. A higher score indicates more frequent and/or greater severity of symptoms self-reported by the test-taker. The graphical display in this section allows for immediate comparison of tests across repetitions of testing within the same individual. The five neuropsychological composite scores from section 3 and the total symptom composite score from section 2 were chosen to differentiate between concussed and non-injured individuals through studies conducted by ImPACT ("The ImPACT Test").

1.4 SURVIVAL ANALYSIS

1.4.1 Model & Hazard Ratios

The Cox proportional hazards model is a semi-parametric model of survival time to event. It examines the relationship between survival time and a set of covariates, which enter the model linearly, without specifying the form of the baseline hazard:

$$h(t|\mathbf{X}) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k)$$

where $h_0(t)$ is the unspecified baseline hazard at time t,

$\mathbf{X} = (x_1, x_2, \ldots, x_k)$ is the set of k potential covariates, and

$\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)$ is the set of k coefficients.

The hazard is the rate of the event at any given time. The hazard ratio measures the degree of association between the binary event and two strata. From the model above, the hazard ratio for the p^{th} covariate may be defined as a function of the p^{th} covariate coefficient:

$$HR_p = \exp(\beta_p).$$

For continuous covariates, the hazard ratio quantifies the hazard for a one-unit change in the associated covariate. For categorical covariates, the hazard ratio quantifies the hazard for one stratum versus the hazard for another stratum (Fox, *Cox Proportional-Hazards*).

1.4.2 Cox Proportional Hazards Model Assumptions

Survival analysis, or time to event analysis, provides an assessment of the association between a time-to-event response variable and a set of potential predicting variables. Through this assessment, we can estimate hazard ratios, while adjusting for other variables and identify individuals or groups of individuals at significant risk for the response variable of interest, time to concussion. The Cox proportional hazards model of survival analysis is a generalized proportional (log-relative) hazards model because it does not make any assumptions about the rate or shape of the hazard within the study population. The assumptions of the Cox model are:

1. Time-to-event outcome
2. Non-informative censoring
3. Independence of observations
4. Proportionality of predictors (proportional hazards).

The first few assumptions are empirically assessable. The outcome of interest must be the time until some binary event. A subject is censored if he has incomplete information regarding his survival time. This incomplete information is because either the participant dropped out of the study (lost to follow-up) or did not experience the outcome of interest prior to the end of the designated study period. Survival models incorporate data from uncensored and censored subjects alike to produce parameter estimates. The dependent variable in a survival model is the time to an event rather than the occurrence of the event itself, as in logistic regression models. This controls for variable follow-up times, i.e., censoring. Survival analysis assumes that any censoring is completely random and therefore non-informative (Brant, “Assumptions of the Cox Model”). This assumption depends on the design of the study. It requires that the pattern of censoring be unrelated to the probability of the outcome. Independence of observations is a byproduct of the methods of data acquisition. The proportional hazards assumption must be assessed subsequent to fitting the model. While the Cox model does not make assumptions about the rate or shape of the hazard within the study population, it does assume that the hazard function is proportional over time and across strata of predictors. The relative hazard between groups is assumed to be constant.

1.5 MISSING DATA

1.5.1 Missing Data Mechanism

Analyzing data in the presence of missing values hinges on the specification of the mechanism that causes missing observations. Data may be missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). In the presence of missing data, most analysis techniques proceed with case-wise deletion. Any case with a missing observation is excluded from analysis and thus all of the existing data associated with that case is lost. Implications of case-wise deletion analysis are dependent on the mechanism of the missing data.

The postulation of missing completely at random (MCAR) requires missing values to be independent and unpredictable; i.e., there is no discernable pattern to the missing values. If data are MCAR, then neither the observed variable values in the data set nor the unobserved values of the variable itself are able to predict whether a value will be missing (Howell, "Treatment of Missing Data"). Case-wise deletion analysis in the presence of MCAR data will produce unbiased parameter estimates but may drastically reduce sample size and thus result in large standard errors. Restricting analysis to only complete cases does not yield biased estimates because if the data are missing in an independent and unpredictable way, then the complete cases should reflect a random sample of the population of cases available, with missing values or otherwise ("Deciding to Impute").

The postulation of missing at random (MAR) permits missing values of a variable to be predicted by other variables in the data set. Observed data values can predict whether or not a variable will be missing. In this case, the complete cases may not be a random sample and case-wise deletion analysis may produce biased parameter estimates ("Deciding to Impute").

The postulation of missing not at random (MNAR) proposes that the unobserved value itself predicts whether or not the variable will be missing. Limiting analysis to complete cases would misrepresent any variables that have missing values by this mechanism. The complete cases would not be a random or representative sample. Case-wise deletion analysis would lead to biased parameter estimates (Howell, "Treatment of Missing Data").

1.5.2 Multiple Imputation

Imputation is the process of filling in each missing observation with a plausible value to allow analysis of incomplete data. Multiple imputation repeats this process several times to produce a set of plausible values for each missing observation (*The Multiple Imputation FAQ Page*).

Statistically, multiple imputation has two distinct advantages over single imputation. First, multiple imputation yields more efficient parameter estimates. Sampling variability causes inefficiency. Producing *m* imputations and averaging the parameter estimates over the multiple data sets reduces this variability and thus yields more efficient estimates. Efficiency increases with *m*, but once a certain value of *m* has been reached, the rate of increase in efficiency diminishes. Second, standard error estimates, and thus confidence intervals and p-values, are dependent on the variability of parameter estimates across *m* imputations. Multiple imputation produces a set of plausible values such that the standard error estimates "accurately reflect the uncertainty about the missing values" (Allison, "Why You Probably").

Practically, multiple imputation analysis is most beneficial when the mechanism of missingness is MCAR or MAR. With either case-wise deletion analysis or multiple imputation analysis, parameter estimates are unbiased when data are MCAR. The benefit of multiple imputation in this case is an increased sample size and thus smaller standard errors. When data are MAR, case-wise deletion may produce biased parameter estimates. Multiple imputation, however, produces unbiased parameter estimates. Standard multiple imputation procedures are lacking in the case of MNAR data because the model of missingness is not included.

Efficiency of parameter estimation and accuracy of standard error estimation drive the discussion on the number of imputations *m* that are necessary. Early research focused almost exclusively on parameter estimation. In a 1999 paper, J.L. Schafer concluded that there was "little or no practical benefit to using more than five to ten imputations" with missing data rates as high as about 50% (qtd. in Allison, "Why You Probably"). More recent research questioned the accuracy of standard error estimates based on ten or fewer observations, especially in the presence of high rates of missing data. Such research generally concluded that the number of imputations necessary was comparable to the number of incomplete cases (Allison, "Why You Probably"); e.g., data with 50% incomplete cases (at least one missing observation) requires about 50 imputations. Regardless of the number of multiple imputations *m*, the benefits of multiple imputation outweigh the ease of single imputation.

There are three phases of multiple imputation evaluation: multiple imputation, analysis, and inference. The multiple imputation phase involves filling in the missing values *m* times to produce *m* complete data sets. The analysis phase involves standard analysis procedures for each of the *m* complete data sets. The inference phase involves combining the results from each of the *m* complete data sets.

1.5.3 Multiple Imputation Methods

Standard multiple imputation for MCAR and MAR yields unbiased estimates. If the data are MNAR, however, analysis must include a model of the missingness to produce unbiased estimates, but developing a valid and accurate model for this mechanism is often impractical

(“Deciding to Impute”). There was no evidence to suggest MNAR in this data set. Henceforth, this study focuses on comparing and contrasting analysis techniques under the assumption that the missingness in the given data set is either MCAR or MAR.

Under the assumption of MCAR or MAR, the most common methods for generating multiple imputation data sets are using the multivariate normal model or chained equations. Either approach, will produce unbiased estimates for data with MCAR or MAR missingness. The two approaches have contrasting strengths and weaknesses. Multivariate normal (MVN) imputation derives imputed values from a multivariate normal distribution. The method is, however, robust to non-normal data. Multiple imputation by chained equations (MICE) derives imputed values from whichever distributions are designated. Each imputed variable may have a separate regression designated. MVN has a stronger basis in theory, which translates to better statistical properties. MICE has proven effective in practice ("Multiple Imputation in Stata, Part I"). Due to its assumption that the variables have a multivariate normal distribution, the MVN method does not retain theoretical and thus statistical advantages when imputing binary and categorical variables (“Deciding to Impute”). The MICE method inherently handles imputing binary and categorical variables by allowing each variable to be imputed by a different distribution.

Multiple imputation according to the joint MVN distribution follows from the assumption that the data are distributed according to a multivariate normal distribution. This distribution assumes all independent variables are continuous in nature. For any non-continuous independent variables, this multiple imputation method models and imputes such variables as continuous. Subsequent to imputation, the variable values are reassigned to appropriate discrete values. Joint

MVN multiple imputation is appropriate when all possible independent variables are continuous and jointly normally distributed. There is no statistical justification to support that multiple imputation will yield accurate parameter estimates if the data do not meet this criteria.

MICE is a conditional approach to multiple imputation. It is an iterative, sequential method that models each variable with missing values conditionally on the observed and imputed values for remaining variables. MVN is a special case of MICE for which all independent variables are designated to follow a normal regression distribution. When all independent variables are continuous and jointly multivariate normally distributed, MICE and MVN should yield similar results. Otherwise, the flexibility inherent in MICE should yield more accurate results. In practical application, the majority of data sets are likely to contain at least one discrete variable or not follow an underlying multivariate normal distribution. In these cases, MICE has better practical applications (Kropko, *Multiple Imputation for Continuous and Categorical Data*).

1.5.4 Multiple Imputation Analysis & Inference

The analysis phase involves independently analyzing each of the *m* complete data sets using standard analysis procedures. The inference phase involves combining the *m* results from the analysis of each complete data set. The overall parameter estimate for the p^{th} covariate is derived by averaging the p^{th} parameter estimates across all of the *m* complete data sets:

$$\beta_p^{bar} = (1/m) \sum_{j=1}^{m} \beta_{jp}^{hat}$$

where *m* is the total number of imputations, and

β_{jp}^{hat} is the parameter estimate for the p^{th} covariate in the j^{th} imputation.

The standard error estimate, for the p^{th} covariate is derived by combining the variability within the imputed data sets:

$$V_{W,p} = (1/m) \sum_{j=1}^{m} SE_j$$

where SE_j is the standard error associated with β_{jp}^{hat},

and the variability between all of the imputed data sets:

$$V_{B,p} = \{1/(m-1)\} \sum_{j=1}^{m} (\beta_{jp}^{hat} - \beta_{p}^{bar})^2.$$

The overall standard error estimate associated with the overall p^{th} parameter estimate is the square of the overall variance, which combines the within and between variabilities:

$$V_p = V_{W,p} + \{1+(1/m)\}V_{B,p} \text{ and}$$

$SE_p = \sqrt{V_p}$ (*The Multiple Imputation FAQ Page*).

1.6 PURPOSE OF THE STUDY

The complete eradication of sports-related concussions is not realistic; however, incidence reduction may be achieved through prevention. Prior studies have shown that concussion is quite prevalent in the population of collegiate athletes (49.7% with at least one acknowledged, unreported, or potential concussion), and that prior history of concussion, specific sport, and gender are the most common significant predictors of incidence of concussion (Harmon et al., "Concussion in Sport").

ImPACT is unquestionably the most widespread computerized neurocognitive test used to track recovery of brain function following a sport-induced concussion (“Overview” *Impact Test*). Despite ongoing debate concerning the effectiveness and validity of ImPACT, it is the single best tool currently available to account for the complex, individualized, and poorly understood nature of concussion. It is currently the best option based on its prevalence, as the use of this concussion evaluation system spans a diverse range of ages and levels of competition from high school to professional.

This study explores the utility of using ImPACT beyond its intended purpose to track recovery progress and determine safe “return to play”. It examines the usefulness of using measures from the baseline ImPACT to differentiate risk levels for future concussion among individuals upon entrance to the Pitt Athletic Department. The purpose is to neither support nor discredit ImPACT in its utility to diagnose concussion, as its function is not to assess validity and reliability. Likewise this study neither supports nor discredits ImPACT in its utility to make “return to play” decisions during recovery, as the data set does not address this topic.

Rather, this study serves to link the risk factors of concussion with a proactive, preventative approach to identify college athletes at increased risk for concussion. It will create a model, which the University of Pittsburgh (Pitt) Athletic Department may use in conjunction with a preexisting protocol, e.g. the ImPACT procedures, to identify athletes at increased risk and thus target individuals (or groups of individuals) for prevention measures. Additionally, it will examine potential risk factors for concussion that have rarely been considered by other studies. Of particular interest are the composite scores generated by ImPACT. These scores are formulated to assess neurocognitive function with respect to five different fields, which may be affected in the event of a concussion. This study examines whether baseline measures of these neurocognitive fields affect performance on the field and thereby affect incidence of concussion. For instance, it is biologically plausible that a slower reaction time or processing speed could affect incidence of concussion through an effect on athletic performance. The ultimate goal of this study is to increase the specificity with which prevention methods target at-risk individuals.

2.0 METHODS

2.1 DATA ACQUISITION

The Pitt Athletic Department has a cooperative relationship with the University of Pittsburgh Medical Center's (UPMC) Sports Medicine Concussion Program where many of the developers of ImPACT are established. As such, the Athletic Department and athletic trainers have a comprehensive concussion diagnosis and treatment protocol, which is closely coordinated with world-renowned concussion specialists from UPMC. As of July 2007 (the beginning of the 2007-08 academic year) all student-athletes were required to complete the baseline ImPACT test prior to involvement in university-sponsored athletics. Following a concussion, whether self-reported or reported by an athletic trainer, the injured athlete was to complete a post-injury impact test as proximate to the time of injury as possible. For this reason, the presence of a post-injury test was the indicator for a concussion and the time between the baseline and the first post-injury test was an approximation of the time to concussion. Subsequent post-injury tests were to be taken at the discretion of athletic trainers and any involved neurocognitive experts. Such tests were performed to evaluate the state of recovery of the injured athlete.

There are two mutually exclusive and exhaustive groups of student athletes, those who suffered a concussion while involved with the Athletic Department as a Division I student athlete and those who did not suffer a concussion prior to the end of the study in April 2014 (the end of the 2013-14 academic year). For the first group, baseline data were to be collected upon entrance to the Athletic Department, and post-injury data were to be collected upon assessment of injury. For the latter group, only baseline data were to be collected. The latter group could be further stratified into two groups: those who were still Pitt student-athletes at the end of the study and those who had completed involvement with the Pitt Athletic Department as student-athletes at or prior to the end of the study.

2.2 DATA SET

This project was approved by the University of Pittsburgh Institutional Review Board (approval #PRO14110155, see Appendix B Institutional Review Board (IRB) Approval).

2.2.1 Original Data Set

Every individual who entered the Pitt Athletic Department as of July 2007 was presumed to have at least a baseline ImPACT test upon entrance. Every individual who suffered a reported concussion while a member of the Pitt Athletic Department as of July 2007 was presumed to have at least one post-injury ImPACT test following injury. Every individual who both entered the Pitt Athletic Department as of July 2007 and suffered a reported concussion was presumed to have a baseline and at least one post-injury test. Many individuals had multiple post-injury tests before "return to play" was permitted.

The original data set was comprised of all information obtained from ImPACT tests taken through the Pitt Athletic Department from July 2007 through April 2014. Each test, whether baseline or post-injury, reported the same fields. Based on the type of initial ImPACT test recorded in the Pitt database, there were three mutually exclusive and exhaustive groups of student-athletes: those with an initial baseline test, those with an initial baseline ++ test (invalid baseline), or those with an initial post-injury test (no distinction was made between different levels of post-injury tests at this stage).

The group of student-athletes characterized by an initial baseline ImPACT could be partitioned into three strata by the type of subsequent ImPACT tests recorded in the Pitt database (Figure 1). Those without any subsequent tests were deemed concussion free at the end of the study. Those with only post-injury tests, one or more, were deemed to have suffered a concussion while at Pitt. These first two strata were expected scenarios. Lastly, though, some student-athletes had both subsequent baseline and post-injury tests. With advisement from a Pitt athletic trainer who administered ImPACT tests through the entire study period, it was presumed that these subsequent baseline tests were either post-injury tests misspecified as baseline due to user error or a reevaluation of baseline following a concussion. Either way, these athletes were deemed to have suffered a concussion.

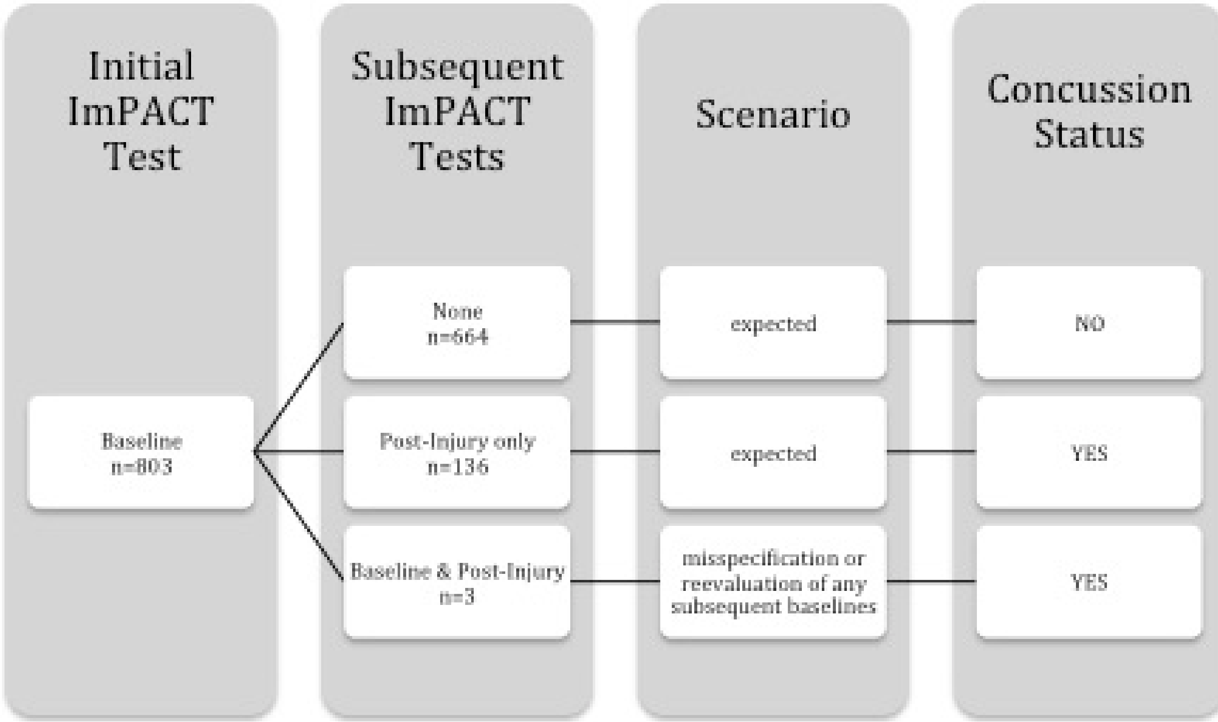

Figure 1. Baseline Initial ImPACT Test

Several athletes had an initial baseline ++ ImPACT test, a designation that indicates an invalid baseline test (Figure 2). A test is determined invalid if it does not provide an accurate account of normal cognitive function (Moser, "Baseline Neuropsychological Tests"). This determination of validity is steered by guidelines but is ultimately subjective. It is the responsibility of the health care professional to determine if there are particular reasons for potentially inaccurate tests. Athletes without a baseline test subsequent to this invalid test, i.e., athletes with no subsequent tests or only subsequent post-injury tests, were excluded from further analysis because baseline values were the variable values of interest in this study. Student-athletes with only a subsequent baseline test were deemed to be concussion free at the end of the study. Student-athletes with a subsequent baseline test and then post-injury tests were deemed to have suffered a concussion while at Pitt.

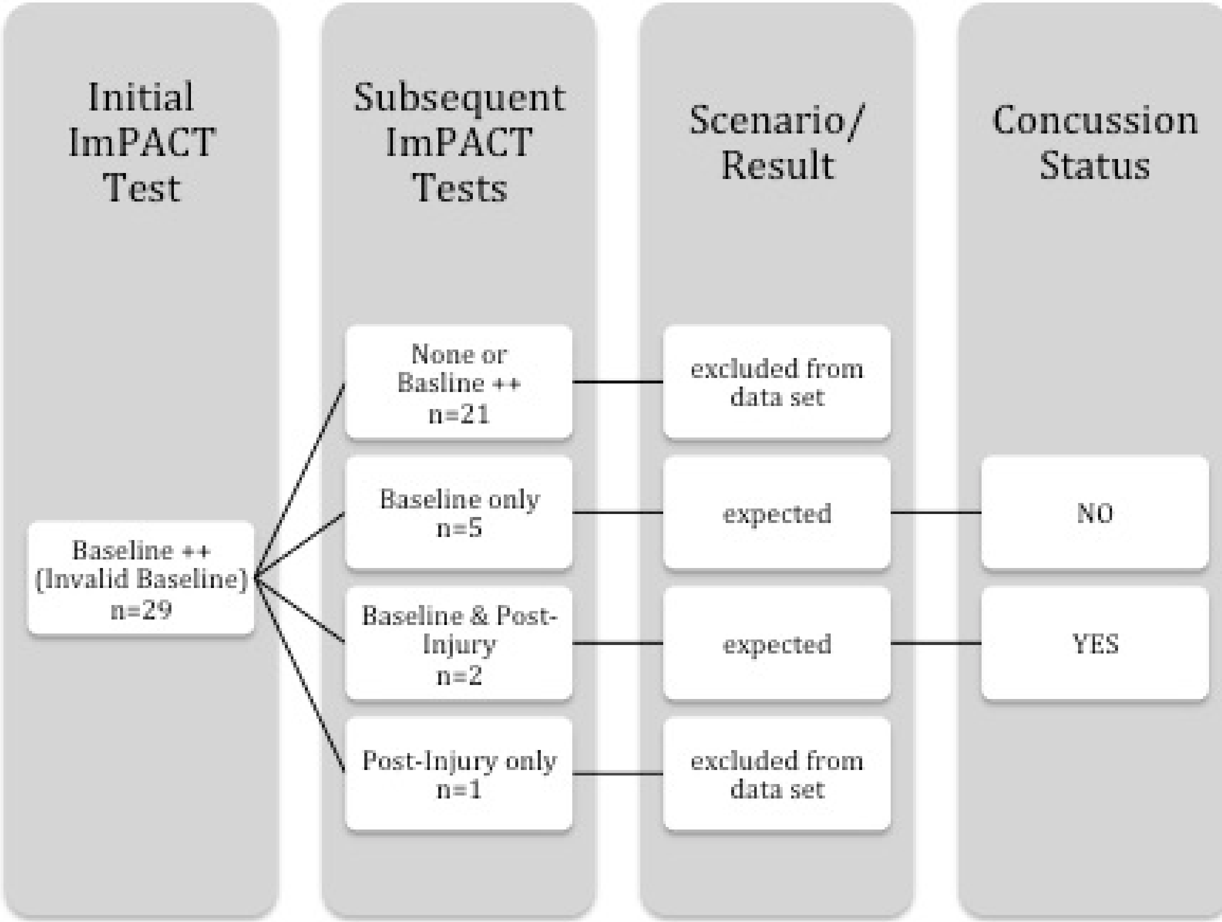

Figure 2. Baseline ++ (Invalid Baseline) Initial ImPACT Test

The last group of athletes had an initial post-injury ImPACT test and either subsequent baseline or post-injury tests or no tests (Figure 3). In other words, athletes belonging to this group did not take a baseline test upon entrance to the Athletic Department. There are three possible explanations for this scenario. First, an athlete may have entered the Athletic Department prior to July 2007 and therefore was not subject to taking a baseline test before participating in athletics. Such athletes would not be representative of the target study population (see 2.4 Study Population). Second, an athlete may have taken a baseline test misspecified as a post-injury test, particularly possible if the athlete had only one post-injury test. Such cases could not be confirmed from the ImPACT data and making assumptions could result in bias. Third, an athlete may not have been subjected to the current protocol. There are multiple trainers attending to hundreds of student athletes at the university in multiple sports. It is possible that some trainers were not as diligent at administering baseline tests for incoming freshmen and transfer students, precluding their inclusion. Such athletes would be missing the data of interest: baseline test values. Data for all athletes without an initial baseline test were dropped from further consideration and analysis, regardless of the category the athlete belonged to.

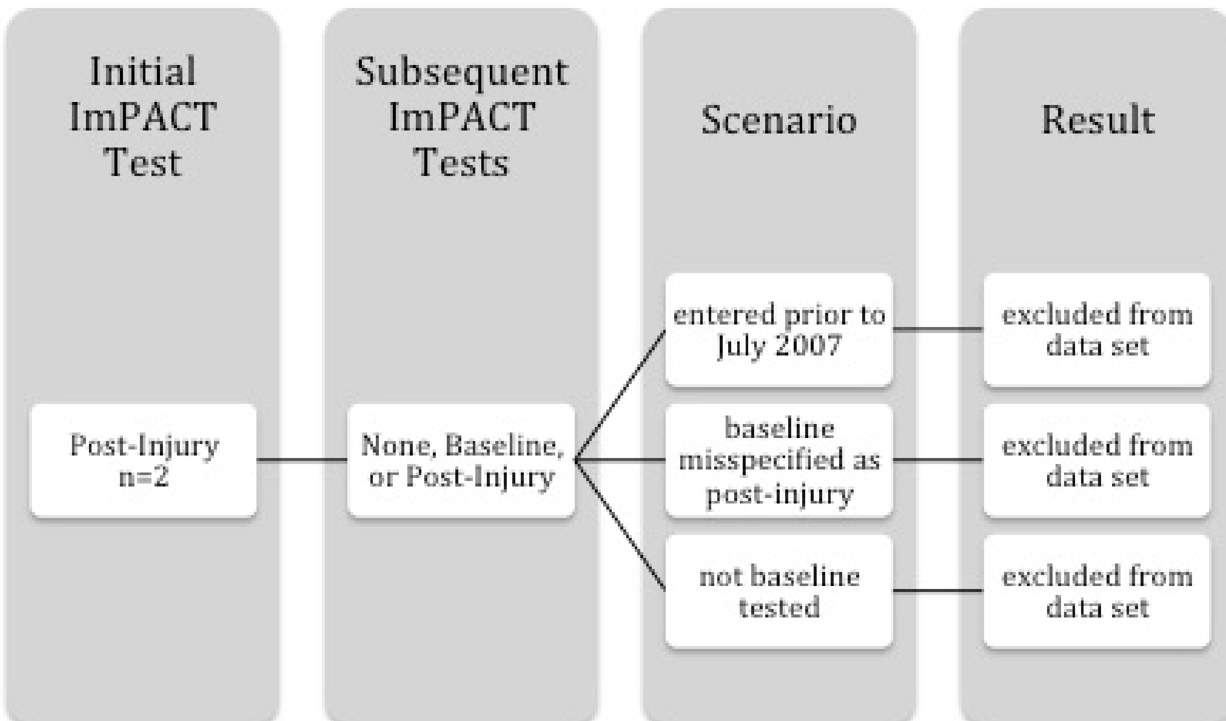

Figure 3. Post-Injury Initial ImPACT Test

2.2.2 New Data Set for Survival Analysis

The format of the original data set (multiple tests per athlete) was not conducive to satisfying the purpose and goals of this study, i.e., the development of a Cox model that assesses risk of concussion for incoming student athletes, and had to be reformatted.

First, the integral variables to survival analysis were established. The cohort entry date (initial event) was the baseline test date, even if the first initial baseline was invalid. The outcome (subsequent event) was concussion or no concussion. The survival time (time to event) for athletes who sustained a concussion was the time from the initial event until the first post-injury test date. The time to event for athletes who were censored, those who did not sustain a concussion, was the time from the initial event until the end of their time at Pitt. The maximum amount of follow-up time was limited to four years (48 months), the typical four-year period of eligibility.

Next, the data needed to be formatted such that each athlete would have one line of data with values reflective of each athlete upon entrance to the Athletic Department. The majority of the original data set was already in this format, including all athletes who had only one baseline test upon entrance and did not suffer a concussion while at Pitt. The remaining data required an adjustment to the original format. Some athletes had an initial baseline test and multiple subsequent post-injury tests (Figure 1). Some athletes had multiple baseline tests (baseline ++ and baseline) and post-injury tests (Figure 2).

By retaining only the valid baseline tests, all invalid baseline test data and all post-injury test data were dropped. A few athletes had multiple baseline tests from either a misspecification or reevaluation. The baseline test with the earliest date was retained. Consequently, all athletes had one line of data that reflected their initial baseline test information in the newly formatted data set.

This reformatting resulted in a data set that would allow for Cox proportional hazards model survival analysis. The new data set was searched for unusual data points through an examination of the variable values (see Appendix A for a detailed description of how unusual observations were managed).

2.3 STUDY PERIOD

The study period was restricted to July 2007 (the beginning of the 2007-08 academic year) through April 2014 (the end of the 2013-14 academic year). Because the baseline values are the potential predicting covariates of interest, this excludes all post-injury tests taken during the study period when there was not an associated valid baseline test taken during the study period and prior to the aforementioned post-injury tests.

2.4 STUDY POPULATION

The target population consisted of all student athletes who competed for one of Pitt's Division I collegiate level sports and entered the Athletic Department during the study period, from the beginning of the 2007-08 academic year to the completion of the 2013-14 academic year. It is unclear how strictly the ImPACT procedures were enforced or followed during this period, thus athletes who were both never subjected to a baseline test and not suspected of a concussion, and therefore without a post-injury test, were excluded from the study population.

2.5 VARIABLES OF INTEREST

2.5.1 Dependent Variable

The Cox proportional hazards model to identify individuals at higher risk for concussion is dependent on the time to incidence of a concussion. The dependent variable has two components: incidence of a concussion and time to the concussion. Incidence of a concussion was determined by the existence of a post-injury ImPACT test. Because the concussion variable is mutually exclusive and exhaustive for two groups, there are no missing values. The time to concussion, given that a concussion was sustained, is approximated by the difference in the cohort entry date (baseline test date) and the first subsequent post-injury test date. The time variable for athletes who did not sustain a concussion is approximated by the difference in the cohort entry date and the end of the study. It was limited to the typical four-year period of eligibility or the end of the study limited the time variable. The data provided no plausible way to differentiate between redshirt athletes, transfer athletes, and four-year athletes (see 4.2.2 Redshirt & Transfer Student-Athletes).

2.5.2 Independent Variables

Given that the overarching purpose of this study is to utilize ImPACT in a preventative capacity, all independent covariates have been derived, either directly or indirectly, from the test itself. The original data set contained over 100 variables, most of which were excluded from analysis due to irrelevancy to the purpose of the study, inaccuracy of responses, or transformation of the data to a newly derived variable. From the ImPACT test data, 23 independent variables were included in subsequent analysis (see Figure 4 for a brief description of included independent variables and see Appendix A for a full description of included and excluded variables).

Variable	Derivation	Description
Sleep	Direct	Amount of sleep the night before the test
Age	Indirect from birthday	Age at the time of the test
Height	Direct	Height in inches at the time of the test
Weight	Direct	Weight in pounds at the time of the test
Verbal Memory	Direct	Composite score from ImPACT
Visual Memory	Direct	Composite score from ImPACT
Processing Speed	Direct	Composite score from ImPACT
Reaction Time	Direct	Composite score from ImPACT
Impulse Control	Direct	Composite score from ImPACT
Total Symptom	Direct	Composite score from ImPACT
Female	Direct	Indicator for female athletes
Nationality	Direct	Indicator for foreign student-athletes
Native Language	Direct	Indicator for non-native English speakers
History	Indirect from number of prior concussions	Indicator for at least one prior concussion
ADD/ADHD	Direct	Indicator for ADD or ADHD
Dyslexia	Direct	Indicator for dyslexia
Special Education	Indirect from special education 1–5	Indicator for self-reported special education designation
Exercise	Direct	Indicator for strenuous exercise prior to the test
Handedness	Direct	Right, left, or ambidextrous
Administrator of Test	Direct	Indicates which medical professional administered the test
Contact Expected	Indirect from sport	Low (*limited contact sports*) swimming, tennis, volleyball Medium (contact sports) softball, baseball, basketball, cheerleading, diving, gymnastics High (contact and collision sports) football, soccer, wrestling
Academic Year of Baseline Test	Indirect from baseline test date	Each academic year spans from July of one year through June of the following year
Symptom Classification	Indirect from total symptom	Based on ImPACT's post-concussion symptom scale for university men & women (Table 15)

Figure 4. Independent Variables Considered

Gender was of particular concern because incidence rates are known to differ by gender within the same sport. The effect of gender is not limited to sports alone. In a model including men and women, the effects of potential predictor variables may be counteracted and balanced by an initial discrepancy due to gender. To preserve as much information as possible and to control for gender as a possible confounder, all of the predictor variables under consideration were examined when stratified by gender.

2.6 COX PROPORTIONAL HAZARDS MODEL ANALYSES

The included independent variables provide demographic information, current level of symptoms, and ImPACT composite scores. All values for the variables were obtained directly or indirectly from the ImPACT tests.

While many of the sports have both male and female athletes, there are a few gender-specific sports with very different expected incidence rates of concussion. Given that prior studies have found gender to be a significant predictor of concussion and the data set includes gender-specific sports, analysis was stratified by gender.

The first step in the actual model building process was running univariable Cox proportional hazards models for each of the potential independent covariates by gender. This provided an assessment of the univariable relationship between each independent covariate and the dependent variable, adjusting only for gender. Thus, these univariable Cox models were a means to ascertain the extent and significance to which each independent covariate could predict the outcome of time to a concussion in male and female athletes.

Using the results from the univariable Cox models, every independent covariate deemed a significant predictor of time to concussion at the α=0.20 level, i.e. p-value<0.20, was included into the appropriate multivariable Cox proportional hazards model. Based on prior studies that have determined history of concussion, gender, and sport to be significant predictors of incidence of concussion, variables that quantified these qualities were forced into the models at every step of selection. Inclusion of gender was achieved through dual models, one for male athletes and one for female athletes. The models including all such covariates were the full models under consideration.

Model selection was conducted using backward elimination with a p-value to remove set at α=0.10. The variable, except for those forced in, with the highest univariable Wald test p-value, for continuous and indicator variables, or likelihood ratio test p-value, for multi-level categorical variables, greater than the specified alpha level was eliminated from the full model and any future models. A new model was fit with all but the previously excluded variable. Again, the variable with the highest p-value greater than the specified alpha was eliminated. The same elimination and refitting process continued until no variables could be eliminated from the model. Selection was finalized when all independent covariates remaining in the model had p-values less than the α=0.10 level, other than those forced in.

The selected models were then scrutinized to confirm the Cox proportional hazards model assumptions were met. The time to event outcome variable, non-informative censoring, and independence of observations assumptions were all empirically assessable through an examination of the study design. The proportionality of predictors assumption was assessed through a Stata revised version of Therneau and Grambsch's test of non-zero slopes as well as an examination of log-log survival plots for low-level categorical variables and Schoenfeld residuals for multi-level categorical variables and continuous variables ("Testing the Proportional Hazard Assumption in Cox Models").

2.7 MULTIPLE IMPUTATION

2.7.1 Multiple Imputation by Chained Equations (MICE)

Due to its practical application and ability to easily incorporate binary and categorical variables, MICE was the chosen approach to impute missing values in the data set. Imputation was done for the entire data set so that all observed values and information were utilized to fill in the missing observations. The number of imputations necessary to yield efficient parameter estimates and accurate standard errors was determined by the percent of incomplete cases in the data. When convergence was not achieved, some of the variables with missing values were excluded from the imputation process. Variables were chosen based on the purpose of the model and univariable Cox proportional hazards model p-values.

2.7.2 Multiple Imputation Analysis & Inference

The analysis and inference of multiply imputed data closely followed the case-delection analysis and inference. Univariable Cox proportional hazards models were run for each of the imputed independent variables by gender. Every imputed independent variable deemed a significant predictor of time to concussion at the $\alpha=0.20$ level, i.e. p-value<0.20, was included into the appropriate multivariable Cox proportional hazards model. Every regular (non-imputed) independent variable deemed a univariably significant predictor from case-wise deletion analysis was also included in the appropriate full model. Variables that quantified history of concussion, gender, and sport were forced into the full model and at every step of selection. Model selection was conducted through backward selection with a p-value to remove set at $\alpha=0.10$. The selected models were scrutinized to confirm the Cox proportional hazards model assumptions were met.

All log-rank tests, simple, pair-wise, and stratified, were analyzed at the $\alpha=0.05$ level.

3.0 RESULTS

3.1 DESCRIPTIVE STATISTICS

The overall Kaplan-Meier survival curve of time to concussion declines at a diminishing rate to just below 80% at the end of four years of follow-up time (Figure 5).

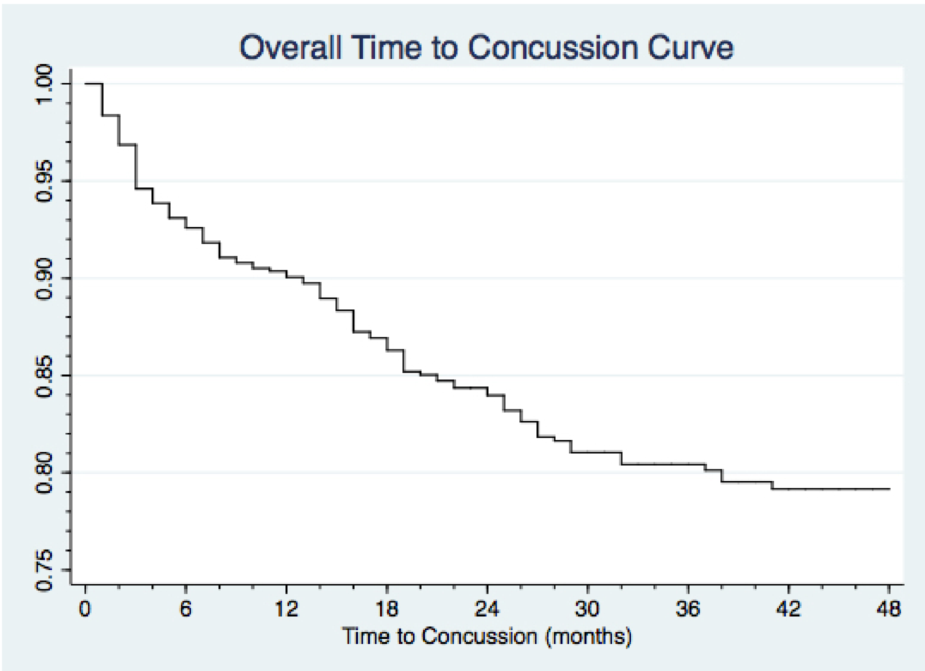

Figure 5. Overall K-M Time to Concussion

Gender is known to be associated with risk level of sports-related concussion. At surface level, the data appears to contradict this commonly held notion. The simple log-rank test for gender indicates that the survival curves for men and women are not significantly different in this population, and therefore, the survival curve for men is not significantly different from the survival curve for women. The Kaplan-Meier survival curves stratified by gender appear more or less equivalent through the first 32 months of follow-up. After this point, the curve for women remains slightly below that for men (Figure 6).

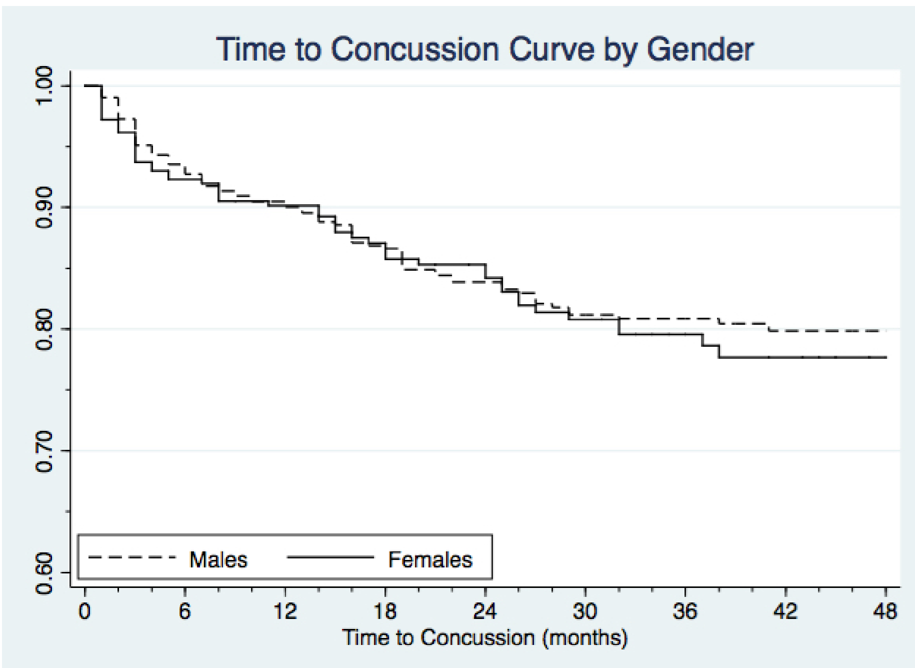

Figure 6. K-M Time to Concussion Stratified by Gender

The simple log-rank test for the three-level categorical redefinition of sport indicates a significant difference between the low and high expected contact groups. Limited contact sports have a significantly different Kaplan-Meier survival curve compared to the survival curve for contact and collision sports (Figure 7).

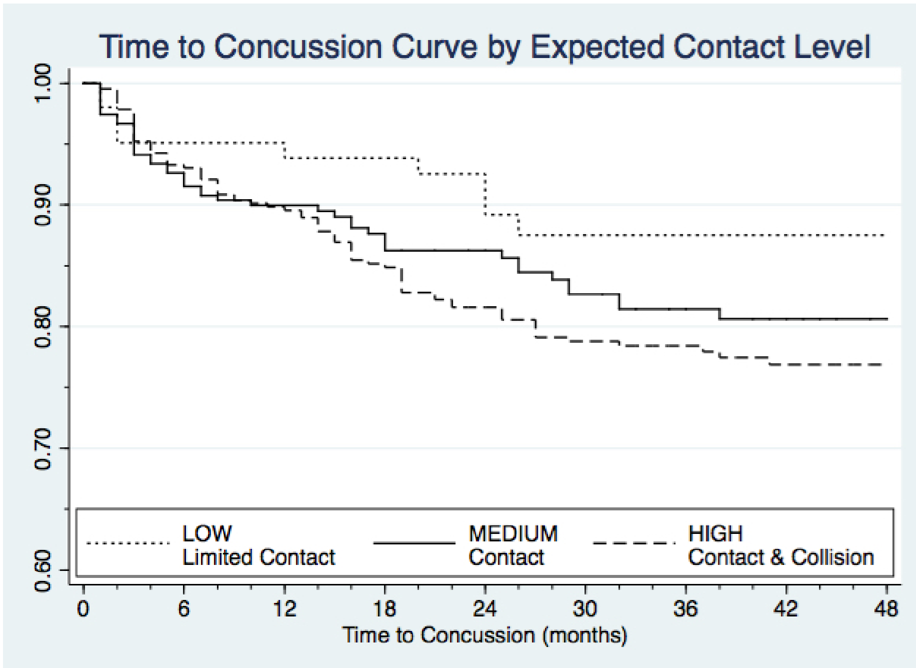

Figure 7. K-M Time to Concussion by Contact Expected

Likewise, a prior history of concussion is associated with risk level of sport-related concussion. The simple log-rank test for history of concussion indicates there is a significant difference in the survival curves of athletes with versus those without a prior history of concussion. The Kaplan-Meier survival curves by prior history of concussion appear equivalent during the first six to eight months of follow-up but diverge thereafter (Figure 8).

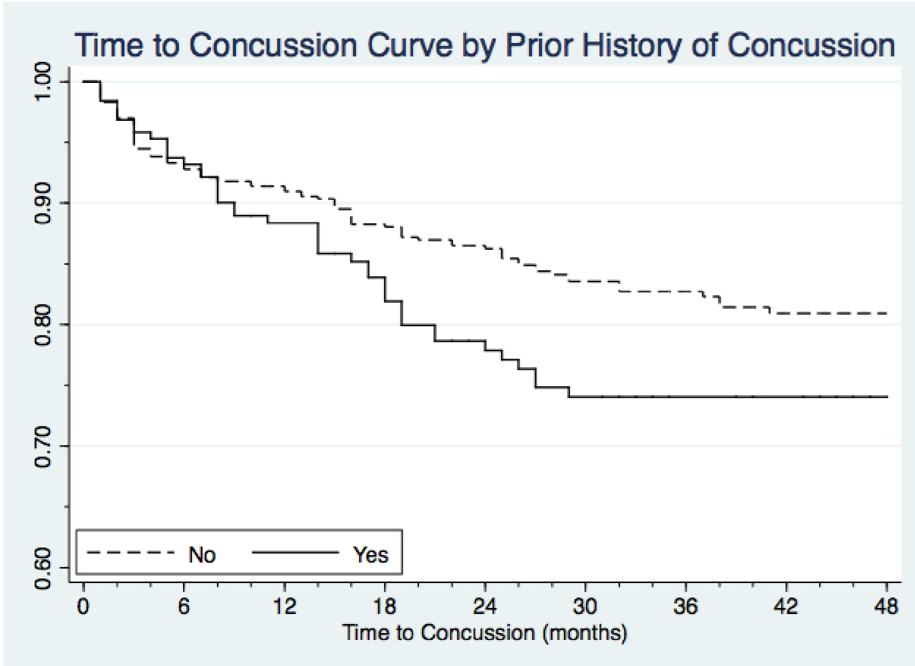

Figure 8. K-M Time to Concussion by Prior History of Concussion

There are a few gender-specific sports with very different person-time incidence rates of concussion. Football and wrestling have only male athletes competing at Pitt with incidence rates of 8.41 and 8.73 concussions per 100 person-years, respectively. Gymnastics, tennis, and volleyball have only female athletes competing at Pitt and incidence rates of 7.08, 0, and 0 concussions per 100 person-years, respectively. For this reason, a comparison of survival based on sport must consider gender. The data were stratified by gender for subsequent analysis because other studies have found gender to be a significant predictor of concussion and the data set includes gender-specific sports. Distributions of potential predictor variables were examined by gender (Table 1) and several are significantly different.

Table 1. Descriptive Statistics by Gender

Variable	Total (N=797)	Gender: *Male* *(n_M=511)*	*Female* *(n_F=286)*	P-value
Sleep*	7 (12)	6.5 (12)	7 (7)	**<0.01**
Age*	18.8 (6.7)	18.9 (6.3)	18.5 (5.6)	**<0.001**
Height*	70 (24)	72 (21)	66 (24)	**<0.001**
Weight*	167 (267)	185 (236)	140 (162)	**<0.001**
Verbal Memory*	87 (49)	86 (49)	88 (46)	**0.02**
Visual Memory*	77 (70)	78 (70)	76.5 (62)	0.24
Processing Speed*	40.8 (41.2)	39.8 (40.2)	42.0 (27.6)	**<0.001**
Reaction Time*	0.56 (0.63)	0.57 (0.6)	0.56 (0.6)	**<0.001**
Impulse Control*	4 (30)	4 (30)	5 (19)	0.46
Total Symptom*	0 (90)	0 (42)	1 (90)	**<0.01**
Nationality†				0.27
American	745 (93.5%)	478 (93.5%)	267 (93.4%)	
Foreigner	43 (5.4%)	24 (4.7%)	19 (6.6%)	
Missing	9 (1.1%)	9 (1.8%)	0 (0%)	
Native Language†				0.69
English	765 (96.0%)	489 (95.7%)	137 (99.28%)	
Other	25 (3.1%)	15 (2.9%)	1 (0.72%)	
Missing	7 (0.9%)	7 (1.4%)	0 (0%)	
History of Prior Concussion†				0.55
No	598 (75.0%)	387 (75.7%)	211 (73.8%)	
Yes	191 (24.0%)	119 (23.3%)	72 (25.2%)	
Missing	8 (1.0%)	5 (1.0%)	3 (1.0%)	
ADD/ADHD†				0.41
No	491 (61.6%)	314 (61.4%)	177 (61.9%)	
Yes	25 (3.1%)	18 (3.5%)	7 (2.4%)	
Missing	281 (35.3%)	179 (35.0%)	102 (35.7%)	

* Continuous, non-normal variables report the median (range) overall and by gender and p-values were calculated using the Kruskal-Wallis one-way analysis of variance by ranks (with ties).

† Categorical variables report the count (column percentage) overall and by gender and p-values were calculated using Pearson's chi-squared test.

P-values that are bold indicate a significant difference (α=0.05).

Table 1. Continued

Variable	Total (N=797)	*Gender*		P-value
		Male *(n_M=511)*	*Female* *(n_F=286)*	
Dyslexia†				0.92
No	511 (64.1%)	327 (64.0%)	184 (64.3%)	
Yes	3 (0.4%)	2 (0.4%)	1 (0.4%)	
Missing	283 (35.5%)	182 (35.6%)	101 (35.3%)	
Special Education†				**<0.01**
No	701 (87.9%)	436 (85.3%)	265 (92.7%)	
Yes	96 (12.1%)	75 (14.7%)	21 (7.3%)	
Exercise†				**<0.01**
No	437 (54.8%)	267 (52.3%)	170 (59.4%)	
Yes	80 (10.0%)	65 (12.7%)	15 (5.2%)	
Missing	280 (35.1%)	179 (35.0%)	101 (35.3%)	
Handedness†				0.24
Right	685 (85.9%)	431 (84.3%)	254 (88.8%)	
Left	79 (9.9%)	55 (10.8%)	24 (8.4%)	
Ambidextrous	31 (3.9%)	23 (4.5%)	8 (2.8%)	
Missing	2 (0.3%)	2 (0.4%)	0 (0%)	
Administrator of Test†				**<0.01**
1	162 (20.3%)	121 (23.7%)	41 (14.3%)	
2	47 (5.9%)	25 (4.9%)	22 (7.7%)	
3	63 (7.9%)	38 (7.4%)	25 (8.7%)	
4	135 (16.9%)	94 (18.4%)	41 (14.3%)	
5	4 (0.5%)	1 (0.2%)	3 (1.0%)	
6	140 (17.6%)	78 (15.3%)	62 (21.7%)	
7	246 (30.9%)	154 (30.1%)	92 (32.2%)	

Table 1. Continued

Variable	Total (N=797)	*Gender*		P-value
		Male *(n_M=511)*	*Female* *(n_F=286)*	
Contact Expected†				**<0.001**
Low	102 (12.8%)	41 (8.0%)	61 (21.3%)	
Swimming	82 (80.4%)	41 (100%)	41 (67.2%)	
Tennis	6 (5.9%)	0 (0%)	6 (9.8%)	
Volleyball	14 (13.7%)	0 (0%)	14 (23.0%)	
Medium	272 (34.1%)	116 (22.7%)	156 (54.5%)	
Baseball/Softball	97 (35.7%)	58 (50.0%)	39 (25.0%)	
Basketball	70 (25.7%)	41 (35.3%)	29 (18.6%)	
Cheerleading	47 (17.3%)	8 (6.9%)	39 (25.0%)	
Diving	23 (8.5%)	9 (7.8%)	14 (9.0%)	
Gymnastics	35 (12.9%)	0 (0%)	35 (22.4%)	
High	418 (52.4%)	352 (68.9%)	66 (23.1%)	
Football	187 (44.7%)	187 (53.1%)	0 (0%)	
Soccer	151 (36.1%)	85 (24.1%)	66 (100%)	
Wrestling	80 (19.1%)	80 (22.7%)	0 (0%)	
Missing	5 (0.6%)	2 (0.4%)	3 (1.0%)	
Academic Year of Baseline†				0.06
2007-08	46 (5.8%)	27 (5.3%)	19 (6.6%)	
2008-09	59 (7.4%)	37 (7.2%)	22 (7.7%)	
2009-10	87 (10.9%)	62 (12.1%)	25 (8.7%)	
2010-11	171 (21.5%)	118 (23.1%)	53 (18.5%)	
2011-12	174 (21.8%)	104 (20.4%)	70 (24.5%)	
2012-13	135 (16.9%)	75 (14.7%)	60 (21.0%)	
2013-14	125 (15.7%)	88 (17.2%)	37 (12.9%)	
Symptom Class†				**<0.01**
Low – Normal	420 (52.7%)	286 (56.0%)	134 (46.9%)	
Normal	238 (29.9%)	127 (24.9%)	111 (38.8%)	
Unusual	88 (11.0%)	62 (12.1%)	26 (9.1%)	
High	33 (4.1%)	24 (4.7%)	9 (3.1%)	
Very High	18 (2.3%)	12 (2.3%)	6 (2.1%)	
Missing	0 (0%)	0 (0%)	0 (0%)	
Concussion†				0.77
No	659 (82.7%)	424 (83.0%)	235 (82.2%)	
Yes	138 (17.3%)	87 (17.0%)	51 (17.8%)	

The stratified log-rank test for sport by gender indicates that the survival curves for male and female athletes are not different for each sport in this population. Because other studies have found gender to be significantly associated with risk of concussion in some sports, comparisons of gender were planned for each sport. The comparison for each sport was thus considered independent of the other sports. The simple log-rank test for gender among only soccer players indicates that the curves are different when restricted to the sport of soccer (Figure 9). All other sports with male and female athletes (softball/baseball, basketball, cheerleading, diving, and swimming) did not have significantly different survival curves for men and women, according to simple log-rank tests restricted to each sport independently.

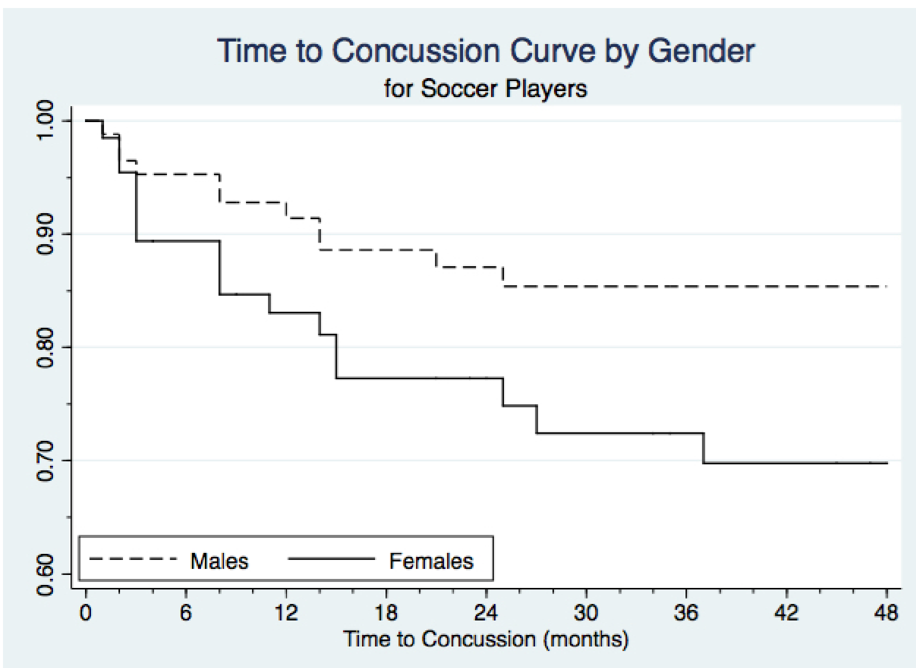

Figure 9. K-M Soccer-Specific Time to Concussion Stratified by Gender

The stratified log-rank test for the contact expected variable by gender echoes the significant results from the simple log-rank test. The null hypothesis of equivalent survival curves for each contact expected group was rejected using the simple log-rank test with $\alpha=0.05$. Based on further exploration with stratified log-rank tests, we hypothesize the differences are dependent on gender. When restricting analysis to male athletes, there is not a difference in survival curves between any of the expected contact groups (Figure 10). When restricting analysis to female athletes, there is a difference in survival by contact expected, specifically among the low and high expected contact groups. (Figure 10).

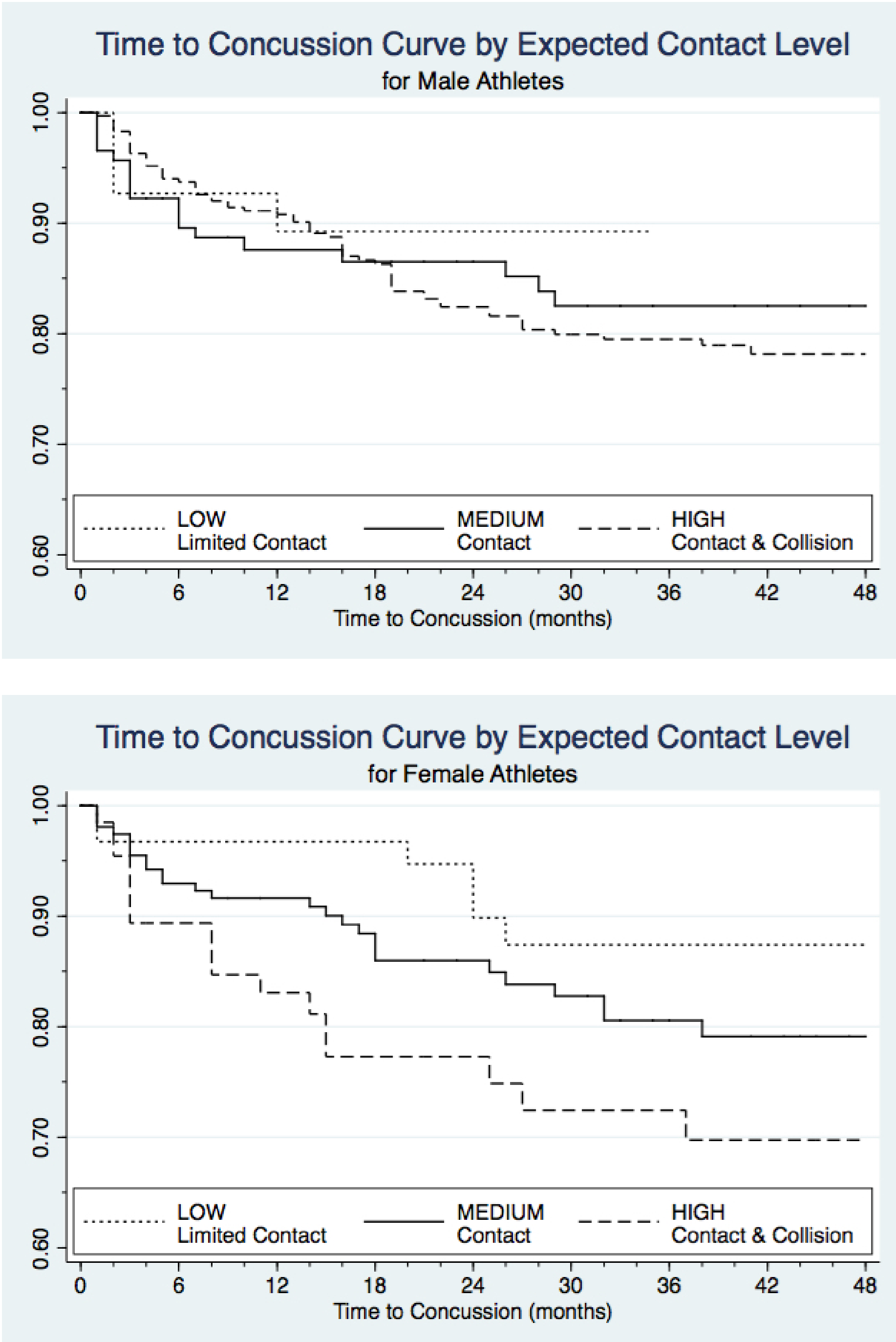

Figure 10. K-M Time to Concussion by Contact Expected, Stratified by Gender

The stratified log-rank test for prior history of concussion by gender reverberates the significant results from the simple log-rank test. The null hypothesis of equivalent survival curves for athletes with versus without a prior history of concussion was rejected using the simple log-rank test with $\alpha=0.05$. Based on further exploration with stratified log-rank tests, we hypothesize the differences are dependent on gender. When restricting analysis to male athletes, there is not a difference in survival curves between athletes with a prior history of concussion and athletes without (Figure 11). When restricting analysis to female athletes, there is a difference in the survival curves between each group (Figure 11).

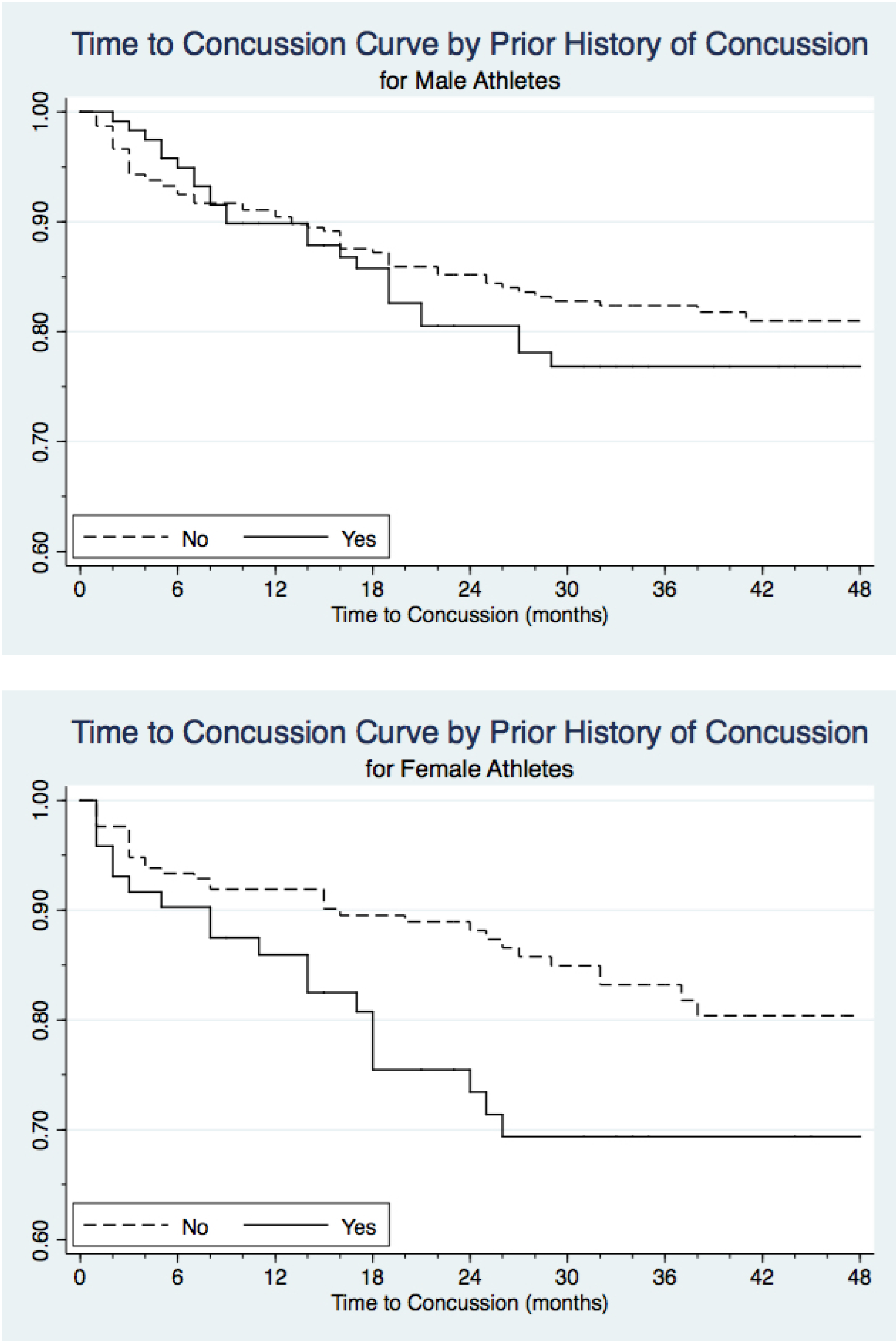

Figure 11. K-M Time to Concussion by Prior History of Concussion, Stratified by Gender

3.2 UNIVARIABLE COX PROPORTIONAL HAZARDS MODELS

Univariable Cox models were run to ascertain the extent and significance to which each independent covariate could predict the outcome of time to concussion. Such models were performed in two different scenarios: unadjusted and adjusted for gender.

The Cox models, without adjusting for gender, assess the unadjusted relationship between each independent covariate and the dependent variable (Table 2). The variables that characterize age, reaction time composite score, history of prior concussion, ADD/ADHD, strenuous exercise prior to the test, and expected level of contact by sport are all univariably significant (at the α=0.20 level) in predicting time to concussion.

Table 2. Unadjusted Results

Variable	Hazard Ratio	P-value	AIC Statistic
Sleep	1.04	0.51*	1298.71
Age	0.87	**0.09***	1759.34
Height	1.01	0.73*	1747.23
Weight	1.00	0.75*	1760.56
Female	1.08	0.65*	1762.22
Verbal Memory	1.00	0.91*	1762.41
Visual Memory	0.99	0.24*	1761.07
Processing Speed	0.99	0.31*	1761.41
Reaction Time	5.49	**0.12***	1760.13
Impulse Control	0.98	0.48*	1761.92
Total Symptom	1.01	0.50*	1762.00
Nationality	0.94	0.88*	1732.9
Native Language	0.47	0.29*	1744.81
History	1.49	**0.03***	1715.59
ADD/ADHD	2.12	**0.04***	953.12
Dyslexia	1.70e-15	>0.99*	942.96
Special Education	1.00	>0.99*	1762.42
Exercise	0.48	**0.048***	951.82
Handedness		0.43†	1762.18
Right (Baseline)			
Left	0.87	0.64*	
Ambidextrous	0.52	0.27*	

* P-values calculated using the univariable Wald test.
† Overall p-values (categorical variables with >2 levels) calculated using the likelihood ratio test.
P-values that are bold indicate a significant difference (α=0.20).

Table 2. Continued

Variable	Hazard Ratio	P-value	AIC Statistic
Administrator of Test		0.97†	1771.01
1 (Baseline)			
2	0.78	0.58*	
3	1.21	0.59*	
4	1.13	0.67*	
5	1.32	0.78*	
6	1.18	0.54*	
7	1.03	0.90*	
Contact Expected		**0.09**†	1731.78
Low (Baseline)			
Medium	1.58	0.195*	
High	1.92	0.05*	
Academic Year of Baseline Test†		0.23†	1764.31
2007-08 (Baseline)			
2008-09	0.52	0.19*	
2009-10	1.24	0.58*	
2010-11	1.05	0.89*	
2011-12	0.72	0.38*	
2012-13	1.19	0.65*	
2013-14	1.06	0.90*	
Symptom Classification		0.41†	1764.44
Low-Normal	1.05	0.81*	
Normal (Baseline)			
Unusual	1.63	0.07*	
High	1.02	0.96*	
Very High	1.55	0.41*	

The Cox models adjusted for gender assess the univariable relationship between each independent covariate and the dependent variable, while controlling for gender. Stratification by gender provides greater specificity in assessing the relationships univariably (Table 3), particularly for independent variables with significantly different distributions by gender (Table 1).

Table 3. Univariable Results by Gender

Males			Variable	Females		
Hazard Ratio	**P-value**	**AIC Statistic**		**Hazard Ratio**	**P-value**	**AIC Statistic**
1.02	0.75*	741.07	***Sleep***	1.06	0.60*	417.44
0.92	0.38*	1032.33	***Age***	0.79	**0.12***	547.13
1.04	**0.19***	1017.77	***Height***	0.99	0.81*	549.31
1.00	**0.20***	1030.08	***Weight***	0.99	**0.18***	547.68
1.00	0.66*	1032.94	***Verbal Memory***	1.00	0.72*	549.75
0.99	**0.09***	1030.4	***Visual Memory***	1.00	0.72*	549.74
0.97	**0.03***	1028.59	***Processing Speed***	1.03	**0.20***	548.17
6.34	**0.16***	1031.24	***Reaction Time***	5.02	0.45*	549.32
0.98	0.53*	1032.71	***Impulse Control***	0.99	0.74*	549.76
1.02	**0.17***	1031.48	***Total Symptom***	1.00	0.81*	549.82
1.00	0.99*	1005.57	***Nationality***	0.86	0.81*	549.81
0.82	0.78*	1017.94	***Native Language***	1.62e-15	1.00*	546.27
1.28	0.30*	1005.81	***History***	1.86	**0.03***	533.26
2.55	**0.03***	522.11	***ADD/ADHD***	1.44	0.62*	322.20
3.41e-14	>.99*	513.59	***Dyslexia***	4.58e-15	>.99*	321.82
1.13	0.67*	1032.96	***Special Education***	0.69	0.54*	549.45
0.72	0.39*	524.95	***Exercise***	1.81e-16	1.00*	313.97
	0.44†	1032.94	***Handedness***		0.94†	551.74
			Right (Baseline)			
0.82	0.60*		*Left*	0.98	0.97*	
0.47	0.29*		*Ambidextrous*	0.71	0.73*	
	0.62†	1038.73	***Administrator of Test***		0.26†	552.13
			1 (Baseline)			
0.73	0.62*		*2*	0.78	0.72*	
1.01	0.99*		*3*	1.44	0.51*	
0.97	0.94*		*4*	1.45	0.45*	
9.28	0.03*		*5*	2.04e-19	1.00*	
0.95	0.88*		*6*	1.42	0.44*	
1.28	0.40*		*7*	0.63	0.34*	

* P-values calculated using the univariable Wald test.
† Overall p-values (categorical variables with >2 levels) calculated using the likelihood ratio test.
P-values that are bold indicate a significant difference (α=0.20).

Table 3. Continued

Males			Variable	Females		
Hazard Ratio	P-value	AIC Statistic		Hazard Ratio	P-value	AIC Statistic
	0.48†	1033.01	***Contact Expected***		**0.07**†	523.29
			Low (Baseline)			
1.39	0.56*		*Medium*	1.72	0.23*	
1.67	0.32*		*High*	2.79	0.03*	
	0.32†	1036.15	***Academic Year of Baseline Test***		**0.03**†	546.08
			2007-08 (Baseline)			
1.75	0.42*		*2008-09*	4.45e-17	1.00*	
2.54	0.14*		*2009-10*	0.64	0.43*	
1.92	0.29*		*2010-11*	0.73	0.50*	
1.22	0.76*		*2011-12*	0.52	0.17*	
2.60	0.14*		*2012-13*	0.61	0.35*	
1.96	0.33*		*2013-14*	0.72	0.61*	
	0.28†	1034.03	***Symptom Classification***		0.44†	552.14
1.22	0.49*		*Low-Normal*	0.92	0.79*	
			Normal (Baseline)			
1.81	0.09*		*Unusual*	1.56	0.31*	
1.14	0.82*		*High*	0.96	0.96*	
2.99	0.05*		*Very High*	4.45e-15	1.00*	

In male athletes, the variables that characterize height, weight, visual memory composite score, processing speed composite score, reaction time composite score, total symptom composite score, and ADD/ADHD are all univariably significant (at the α=0.20 level) in predicting time to concussion. In female athletes, the variables that characterize age, weight, history of prior concussion, processing speed composite score, expected level of contact by sport, and academic year of the baseline test are all univariably significant (at the α=0.20 level) in predicting time to concussion. The weight and processing speed composite score variables are the only variables significant for both genders.

3.3 MULTIVARIABLE COX PROPORTIONAL HAZARDS MODELS

3.3.1 Full Models

Multivariable model selection analysis only considered variables with univariable p-value less than an α=0.20 level and forced-in variables. A full Cox model for male athletes was fit with these significant and known predictors of time to concussion: height, weight, visual memory composite score, processing speed composite score, reaction time composite score, total symptom composite score, ADD/ADHD, history of prior concussion, and expected level of contact by sport. The latter two variables were forced into the full model based on results from previous studies.

Likewise, a full Cox model for female athletes was fit with significant and known predictors of time to concussion: age, weight, processing speed composite score, history of prior concussion, and expected level of contact by sport.

3.3.2 Model Selection

The initial four steps of model selection for male athletes eliminated each of the univariably statistically significant ImPACT composite scores. The last two steps eliminated weight and height from the model (Table 4).

Table 4. Backward Selection Results for Male Athletes

Variable	Step Removed	Hazard Ratio	P-value*
Reaction Time Composite Score	1	0.91	0.96
Total Symptom Composite Score	2	0.99	0.80
Memory Visual Composite Score	3	1.00	0.79
Processing speed Composite Score	4	0.98	0.35
Weight	5	0.99	0.24
Height	6	1.06	0.19

* P-values calculated using the univariable Wald test.

There were only two steps of model selection for female athletes (Table 5). The first step eliminated weight and the second (final) step eliminated the processing speed composite score from the model.

Table 5. Backward Selection Results for Female Athletes

Variable	Step Removed	Hazard Ratio	P-value*
Weight	1	1.00	0.54
Processing speed Composite Score	2	1.04	0.12

* P-values calculated using the univariable Wald test.

3.3.3 Selected Models

The selected model for male athletes is comprised of the ADD/ADHD indicator, history of prior concussion indicator, and expected level of contact by sport variables (Table 6). The latter two variables were forced into the model at one or more stages of selection.

Male athletes who self-reported ADD or ADHD have a hazard that is 2.52 times the hazard for male athletes who did not report having either disorder. Males with a prior history of concussion have a hazard that is 1.41 times the hazard for males without any prior concussions. Males competing in medium contact expected sports (baseball, basketball, cheerleading, and diving) and males competing in high contact expected sports (football, soccer, and wrestling) have higher hazards, 1.75 and 1.35 times respectively, than males competing in low contact expected sports (swimming). The effect size of medium versus low contact sports (1.75) is actually greater than the effect size of high versus low contact sports (1.35).

Table 6. Variables in Selected Model for Male Athletes

Variable	Hazard Ratio	P-value
ADD/ADHD	2.52	**0.04***
History	1.41	0.28*
Contact Expected		0.57†
Low (Baseline)		
Medium	1.75	0.33*
High	1.35	0.57*

* P-values calculated using the univariable Wald test.
† Overall p-values (categorical variables with 2 or more levels) calculated using the likelihood ratio test.
P-values that are bold indicate a significant difference (α=0.10).

The selected model for female athletes is comprised of the age, history of prior concussion, and expected level of contact by sport variables (Table 7). The latter variable was the only one that had to be forced into the model at one or more stages of selection.

The hazard with respect to age is inversely related in female athletes. With each increasing year, the hazard declines 0.68 times. Equivalently, female athletes one year younger have a hazard that is 1.47 times higher. Females with a prior history of concussion have a hazard that is 1.92 times the hazard for females without any prior concussions. Females competing in medium contact expected sports (softball, basketball, cheerleading, diving, and gymnastics) and females competing in high contact expected sports (soccer) have higher hazards, 1.38 and 1.78 times respectively, than females competing in low contact expected sports (swimming, tennis, and volleyball).

Table 7. Variables in Selected Model for Female Athletes

Variable	Hazard Ratio	P-value
Age	0.68	**0.04***
History	1.92	**0.03***
Contact Expected		0.49†
Low (Baseline)		
Medium	1.38	0.49*
High	1.78	0.25*

* P-values calculated using the univariable Wald test.
† Overall p-values (categorical variables with 2 or more levels) calculated using the likelihood ratio test.
P-values that are bold indicate a significant difference (α=0.10).

The female and male models selected both contain the history of prior concussion(s) indicator and the contact expected by sport variables. During selection for the male model, both of these variables were forced into the model at one or more stages. During selection for the female model, only the contact expected variable had to be forced in. As expected in this scenario, the history variable is more significant in the female model. The effect size is also larger in the female model such that prior concussion history versus no history has a 1.92 times higher hazard for females and only a 1.41 times higher hazard for males. In males and females, the contact expected variable is non-significant but was forced into the model. As noted earlier, the effect size of medium versus low contact sports (1.75) is actually higher than the effect size of high versus low contact sports (1.35) for male athletes. For female athletes, there are similar effect sizes but they are flipped and more in line with what would be expected: 1.38 times higher in medium versus low contact sports and 1.78 times higher in high versus low contact sports.

3.3.4 Checking Model Assumptions

The first few assumptions are empirically assessable. The outcome of interest for this study is whether or not an athlete sustained a concussion while competing at Pitt. By the nature of this study, the response variable is dependent on time because some athletes were censored due to the end of the study. Thus, the outcome must be addressed as the time until a concussion was sustained, or time to the event, to account for this censoring. The primary reason to use survival analysis for this data set is to control for athletes with variable follow-up times, i.e., censored data.

Survival analysis assumes that any such censoring is completely random and therefore non-informative (Brant, “Assumptions”). This assumption depends on the design of the study. It requires that the pattern of censoring be unrelated to the probability of the outcome. In this data set, all censoring is dependent on the cohort entry date. Athletes who had not sustained a concussion by the end of the study were censored. The maximum amount of follow-up time was four years, or equivalently 48 months, to reflect the typical four-year period of eligibility. Thus, any athletes with follow-up time that exceeded this were reduced to the maximum.

Independence of observations is a byproduct of the methods of data acquisition. For this study, there is some concern regarding this assumption. Although uncommon, there are siblings who attended Pitt and participated in the Athletic Department. Since the data set provides no way to determine which observations are related, it must be assumed that dependence among observations is quite rare. In this case, rare dependence appears to be a valid assumption. This, however, should be considered during application of the model and any further development.

The proportional hazards assumption must be assessed subsequent to fitting the model. While the Cox model does not make assumptions about the rate or shape of the hazard within the study population, it does assume that the hazard function is proportional over time and across strata of predictors. The relative hazard between groups is assumed to be constant.

For the model of male athletes, the Stata revised version of Therneau and Grambsch’s test of non-zero slopes indicates that the ADD/ADHD indicator variable is the only predictor to violate the proportional hazards assumption. While the log-log survival plots for the ADD/ADHD indicator, history indicator, and contact expected variables are inconclusive (Figure 12), the plots of Schoenfeld residuals versus time for each predictor, except history, have a few erratic observations (Figure 13).

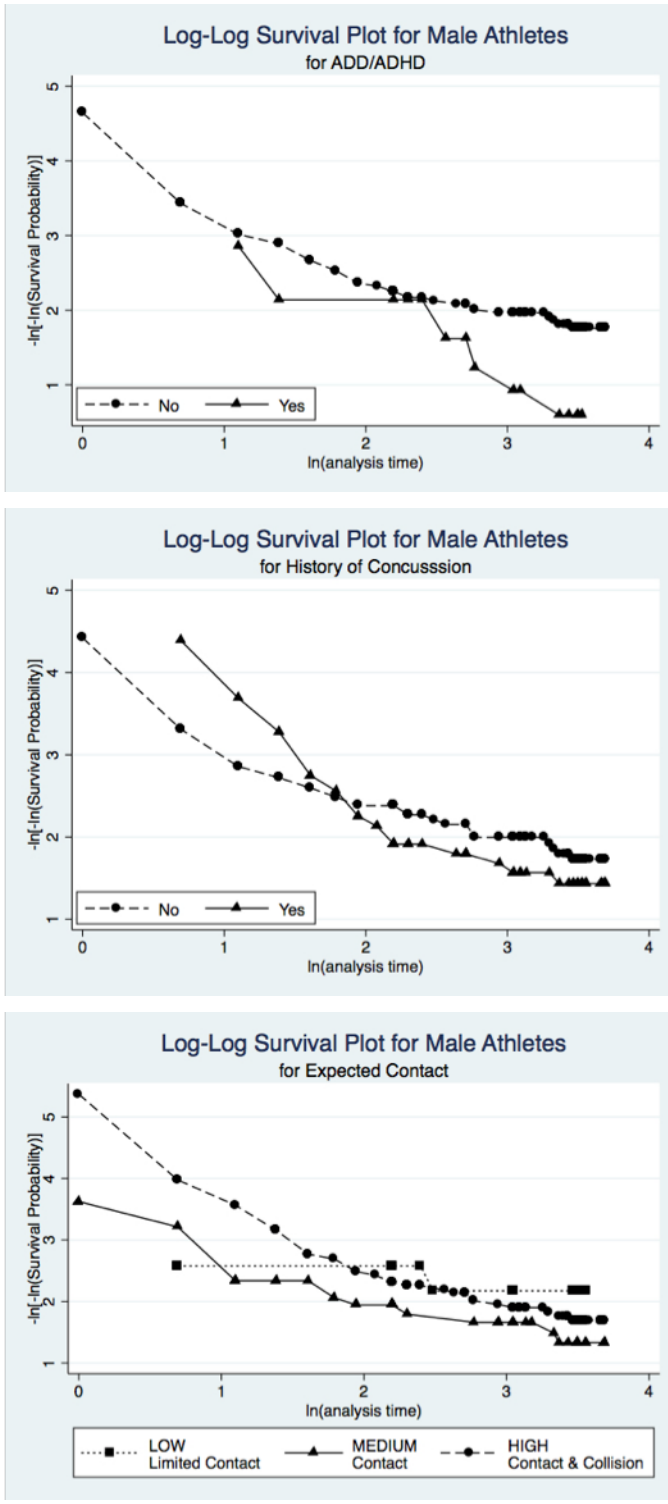

Figure 12. Log-Log Survival Plots for Selected Model of Male Athletes

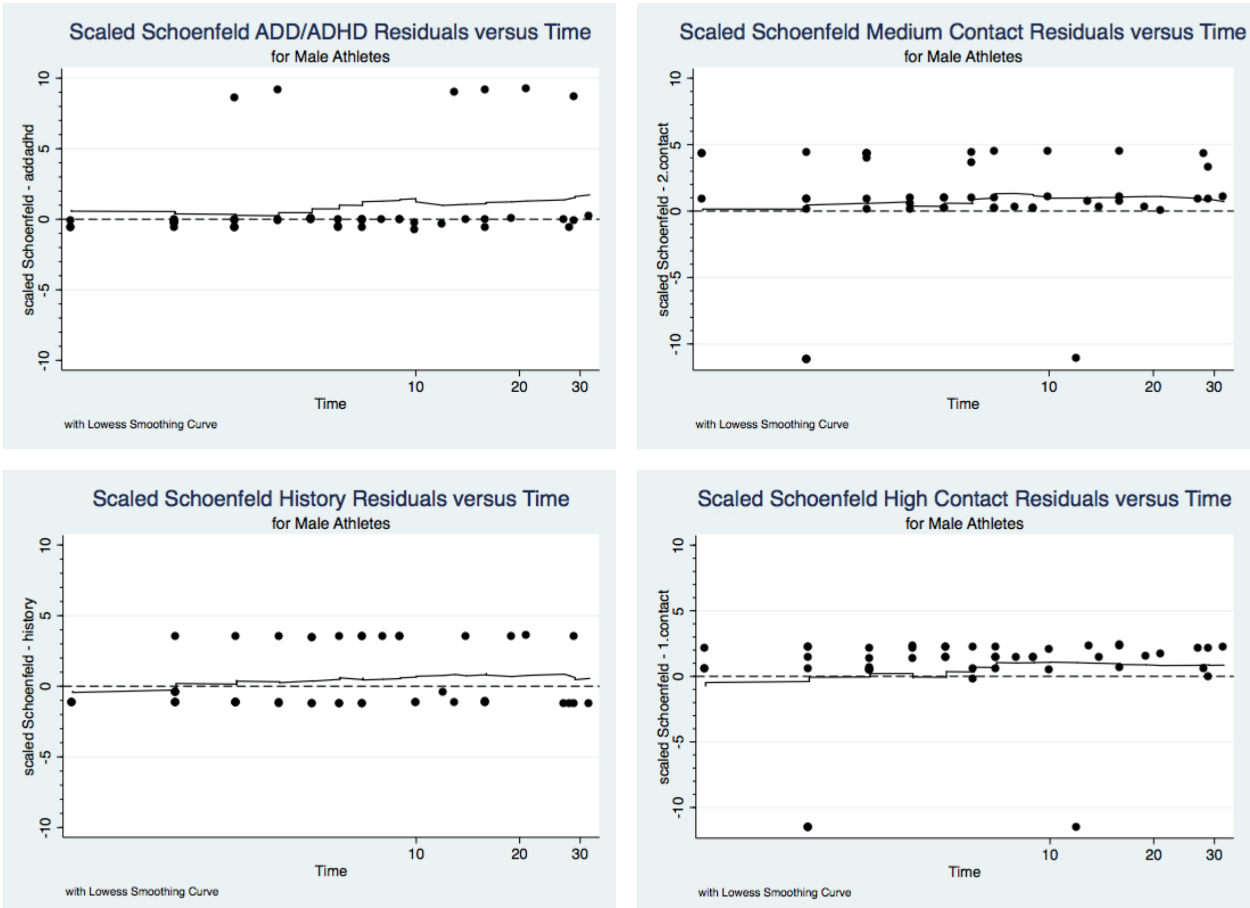

Figure 13. Schoenfeld Residuals versus Time for Selected Model of Male Athletes

For the model of female athletes, the Stata version of Therneau and Grambsch's test of non-zero slopes indicates that the age variable is the only predictor to violate the proportional hazards assumption. The log-log survival plot for the history indicator variable shows parallel lines for a prior history versus no prior history and thus indicates that the proportional hazards assumption has not been violated (Figure 14). The log-log survival plot for the contact expected variable is inconclusive (Figure 14). The plots of Schoenfeld residuals versus time for each predictor, except history, have a few irregular observations (Figure 15).

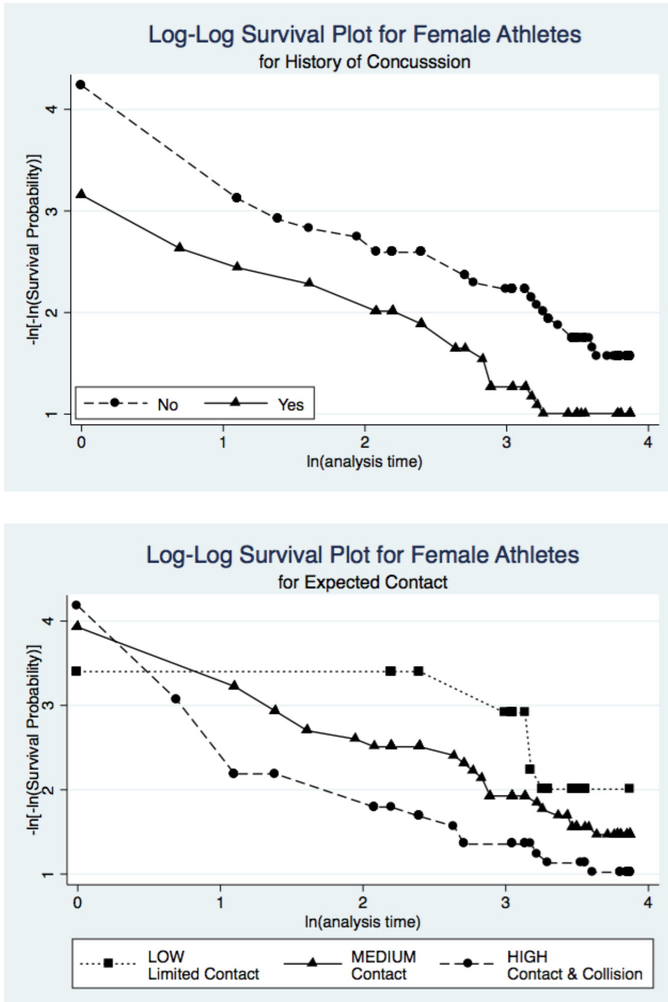

Figure 14. Log-Log Survival Plots for Selected Model of Female Athletes

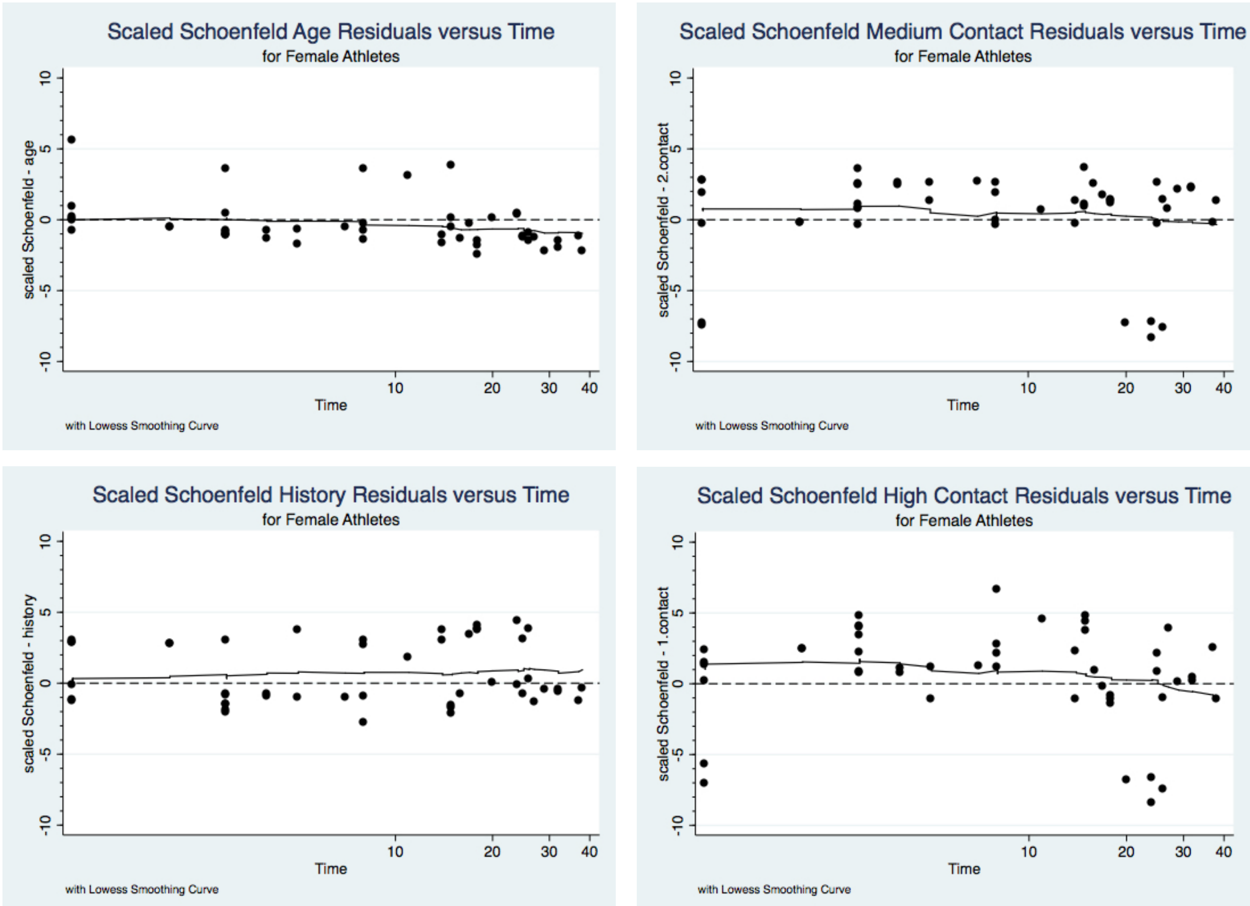

Figure 15. Schoenfeld Residuals versus Time for Selected Model of Female Athletes

Based on the Stata implemented version of Therneau and Grambsch's test of non-zero slopes, there is one predictor variable in each model that violates the assumption of proportional hazards.

3.4 MISSING DATA ANALYSES

3.4.1 Imputed Variables

Slightly less than half of the cases in the data set are complete. Because 52.3% of cases have at least one missing observation, about 50 imputations are required to yield accurate and stable standard errors. Imputation models, henceforth, are based on 50 imputations.

Of the 23 independent variables under consideration in this study, 11 had missing data with rates ranging from 0.3% to 35.5% (Figure 16). For 50 imputations, convergence was not achieved. Therefore, some of the variables with missing values were excluded from the imputation process. Variables were chosen based on the purpose of the model and univariable regular (non-imputed) Cox proportional hazards model p-values.

The history of prior concussion indicator and contact expected variables were both included in the imputation process due to the purpose of the model and findings of prior studies. These variables are to be forced into every step of model selection and thus the final model. The

ADD/ADHD indicator, height, and weight variables were included due to non-imputed univariable significance or borderline significance when adjusted for gender (Table 3).

The sleep, nationality indicator, native language indicator, dyslexia indicator, strenuous exercise indicator, and handedness variables were all considered for exclusion. Ultimately, the dyslexia indicator, sleep, and handedness variables were chosen for exclusion from the imputation process based on high non-imputed univariable p-values, complexity of regression designation for imputations, and/or expected relationship, or lack thereof, between the variable and time to concussion.

Variables	**Missing**	**Regression Designation for MICE**
Sleep	200 (25.1%)	pmm (excluded)
Height	8 (1.0%)	pmm
Weight	5 (0.6%)	pmm
Nationality	9 (1.1%)	logit
Native Language	7 (0.9%)	logit
History	8 (1.0%)	logit
ADD/ADHD	281 (35.3%)	logit
Dyslexia	283 (35.5%)	logit (excluded)
Exercise	280 (35.1%)	logit
Handedness	2 (0.3%)	mlogit (excluded)
Contact Expected	5 (0.6%)	ologit

Figure 16. Independent Variables with Missing Values

3.4.2 Univariable Cox Proportional Hazards Models for Imputed Variables

Univariable Cox models, adjusted for gender, were run for imputed variables to assess whether the imputation process affected the relationship (effect size and significance) between each imputed covariate and the time to concussion (Table 8). The univariable Cox models, adjusted for gender, for regular (non-imputed) variables are consistent with results from case-wise deletion analysis (Table 3).

Of the imputed variables, the ADD/ADHD indicator, height, and weight variables are univariably significant in case-wise deletion analysis (Table 3) for male athletes. Height is the only one of these three to be univariably significant (at the α=0.20 level) in the multiple imputation analysis. The imputed ADD/ADHD indicator and imputed weight variables are not univariably significant. Since weight is borderline significant in case-wise deletion analysis, this result is not entirely noteworthy. However, the ADD/ADHD is highly univariably significant in case-wise deletion analysis and is one of the variables in the associated selected model. This departure in significance for the imputation analysis compared to case-wise deletion analysis is a noteworthy shift.

Of the imputed variables, the weight, history of prior concussion, and expected level of contact by sport variables are univariably significant in case-wise deletion analysis (Table 3) for female athletes. The same variables are univariably significant (at the α=0.20 level) in the multiple imputation analysis (Table 8). Effect size and significance level are relatively similar between imputation analysis and case-wise deletion analysis for these three variables.

Table 8. Univariable Imputation Results by Gender

Males		Imputed Variable	Females	
Hazard Ratio	**P-value**		**Hazard Ratio**	**P-value**
1.02	0.75*	***Sleep***	1.06	0.60*
1.05	**0.12***	***Height***	1.04	0.84*
1.00	**0.18***	***Weight***	1.00	**0.18***
1.15	0.78*	***Nationality***	0.86	0.81*
0.83	0.79*	***Native Language***	1.62e-15	1.00*
1.24	0.37*	***History***	1.81	**0.04***
1.61	0.21*	***ADD/ADHD***	0.95	0.93*
3.41e-14	1.00*	***Dyslexia***	4.58e-15	1.00*
1.05	0.86*	***Exercise***	0.51	0.29*
	0.49†	***Handedness***		0.94†
		Right (Baseline)		
0.81	0.58*	*Left*	0.98	0.97*
0.46	0.28*	*Ambidextrous*	0.71	0.73*
	0.57†	***Contact Expected***		**0.13**†
		Low (Baseline)		
1.47	0.49*	*Medium*	1.54	0.33*
1.67	0.32*	*High*	2.42	0.06*

* P-values calculated using the univariable Wald test.

† Overall p-values (categorical variables with >2 levels) calculated using the likelihood ratio test.

P-values that are bold indicate a significant difference (α=0.20).

3.4.3 Multivariable Cox Proportional Hazards Models

Multivariable model selection analysis only considered variables with univariable p-value less than an α=0.20 level and any variables forced in. A full Cox model for male athletes was fit with imputed univariably significant variables, regular univariably significant variables, and known predictors of time to concussion. The variables included were height, weight, visual memory composite score, processing speed composite score, reaction time composite score, total symptom composite score, history of prior concussion, and expected level of contact by sport. The latter two variables were forced into the full model based on results from previous studies.

Model selection for each gender was conducted using backward elimination with a p-value to remove set at α=0.10. The steps of model selection for male athletes eliminated the reaction time composite score, weight, total symptom composite score, and visual memory composite score, in that order (Table 9).

Table 9. Imputation Backward Selection Results for Male Athletes

Variable	Step Removed	Hazard Ratio	P-value*
Reaction Time Composite Score	1	1.15	0.86
Weight	2	1.00	0.70
Total Symptom Composite Score	3	1.01	0.54
Memory Visual Composite Score	4	0.99	0.40

* P-values calculated using the univariable Wald test.

The selected imputation model for male athletes is comprised of the height variable, processing speed composite score, history of prior concussion indicator, and expected level of contact by sport variable (Table 10). The latter two variables were forced into the model at one or more stages of selection.

The processing speed composite score is directly related to performance in the sections of the ImPACT test that combine to produce that score. This fact along with a hazard ratio less than one indicates that processing speed performance is a protective factor in male athletes. A one-unit increase in the processing speed composite score is associated with a hazard that is 0.97 times lower. A male athlete who is one inch taller than another male athlete has a hazard that is 1.06 times higher. Equivalently, a male athlete who is six inches taller than another male athlete has a hazard that is 1.39 times higher. Male athletes with a prior history of concussion have a hazard that is 1.22 times the hazard for male athletes without any prior concussions. Males competing in medium contact expected sports (baseball, basketball, cheerleading, and diving) have a hazard that is 1.31 times the hazard for males competing in low contact expected sports (swimming). Males competing in high contact expected sports (football, soccer, and wrestling) have a hazard that is 1.51 times the hazard for males competing in low contact expected sports (swimming).

Table 10. Variables in Selected Imputation Model for Male Athletes

Variable	Hazard Ratio	P-value
Processing Speed Composite Score	0.97	**0.05***
Height	1.06	**0.09***
History	1.22	0.42*
Contact Expected		–†
Low (Baseline)		
Medium	1.31	0.63*
High	1.51	0.43*

* P-values calculated using the univariable Wald test.
† Overall p-value for multi-level categorical variables could not be calculated using the likelihood ratio test with a multivariable imputation model.
P-values that are bold indicate a significant difference (α=0.10).

A full Cox model for female athletes was fit with univariably significant imputed variables, univariably significant regular variables, and known predictors of time to concussion. The variables included were age, weight, processing speed composite score, history of prior concussion, and expected level of contact by sport.

There were only three steps of model selection for female athletes. The first step eliminated weight, the second step eliminated the processing speed composite score, and the final step eliminated the age variable from the model (Table 11).

Table 11. Imputation Backward Selection Results for Female Athletes

Variable	Step Removed	Hazard Ratio	P-value*
Weight	1	0.99	0.31
Processing Speed Composite Score	2	1.03	0.23
Age	3	0.81	0.18

* P-values calculated using the univariable Wald test.

The selected imputation model for female athletes is comprised of the history of prior concussion and expected level of contact by sport variables (Table 12). The only two variables had to be forced into the model at one or more stages of selection.

Female athletes with a prior history of concussion have a hazard that is 1.60 times higher than the hazard for female athletes without any prior concussions. Females competing in medium contact expected sports (softball, basketball, cheerleading, diving, and gymnastics) have a hazard that is 1.41 times the hazard for females competing in low contact expected sports (swimming, tennis, and volleyball). Females competing in high contact expected sports (soccer) have a hazard that is 2.06 times the hazard for females competing in low contact expected sports (swimming, tennis, and volleyball).

Table 12. Variables in Selected Imputation Model for Female Athletes

Variable	Hazard Ratio	P-value
History	1.60	0.12*
Contact Expected		–†
Low (Baseline)		
Medium	1.41	0.45*
High	2.06	0.13*

* P-values calculated using the univariable Wald test.
† Overall p-value for multi-level categorical variables could not be calculated using the likelihood ratio test with a multivariable imputation model.
P-values that are bold indicate a significant difference (α=0.10).

The female and male models selected both contain the history of prior concussion(s) indicator and the contact expected by sport variables. During selection for both of the models, both of these variables were forced into the model at one or more stages. History and the high contact expected level are borderline significant for females and highly insignificant for males. The effect sizes are also larger in the female model compared to those in the male model. History of prior concussion versus no concussion has a 1.22 times higher hazard for males and a 1.60 times higher hazard for females. For males, the hazard is 1.31 times higher in medium versus low contact sports and 1.51 times higher in high versus low contact sports. For female athletes, the hazard is 1.41 times higher in medium versus low contact sports and 2.06 times higher in high versus low contact sports.

3.4.4 Checking Model Assumptions

The first few assumptions are empirically assessable and independent of analysis technique (case-wise deletion versus multiple imputation). The time to event outcome, non-informative censoring, and independence of observations assumptions were assessed subsequent to the selection of case-wise deletion models (see 3.3.4).

The proportional hazards assumption is dependent on analysis technique and the selected models. It, therefore, must be assessed subsequent to fitting the model. The Cox model assumes a constant relative hazard between groups.

For the multiply imputed models, the proportional hazards assumption for continuous and binary variables was assessed through the significance of an interaction term between time and the predictor variable. For the model of male athletes, none of the three interaction terms between time and visual motor composite score, history indicator, and height each are significant at the $\alpha=0.05$ level. For the model of female athletes, the interaction term between time and history indicator variable is not significant at the $\alpha=0.05$ level. The proportional hazards assumption for the multi-level categorical contact expected variable was informally assessed by extracting a few of the imputed data sets and examining the Stata version of Therneau and Grambsch's test of non-zero slopes. Neither the male nor female athlete models had violations in the proportional hazards assumption for the contact expected variable in any of the imputed data sets arbitrarily selected for extraction.

An additional assumption for multiple imputation analysis is a correct specification of the imputation process. The chained equations approach was used to account for the complexity of missingness in the data set and to handle imputing continuous and categorical variables iteratively. The nationality, native language, history, ADD/ADHD, and exercise indicator variables were all imputed under the logit (logistic) method to account for the underlying binomial distribution of each. The handedness categorical variable was imputed under the mlogit (multinomial logistic) method to account for its nominal nature. The contact expected categorical variable was imputed under the ologit (ordinal logistic) method to account for ordinal nature. The height and weight continuous variables were imputed under the pmm (predictive mean matching) method to account for their non-normal distributions (Figure 16).

3.5 CASE-WISE DELETION VERSUS MULTIPLE IMPUTATION

The selected imputation model for male athletes and the selected case-wise deletion model for male athletes only have the history indicator and contact expected variables in common, both of which were forced into each model during selection. In the case-wise deletion model, the effect sizes for the contact expected variable are in contrast to what may be expected (Table 13). The effect size (hazard) of medium versus low contact sports (1.75) is actually higher than the effect size of high versus low contact sports (1.35). In the imputation model, males competing in medium contact expected sports have a hazard that is 1.31 times the hazard for males competing in low contact expected sports. Males competing in high contact expected sports have a hazard that is 1.51 times the hazard for males competing in low contact expected sports. The hazard ratios are reversed in the two models.

The two models are somewhat different with respect to variables that are included but were not forced in. The case-wise deletion model only includes the ADD/ADHD indicator variable additionally. This variable is highly significant for the model but it also has a relatively high rate of missing data and therefore limits the sample size for the model. The imputation model excluded this indicator at the univariable level prior to fitting a full model. Instead, the processing speed composite score and height variables are included.

Table 13. Variables in Selected Model for Male Athletes by Analysis Strategy

Variable	***Case-Wise Deletion***		***Multiple Imputation***	
	Hazard Ratio	**P-value**	**Hazard Ratio**	**P-value**
ADD/ADHD	2.52	**0.04***	–	–
Processing Speed Composite Score	–	–	0.97	**0.05***
Height	–	–	1.06	**0.09***
History	1.41	0.28*	1.22	0.42*
Contact Expected		0.57†		–†
Low (Baseline)				
Medium	1.75	0.33*	1.31	0.63*
High	1.35	0.57*	1.51	0.43*

* P-values calculated using the univariable Wald test.
† Overall p-values (categorical variables with >2 levels) calculated using the likelihood ratio test (could not be calculated for a multivariable imputation model).
P-values that are bold indicate a significant difference (α=0.10).

The selected imputation model for female athletes and the selected case-wise deletion model for female athletes only have the history indicator and contact expected variables in common (Table 14). History is highly significant in the case-wise deletion model but had to be forced into the multiple imputation model. The effect size of history is slightly greater in the case-wise deletion model while the effect size of contact expected is marginally greater in the multiple imputation model. The only other variable present is age in the case-wise deletion model. It is a highly significant protective variable since younger females have a higher hazard than older female athletes.

Table 14. Variables in Selected Model for Female Athletes by Analysis Strategy

Variable	***Case-Wise Deletion***		***Multiple Imputation***	
	Hazard Ratio	**P-value***	**Hazard Ratio**	**P-value***
Age	0.68	**0.04***	–	–
History	1.92	**0.03***	1.60	0.12*
Contact Expected		0.49†		–†
Low (Baseline)				
Medium	1.38	0.49*	1.41	0.45*
High	1.78	0.25*	2.06	0.13*

* P-values calculated using the univariable Wald test.
† Overall p-values (categorical variables with >2 levels) calculated using the likelihood ratio test (could not be calculated for a multivariable imputation model).
P-values that are bold indicate a significant difference (α=0.10).

4.0 DISCUSSION

4.1 ANALYSIS OF RESULTS

4.1.1 Exclusion of Cohort Variable

Although the academic year of the baseline test variable was univariably significant for female athletes, it was excluded from the consideration because this variable, as it is formatted, is inconsistent with the purpose of the study: to create a model that may be used to identify groups of incoming student-athletes at increased risk of concussion. After consultation with a Pitt athletic trainer, it was determined that there was no change in the way concussion was handled by the athletic department with respect to this particular dataset. Incoming athletes would be among a new level of the cohort variable that characterizes the academic year of the baseline test. Because there is no preexisting information for the cohort of athletes which the model is meant for, a meaningful relationship between cohort and time to concussion cannot be discerned.

4.1.2 Case-Wise Deletion versus Multiple Imputation

Implications of case-wise deletion analysis are dependent on the cause of the missing values. When data are MCAR, missing values are independent and unpredictable. Neither the observed variable values in the data set nor the unobserved values of the variable itself are able to predict whether a value will be missing. When data are MAR, the missing values may be predicted by other variables in the data set; i.e., observed data values can predict whether or not a variable will be missing.

The three variables with the highest rates of missing data were the ADD/ADHD indicator, dyslexia indicator, and strenuous exercise indicator variables. None of these variables had an observed value when the cohort entry date preceded 2011. Although some variables with missing values may be MCAR, these three variables should certainly be considered MAR.

In case-wise deletion analysis, any case with a missing value is excluded from analysis and thus all of the existing data associated with that case is lost. Case-wise deletion analysis in the presence of MCAR data will produce unbiased parameter estimates but may drastically reduce sample size and thus produce large standard errors. For MAR data, the complete cases may not be a random sample and thus case-wise deletion analysis may produce biased parameter estimates.

These distinctions are evident in the results of this study. In case-wise deletion analysis, the model for male athletes was driven by the ADD/ADHD indicator variable. As long as it remained in the model, all data from the years 2007-2010 was lost and thus the sample size was reduced by approximately 35%. The final model consisted of this ADD/ADHD variable along with the history of prior concussion and contact expected by sport variables. The effect sizes for the contact expected variable are in contrast to what might have been predicted: 1.75 times higher hazard in medium contact versus low contact sports and 1.35 times higher hazard in high contact versus low contact sports. The overall person-time incidence rates observed in this data set are 5.02 concussions per 100 person-years in low contact sports, 6.99 concussions per 100 person-years in medium contact sports, and 8.19 concussions per 100 person-years in high contact sports. However, these incidence rates are very different when the ADD/ADHD variable is not missing. They are 5.12, 11.39, and 8.74 concussions per 100 person-years in low, medium, and high contact sports, respectively. Clearly male athletes with missing values for the ADD/ADHD variable do not represent a random sample of male athletes from the population in the data set.

Because the data are more likely to be MAR than MCAR, case-wise deletion models pose the risk of being subject to biased estimates. For this data set, multiple imputation models should be used.

4.1.3 Multiple Imputation Models

In the multiple imputation model for male athletes, the processing speed composite score is a protective factor. A one-unit increase in the processing speed composite score is associated with a hazard that is 0.97 times lower. Males with relatively low values for this composite score could be targeted for prevention methods. The person-time incidence rate for male athletes with scores greater than or equal to 39 (the mean in the study population) is 5.24 concussions per 100 person-years while it is 9.08 concussions per 100 person-years for male athletes with composite score less than the average.

A male athlete who is six inches taller than another male athlete has a hazard that is 1.39 times higher. The person-time incidence rate for male athletes with height greater than or equal to six feet (72 inches) (the mean in the study population) is 7.90 concussions per 100 person-years while it is 6.38 concussions per 100 person-years for male athletes shorter than the mean height.

Male athletes with a prior history of concussion have a hazard that is 1.22 times the hazard for male athletes without any prior concussions. According to this model, any male athletes with a prior history of concussion should be targeted for prevention methods.

Males competing in high expected contact sports (football, soccer, and wrestling) have a hazard that is 1.51 times the hazard for males competing in low expected contact sports (swimming) and 1.15 times the hazard for males competing in medium expected contact sports (baseball, basketball, cheerleading, and diving). Football players, wrestlers, and men’s soccer players could all be targeted for prevention methods.

According to the multiple imputation model for female athletes, those with a prior history of concussion have a hazard that is 1.60 times higher than the hazard for female athletes without any prior concussions. Thus, any female athletes with a prior history of concussion should be targeted for prevention methods.

Females competing in high expected contact sports (soccer) have a hazard that is 2.06 times the hazard for females competing in low expected contact sports (swimming, tennis, and volleyball) and 1.46 times the hazard for females competing in medium expected contact sports (softball, basketball, cheerleading, diving, and gymnastics). Women's soccer players could be targeted for prevention methods.

4.1.4 Compare & Contrast to Other Studies

Although other studies have found prior history of concussion and type of sport to be significant predictors of risk for concussion, multiple imputation analysis of this data set does not support these findings for either gender.

This discrepancy with prior studies may be attributed to the sports under consideration. There are a few gender-specific sports with very different person-time incidence rates of concussion. Football and wrestling have only male athletes competing at Pitt with incidence rates of 8.41 and 8.73 concussions per 100 person-years, respectively. Gymnastics, tennis, and volleyball have only female athletes competing at Pitt and incidence rates of 7.08, 0, and 0 concussions per 100 person-years, respectively. For this reason, a comparison of survival based on sport must consider gender.

Type of sport was not analyzed as an independent variable; rather, it was categorized based on expected level of impact. This redefinition of sport may account for the discrepancy. An alternative categorization of sports might yield such categorization as a significant predictor. Perhaps only a small number sports are significant predictors while all other sports should all be grouped together in a more meaningful way.

Also, it seems likely that history of prior concussion and sport may be associated. If incidence rate is higher in a particular sport, it is also more likely that athletes of that sport will have a prior history of concussion. The presence of both variables in the model may detract from each other relative to significance levels.

In the unadjusted case-wise deletion Cox proportional hazards models (Table 2), history and contact expected (pseudo sport variable) were significant, however gender was not. While many of the sports have both male and female athletes, there are a few gender-specific sports with very different person-time incidence rates of concussion. If univariable analysis is restricted to each sport independent of the others, gender is significant only in soccer.

While this study does not find history of concussion, expected level of contact by sport, or gender to be significant in a multivariable multiple imputation model, each is univariably significant, whether overall or within some sports.

This study presented an analysis of some variables that have not been used in predicting incidence of concussion. Since the purpose of the ImPACT test protocol is to facilitate “return to play” decisions, the composite scores obtained from each test are used in the same manner, i.e., with respect to analysis post-concussion. This study analyzes the utility of such variables in a novel way, prior to incidence. The selected multiple imputation model for male athletes found the processing speed composite score to be a significant protective factor with regards to time to concussion. Four of the six composite scores (visual memory, processing speed, reaction time, and total symptom) were univariably significant (α=0.20) for male athletes (Table 3). Processing speed was univariably significant (α=0.20) for female athletes (Table 3).

4.2 LIMITATIONS OF THE STUDY

4.2.1 Generalization of Results

The results of this study are valid for the population represented by the data set upon which the results were derived. To generalize the results to other populations of collegiate athletes is only as valid as the methods of data collection used to produce the data set. It was assumed that every incoming athlete took a baseline ImPACT test as of the beginning of the 2007-08 academic year; however, many sports were not represented in the earliest cohorts of data. It is possible that student-athletes were baseline tested in subsequent years. Neither representation nor degree of representation can definitively indicate whether this assumption of data collection was upheld. The number of athletes per class can fluctuate immensely by sport. The data set itself neither validates nor invalidates the methods of data collection. It can, however, call into question such methods.

The results may not be generalized to any collegiate sports not represented in the study. For instance, track & field was excluded because there were only three male and three female track & field athletes represented in the data set over a seven-year study period. It was determined that this sport either did not follow the baseline ImPACT test protocol or were not included in the protocol intentionally. Either way, data from these six athletes would not be representative of the track & field program in its entirety. Thus, the results are not appropriate for track & field athletes, or for any athletes who participate in a collegiate sport not analyzed in this study.

This data set is presumed to represent the target population of Pitt athletes who entered the Athletic Department between July 2007 and April 2014 if the ImPACT test procedures were implemented in an all-inclusive manner, as intended. If the data set includes all Pitt Division I college athletes during this period, then it could represent a sample population for the target population of incoming Pitt student-athletes. On a larger scale, the data set could represent a sample population for the target population of all incoming Division I college athletes. The purpose of statistical analysis is to generalize results from a sample population to a target population. The ability to extrapolate the results of this study to either target population depends on the external validity of the study, which in turn depends on the make-up of the study population.

The integrity of the data set cannot be ascertained from the data set itself or from consultation with a member of the Pitt athletic training staff. Based on an examination of the sports represented for each academic year, it appears that the Pitt Athletic Department's ImPACT test protocol (outlined in 2.1) may have been implemented in a stepwise fashion over a few years. This conclusion limits the ability to generalize specific results of the study to larger populations.

4.2.2 Redshirt & Transfer Student-Athletes

The data set provided no indication of athletes who “redshirted” a year or transferred into or out of their program. A “redshirt” year may be granted to incoming student-athletes to extend a player’s four years of eligibility. During a “redshirt” year, the student-athlete may take courses at the university and participate with the collegiate team during practices but may not compete in any games or contests. A medical “redshirt” may be granted to a student who has suffered a debilitating injury during the course of a season to extend a player’s eligibility (“NCAA Eligibility Requirements”). To negate the effects of “redshirt” athletes, the study performs a four-year survival analysis. Since incidence rates of concussion are shown to be higher in games compared to practices (Daneshvar, “The Epidemiology of Sport-Related Concussion”), identifying athletes who “redshirt” could marginally affect the model.

By analyzing a four-year survival from the point of the baseline test, survival times are over-estimated when transfers are included in the study but not controlled for. Survival times will only be over-estimated for transfer students who do not sustain a concussion during the study period. In the 2008-09 academic year, the transfer rate for all sports was 6.5% (Alexander, “College Athletes”). If the incidence rate of concussion is not significantly different for transfer versus non-transfer students and assuming the self-reported rate of concussion is 33.5% (Llewellyn TL, “Concussion Reporting Rates”), then approximately 4.3% of the collegiate athlete population will have over-estimated survival times; i.e., 4.3% of the population will be transfers and not sustain a concussion. More precise entrance and exit times to the Pitt Athletic Department should be documented for more accurate modeling.

4.2.3 Multiple Comparisons

The finding that the survival curves for male and female soccer players differ may be due simply to chance because of multiple comparisons. The further examination of gender-stratified survival curves by expected contact level and history of prior concussion preserve the probability of a type I error at 0.05 (α=0.05). Stratified log-rank tests and pair-wise stratified log-rank tests subsequent to the simple log-rank tests were used to determine where the difference was.

4.3 FUTURE WORK

This study is a preliminary analysis for several inter-related tracks. First and foremost, it would be advantageous for future studies to be prospective in nature to ensure a representative and valid data set. Although not necessary, it would also be advantageous for future studies to analyze larger data sets.

A natural extension, particularly with a larger data set, is to perform similar analyses in this study but by sport, not gender. Sport-specific models would have the benefit of precision when assessing the association between incidence of concussion and potential predictor variables. Because the definition of concussion is particularly general, an increased focus on the exact cause of concussion may be fundamental to identifying groups of athletes at increased risk. For instance, in soccer, the primary cause of bumps, blows, or jolts to the head are collisions between players when heading the ball. There is a discrepancy in the rate of collisions expected by position. By restricting analysis to individual sports, the unique properties of each sport may be used to define a model of incidence of concussion.

Another track is to perform survival analysis on time to recovery. In addition to a representative and valid data set of ImPACT test scores, this analysis would necessitate determining how to represent when recovery is actually achieved. In the middle of this study, the Pitt protocol was amended to require a reevaluation of the baseline test upon recovery. The date of this baseline reevaluation could potentially be used to represent recovery. Test scores may require manipulation to create variables for the change in test scores for each subsequent post-injury test.

Due diligence and care must be taken in procuring a representative data set. This should be accomplished by complying with Pitt's ImPACT test protocol: persistent baseline testing of incoming athletes and post-injury testing as proximate to injury as possible. Additionally, data should be aggregated across populations with comparable protocols to increase sample sizes, particularly for sports with naturally fewer athletes. A larger sample would enable sport-specific analysis, which would potentially further the specificity of identifying at risk individuals and thus the specificity with which prevention methods could target at-risk individuals.

5.0 CONCLUSIONS

The purpose of this study was to increase the specificity with which prevention methods could target college athletes who are at increased risk for concussion. Two models were created, one for male and one for female athletes, which identify risk and protective factors associated with concussion for the given population.

Due to a large percentage of incomplete cases and data that are MAR, multiple imputation analysis was preferred to case-wise deletion analysis. Because the data set was comprised of multiple discrete variables, categorical and binary, MICE analysis was chosen to provide flexibility in the imputation process. With MICE, each imputed variable was given its own regression designation based on the distribution of the variable itself. For example, binary variables were designated with logistic regression to account for the underlying binomial distribution. Independent designation of regression methods for each imputed variable provides the flexibility for imputation when compared to the MVN method, which has a stricter joint distribution assumption.

Analysis of the multiply imputed data utilized nearly all of the observed data. By imputing the data multiple times and aggregating the data sets to generate overall parameter and standard error estimates, we were able to use nearly all of the observed data while accounting for the uncertainty of imputation.

Model selection was performed based on backward elimination with a p-value to remove set at $\alpha=0.10$ and forcing in known predictors of concussion.

With respect to female athletes, the analysis did not yield any novel risk or protective factors associated with concussion. With respect to male athletes, the analysis showed that the processing speed composite score is a protective factor associated with concussion and height is a risk factor associated with concussion. The assumptions of the survival analysis and multiple imputation were met and thus the results are internally reliable and valid.

The significance of ImPACT test composite scores in univariable associations between each score and time to concussion is noteworthy and novel. This study provides a preliminary analysis of the utility of ImPACT in predicting incidence of concussion.

The two models produced, one for male athletes and one for female athletes, identify athletes at increased risk for concussion in the study population. It examined known risk factors (history, gender, and sport) and identified novel risk factors (namely, the processing speed ImPACT composite score) that are significantly associated with incidence of concussion in this study population. Although the ability to generalize these findings is indiscernible with any degree of certainty, the study posits that the ImPACT test is useful in identifying college athletes at increased risk for concussion in this population. This finding necessitates subsequent studies and analysis to corroborate these findings and to do so with data collection methods that ensure the ability to generalize results.

APPENDIX A

INDEPENDENT VARIABLES

All independent variables were derived either directly or indirectly from the ImPACT tests. This section provides a description of how variables were indirectly derived, how unusual values were handled, and descriptions of the included and excluded variables in the data analysis.

A.1 INCLUDED

This section provides a description of variables that were included in data analysis. It also describes how each variable was derived, either directly or indirectly, from the ImPACT test.

A.1.1 Sleep

The sleep variable is a self-reported field for the number of hours of sleep the test taker had the night prior to the test being taken.

A.1.2 Age

The age variable was derived from the date of birth and test date variables on the ImPACT test. This indirectly derived variable quantifies the exact (without rounding) age of the athlete at the time each test was taken. In the data set used for analysis, the age variable quantifies that age at which the athlete entered the study (at the time of the baseline test).

The study population of interest was specified to be college-aged athletes. It was unclear whether extreme ages were simply logged erroneously or if associated tests belonged to individuals outside the study population of interest. Some variation is expected, however anyone with tests at an age younger than 17 or older than 27 were excluded from the data set (n=4).

A.1.3 Height

The height variable came directly from the ImPACT test, however a few of the values were converted from metric to US standard (imperial) measurements (specifically inches) for the sake of consistency and comparison.

A.1.4 Weight

The weight variable came directly from the ImPACT tests. A few of the values were converted from metric to US standard (imperial) measurements (specifically pounds) for the sake of consistency and comparison.

A.1.5 Composite Scores

ImPACT derives five composite scores from the neurocognitive portion of the test, section 3. The values of the verbal memory, visual memory, processing speed, reaction time, and impulse control composite scores were acquired directly from the test. Each of the five composite scores is a separate variable in the analysis. The verbal memory composite, visual memory composite, and visual processing speed composite scores are directly related to performance in these categories. The reaction time composite and impulse control composite scores are inversely related to performance ("Overview" *ImPACT Test*).

A.1.6 Total Symptom Score

The total symptom score variable and the symptom score variable have equivalent values for every test in the data set. Total symptom score was chosen for analysis in lieu of symptom score. The value for this variable was taken directly from the ImPACT test.

A.1.7 Female (Gender)

The female indicator variable was derived from the gender variable acquired from ImPACT. It is a dichotomous variable accounting for males and females.

A.1.8 Nationality

The nationality indicator, or dummy variable, is a variable derived indirectly from the country variable, which was acquired directly from ImPACT. It is a dichotomized variable accounting for American student-athletes and foreign student-athletes.

A.1.9 Native Language

The native language indicator, or dummy variable, is a variable derived indirectly from the language variable, which is acquired directly from ImPACT. It is a dichotomized variable accounting for native English speakers and non-native speakers.

A.1.10 History

The history of prior concussion indicator variable was derived indirectly from the number variable, which lists the number of concussions the associated athlete had sustained prior to the test. Student-athletes without any prior concussions (number=0) were concluded to have no prior history. Student-athletes with one or more prior concussions (number>0) were concluded to have a prior history of concussion.

A.1.11 ADD/ADHD

The ADD/ADHD indicator, or dummy variable, is a dichotomized variable taken directly from ImPACT. The variable represents a self-reported diagnosis of ADD or ADHD.

A.1.12 Dyslexia

The dyslexia indicator, or dummy variable, is a dichotomized variable taken directly from ImPACT. The variable represents a self-reported diagnosis of dyslexia.

A.1.13 Special Education

The special education indicator variable was derived from the special education 1 through 5 indicator variables. It was used to create the symptom classification variable. It distinguishes between athletes who self-reported a special education designation and those who did not.

A.1.14 Exercise

The exercise indicator, or dummy variable, is a dichotomized variable taken directly from ImPACT. It distinguishes between athletes who had no strenuous exercise immediately prior to taking the test and athletes who did exercise strenuously.

A.1.15 Handedness

Handedness is a tri-level categorical variable with levels for right-handers, left-handers, those who are ambidextrous. The right-handed group was used as a baseline for comparison.

A.1.16 Administrator of Test

The administrator variable is a multi-level categorical variable acquired directly from the ImPACT test. It distinguishes between which Pitt athletic training staff member or UPMC neurocognitive expert administered the test. To maintain anonymity of these health care professionals, this variable is presented numerically.

A.1.17 Contact Expected

Due to a limited sample size overall, and particularly within some of the sports, the 12 represented sports were categorized according to the expected level of contact. The high expected contact group is defined as collision and contact sports and includes football, soccer, and wrestling. The medium expected contact group is defined as contact sports and includes baseball, softball, basketball, cheerleading, diving, and gymnastics. The low expected contact groups is defined as limited contact sports and includes swimming, tennis, and volleyball (Klossner, *2013-14 NCAA Sports Medicine Handbook*, 56-58). The low expected contact group was used as a baseline for comparisons.

A.1.18 Academic Year of Baseline Test

This cohort variable categorizes student-athletes based on the academic year of the baseline test date. Each academic year was defined as July of one year through June of the following year. The first cohort, the 2007-08 academic year, was used as a baseline for comparisons.

A.1.19 Symptom Classification

The symptom classification variable was derived based on the guidelines of ImPACT's post-concussion symptom scale (*Immediate Post-Concussion Assessment Testing (ImPACT®) Test: Technical Manual*). The total symptom score variable, female indicator, ADD/ADHD indicator, dyslexia indicator, autism indicator, and special education variables were used to define the levels in the symptom classification variable (Table 15).

Table 15. Post-Concussion Symptom Scale for University Men & Women

Classification	***Raw Scores***			
	Regular Education Men	*Regular Education Women*	*Special Education* Men*	*Special Education* Women*
Low-Normal	0	0	0	0
Normal	1 – 5	1 – 10	1 – 12	1 – 13
Unusual	6 – 12	11 – 21	13 – 28	14 – 21
High	13 – 20	22 – 31	29 – 41	22 – 31
Very High	21+	32+	42+	32+

* All subjects who self-reported any past speech therapy, learning problems, ADHD, or special education placement were included in the Special Education groups.

A.2 TRANSFORMED VARIABLES

This section provides a description of variables that were used to transform the data into variables that were included in data analysis.

A.2.1 Test Type

The test type variable makes a distinction between valid baseline, invalid baseline, and post-injury tests. It was used to determine the outcome of interest in the study: concussion or no concussion. It was also used to determine the cohort entry date and the survival time.

A.2.2 Sport

The sport variable identifies the sport that each athlete entered the Pitt Athletic Department to compete in. It was used to create the level of contact expected variable.

Football, varsity football, and freshman football were all merged under the umbrella category of football. There was one athlete with football designated for sport and female designated for gender. The country of this athlete was Finland. It was deduced that the athlete is a female from Finland who refers to her sport as football but in the US it is commonly known as soccer.

Boxing (n=2), bull riding (n=1), lacrosse (n=3), rugby (n=1), skiing (n=2), and snowboarding (n=1) are not Division I sports at Pitt and thus, any corresponding athletes were excluded from the data set. Track & field only had three male athletes and three female athletes represented in the data set. Observably, track & field athletes were not being baseline tested in accordance to the ImPACT test protocol. All track & field athletes (n=6) were excluded from the data set because the six athletes with baseline tests over the seven-year study period are unlikely to be a random and representative sample.

A.2.3 Country

The country variable identifies the country of origin for each athlete. It was used to dichotomize this information and create the nationality indicator.

A.2.4 Language

The language variable identifies the native language for each athlete. It was used to dichotomize this information and create the native language indicator.

A.2.5 Date of Birth

The date of birth variable was used to identify any athletes who are mistakenly linked together by the third party who de-identified the data. It was also used to create the age variable.

A.2.6 Number

The number of concussions variable comes directly from ImPACT. It quantifies the number of concussions an athlete had sustained prior to entrance to the Pitt Athletic Department. It was used to create the history of prior concussion indicator variable.

A.2.7 Special Education 1, …, Special Education 5

The special education 1 through 5 indicator variables were acquired directly from the ImPACT test. They were used to create the special education indicator which was in turn used to create the symptom classification variable. It distinguishes between athletes who self-reported special education designations and those who did not.

A.2.8 Test Date

The test date variable was used to quantify survival time through the cohort entry date and subsequent event date. It was also used to create the cohort and age variables.

A.3 EXCLUDED

This section provides a brief description of variables that were excluded from data analysis. Reasons for exclusion vary from irrelevancy to the purpose of the study, inaccuracy of responses, complexity or responses, and constancy throughout the study population.

A.3.1 Irrelevant to the Purpose of the Study

Neither the administrator email nor the exam language nor the test version variables are expected to be relevant in quantifying the risk of concussion. There were over 50 variables representing ImPACT test scores that combined to form the composite scores. These scores may be relevant in future studies should one or more of the composite scores prove to be a predictor of incidence rate of concussion. For this preliminary study, though, these scores were too specific.

The missed games and position variables were excluded from analysis for this study but would be appropriate in sport-specific analysis. In its original form from the ImPACT test, the missed variable quantified the number of games the athlete had missed. The variable was transformed into a quad-level categorical variable indicating 0 games missed, 1-9 games missed, 10-19 games missed, and 20 or more games missed. Because the rate of games played varies with the type of sport, this variable would be more appropriate when sport is designated, e.g., sport-specific analysis. The position variable would need to be grouped together in a meaningful way for each sport. Because position is not consistent across all sports, this variable would be more appropriate in sport-specific models.

A.3.2 Suspected Inaccurate Responses

The race, education, years, last concussion date, and type 1 through type 4 variables were all excluded due to a suspected high rate of inaccurate responses. Some variables had values inconsistent with the variable field and some had a span of values far greater than expected indicating that the intended response for this field was not clear.

A.3.3 High Complexity of Responses

The medications and history date variables were not submitted into ImPACT in a manner that would be conducive for analysis. There were no measures in place to ensure reporting this fields in a consistent way. It appears the test taker simply typed responses into these fields. For instance, test takers typed the name of any current medications rather than simply selecting a category like birth control or anti-depressants.

A.3.4 Constant Variables

The organization, level of participation, and autism variables were constant throughout the entire data set. The organization was the University of Pittsburgh and the level of all test-takers kept in the data set was college.

Any test-takers with a semi-professional or professional designation were examined and subsequently dropped. All test-takers with a high school designated level were examined and reassigned to a college level. This was done because the high school level tests were presumed to be incoming freshmen who had not yet participated at a college level.

APPENDIX B

INSTITUTIONAL REVIEW BOARD (IRB) APPROVAL

University of Pittsburgh
Institutional Review Board

3500 Fifth Avenue
Pittsburgh, PA 15213
(412) 383-1480
(412) 383-1508 (fax)
http://www.irb.pitt.edu

Memorandum

To: Laura Berbert

From: IRB Office

Date: 12/4/2014

IRB#: PRO14110155

Subject: Analysis of Utility of ImPACT Tests in Predicting Time to Concussion in University of Pittsburgh Athletes

---

The above-referenced protocol has been reviewed by the University of Pittsburgh Institutional Review Board. Based on the information provided to the IRB, this project includes no involvement of human subjects, according to the federal regulations [§45 CFR 46.102(f)]. That is, the investigator conducting research will not obtain information about research subjects via an interaction with them, nor will the investigator obtain identifiable private information. Should that situation change, the investigator must notify the IRB immediately.

Given this determination, you may now begin your project.

Please note the following information:

- If any modifications are made to this project, use the "**Send Comments to IRB Staff**" process from the project workspace to request a review to ensure it continues to meet the determination.
- Upon completion of your project, be sure to finalize the project by submitting a "**Study Completed**" report from the project workspace.

Please be advised that your research study may be audited periodically by the University of Pittsburgh Research Conduct and Compliance Office.

BIBLIOGRAPHY

"About ImPACT: Concussion Signs and Symptoms." *ImPACT*. ImPACT Applications, n.d. Web. 11 Nov. 2014.

"About ImPACT: ImPACT Founders." *ImPACT*. ImPACT Applications, n.d. Web. 11 Nov. 2014.

"About ImPACT: Overview." *ImPACT*. ImPACT Applications, n.d. Web. 11 Nov. 2014.

"About ImPACT: The ImPACT Test" *ImPACT*. ImPACT Applications, n.d. Web. 13 Nov. 2014.

Alexander, Elton. "College Athletes Turn to Transfers in Search of Playing Time." *Cleveland.com*. Plain Dealer, 04 Mar. 2011. Web. 11 Nov. 2014.

Allison, Paul. "Why You Probably Need More Imputations Than You Think." *Statistical Horizons*. N.p., 09 Nov. 2012. Web. 11 Nov. 2014.

Associated Press. "Concussion Study Says NCAA Needs Improvement." *KXAN*. LIN Television of Texas, LP, 23 Oct. 2014. Web. 11 Nov. 2014.

Brant, Rollin. "Assumptions of the Cox Model." N.p., 24 Mar. 2004. Web. 11 Nov. 2014.

Daneshvar, Daniel H., Christopher J. Nowinskii, Ann C. McKee, and Robert C. Cantu. "The Epidemiology of Sport-Related Concussion." *Clinics in Sports Medicine* 30.1 (2011): 1-17. *Clinics in Sports Medicine*. Elsevier. Web. 11 Nov. 2014.

"Deciding to Impute." *Multiple Imputation in Stata*. UW Board of Regents, University of Wisconsin - Madison, 13 Jan. 2013. Web. 11 Nov. 2014.

Facts about Concussion and Brain Injury: Where to Get Help. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2010. PDF.

Fox, John. *Cox Proportional-Hazards Regression for Survival Data*. N.p.: n.p., Feb. 2002. PDF.

Heads Up: Concussion in High School Sports, Guide for Coaches. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Feb. 2005. PDF.

Heads Up: Facts for Physicians About Mild Traumatic Brain Injury (MTBI). Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2013. PDF.

Howell, David C. "Treatment of Missing Data--Part 1." *Treatment of Missing Data*. N.p., 09 Dec. 2012. Web. 12 Nov. 2014.

"IHRSA - About the Industry." *IHRSA - About the Industry*. International Health, Racquet & Sportsclub Association, 17 June 2013. Web. 11 Nov. 2014.

Immediate Post-Concussion Assessment Testing (ImPACT®) Test: Technical Manual. N.p.: ImPACT Applications, 2011. PDF.

"Injury Prevention & Control: Traumatic Brain Injury." *Centers for Disease Control and Prevention*. Centers for Disease Control and Prevention, 22 July 2013. Web. 09 Nov. 2014.

Kelley, Bruce, and Carl Carchia. ""Hey, Data Data -- swing!"" *ESPN*. ESPN Internet Ventures, 11 July 2013. Web. 12 Nov. 2014.

Klossner, David. *2013-14 NCAA® Sports Medicine Handbook*. Indianapolis, IN: National Collegiate Athletic Association, Aug. 2013. PDF.

Kropko, Jonathan, Ben Goodrich, Andrew Gelman, and Jennifer HIll. *Multiple Imputation for Continuous and Categorical Data: Comparing Joint and Conditional Approaches*. N.p.: n.p., 4 Oct. 2013. PDF.

Kuczynski-Brown, Alex. "Physical Education Programs Stalled By State Loopholes: 2012 Shape Of The Nation Report." *The Huffington Post*. TheHuffingtonPost.com, 14 Nov. 2012. Web. 12 Nov. 2014.

Llewellyn TL, Burdette GT, Joyner AB, and Buckley TA. "Concussion Reporting Rates at the Conclusion of an Intercollegiate Athletic Career." *Clin J Sport Med* (2014): n. pag. *PubMed*. Web. 10 Nov. 2014.

Moser, Rosemarie S. "Baseline Neuropsychological Tests: Getting Valid Results Poses Challenge." *MomsTeam*. N.p., 13 Sept. 2013. Web. 12 Nov. 2014.

Multiple Imputation FAQ Page. N.p., n.d. Web. 11 Nov. 2014.

"Multiple Imputation in Stata, Part 1." *UCLA: Statistical Consulting Group*. UC Regents, n.d. Web. 11 Nov. 2014.

"NCAA Eligibility Requirements." *FindLaw Sports*. FindLaw, n.d. Web. 11 Nov. 2014.

"NCAA Participation Rates Going Up." *National Collegiate Athletic Association*. Turner Sports Interactive, 02 Nov. 2011. Web. 09 Nov. 2014.

Outdoor Participation Report 2013 (2013): 5. The Outdoor Foundation. Web. 09 Nov. 2014.

"Testing the Proportional Hazard Assumption in Cox Models." *UCLA: Statistical Consulting Group*. UC Regents, n.d. Web. 11 Nov. 2014.