Statistical Analysis Plan for the Prevalence of Brain Health in Former Professional Football Players

(Brain Health Initiative)

NCT Number: NCT04708613

Date: 08/09/2018

Study Overview

This project will define, for the first time, the prevalence of brain health (i.e., normal cognitive, neuromotor, behavioral function) versus neurodegeneration in living former professional football players and group-matched controls through a comprehensive assessment of clinical, neuroimaging, and biomarker measures.

More than 150 million youth have played football in the United States. To date, there are 202 documented cases of chronic traumatic encephalopathy (CTE) in deceased, former football players (i.e., a rate of 0.00015%). Researchers have documented long-term mental health outcomes in former professional football players. However, these studies have yet to address the prevalence of chronic traumatic encephalopathy (CTE), because at present the diagnosis of CTE can only be confirmed at autopsy. To date, no statistically valid prevalence study of brain health versus neurodegeneration in former professional football players and controls has been completed to establish the prevalence of normal cognitive function versus neurodegeneration that can be extrapolated to the population of former professional football players. Prevalence determination will form the foundation for powering and statistically analyzing future studies of diagnostic, prognostic, and treatment modalities currently under development and investigation.

Our study has three Specific Aims:

Specific Aim 1: Establish the prevalence of brain health and neurodegeneration through blinded review of a comprehensive assessment of clinical (cognitive, neuromotor, behavioral function), neuroimaging (PET, MRI), and biomarker measures in former professional football players using a random and non-random sample of eligible former players.

Specific Aim 2: Compare the prevalence and characteristics of neurodegeneration in former professional football players to unexposed groups [1) friend and brother controls and 2) community controls] using the comprehensive assessment.

Specific Aim 3: Analyze factors (exposure to contact sports, concussion history, substance abuse, psychiatric disorders, cardiovascular disease, etc.) associated with increased cognitive impairment and neurodegeneration in former professional football players and unexposed groups [1) friend and brother controls and 2) community controls].

Our hypothesis is that rates of cognitive impairment (primary outcome measure) and clinical syndromes of neurodegeneration (secondary outcome measure) will be similar in former professional football players to rates in group-matched controls. Based on our power analysis, we will randomly sample 150 former professional football players and 150 controls over four years.

Study Population
This is a 4-year, observational, multisite study enrolling 300 study participants.

<u>Exposed Group</u>

Participant Identification: Former professional football players aged 29-59 who are eligible for a professional football pension (i.e., played a minimum of 3 seasons, with a minimum of 3 games in each season) will comprise the “exposed’ group. There will be 2 cohorts of former professional football players in this group: (a) random sample, and (b) non-random sample. These are described below.

a. Random-Sample participants

Recruitment of random-sample former professional football players will be conducted via collaboration with the Football Players Health Study (FPHS) at Harvard University. The FPHS includes a cohort of former NFL players who played since 1960. Subjects for the Brain Health Initiative (BHI) will include a random sample of enrollees in the FPHS who meet basic eligibility for BHI.

b. Non-Random participants:

We will also enroll a non-random cohort of former professional football players. These players will not be generated from a randomized list, but rather will comprise volunteers who approach the teams at Harvard and/or Pitt interested in participating.

Inclusion Criteria: Former NFL player aged 29-59, who played a minimum of 3 seasons, with a minimum of 3 games in each season.

Exclusion Criteria: Prior history of schizophrenia or other major neurological disorder.

Stratification Plan for Random Sampling: To ensure accurate and valid prevalence estimates, potential participants will include those from a randomly selected list of former players 29-59 years old and eligible for a professional football pension. Stratified random sampling will be used to minimize potential confounding of age (29-44 and 45-59), race (African-American vs. Not) and professional football positioning grouping. The professional football position grouping is defined as follows:

1. Offensive/Defensive Line
2. Linebacker/Quarterback/Running Back
3. Defensive Back/Tight End/Wide Receiver/Special Teams Specialists
4. Kicker/Punter

The sampling of participants within each stratum will be equal except for the position grouping, which will be sampled in a 3:3:3:1 ratio, which reflects the distribution in the overall sample.

Within each of the 16 strata, a uniform random number will be assigned to each participant in the FPHS. Within each stratum, the population will be sorted by the random number, lowest to highest. The participants will then be contacted until in order, by the randomly assigned number – lowest to highest, until the required number of participants for each stratum is obtained.

In order to achieve the 3:3:3:1 ratio for the interim analysis of 50 participants, there will need to be fifteen participants per stratum in the position strata that exclude the kicker/punter. There will need to be five participants for the kickers/punters strata. This would yield 50 participants. Additional strata will include age groups in a 1:1 ratio (29-44 and 45-59), as well as race in a 2:1 ratio (Black:White). We will not be randomizing with these additional strata prior to the interim analysis, but will weight the stratification as needed after the first 50 former player participants have been enrolled to ensure the appropriate balance in our final cohort.

<u>Unexposed Groups</u>

Participant Identification: There will be 2 cohorts of unexposed groups: 1) friend and brother controls and 2) community controls. All unexposed controls will be group-matched to former NFL participants by age, ethnicity, and education.

Inclusion Criteria: Males, group-matched by age (29-59 years), educational background and ethnicity to cohort of former NFL players.

Exclusion Criteria: History of participation in organized football beyond high school, history of severe TBI, and prior history of schizophrenia or other major neurological disorder.

Outcomes of Interest
The presence of cognitive impairment, neurodegeneration and other clinical syndromes will be adjudicated at a Diagnostic Consensus Conference (DCC) through the Alzheimer's Disease Research

Center (ADRC) of the University of Pittsburgh. The DCC will be comprised of a neurologist, neuropsychiatrist, neuropsychologist, and radiologist and will be provided results of the comprehensive assessment of clinical, neuroimaging, and biomarker tests blinded to group-identifying information. This blinded consensus conference methodology has been successfully used at the ADRC (Lopez 2000a, 2000b), in the Cardiovascular Health Study (Lopez 2003a, 2003b), and in other large-scale studies that focused on neurological disorders and cognition (Lopez 2014; Dekosky 2008). The components of the comprehensive assessment panel are drawn from extensive work of the Alzheimer's Disease Neuroimaging Network (ADNI) and from the TBI clinical and scientific communities.

Diagnostic Consensus Conference: All participants will be discussed and classified at a Diagnostic Consensus Conference. Presence or absence of cognitive impairment (primary outcome measure) and of neurodegeneration (secondary outcome measure) will be adjudicated through blinded review of the comprehensive assessment data by a team of neurology, neuropsychiatry, neuropsychology, and radiology and experts.

Mild Cognitive Impairment Criteria (Primary Outcome): The diagnosis of MCI is based on the ADRC National Alzheimer's Coordinating Center (NACC) MCI criteria (Petersen, 2003; Lopez, 2003) and uses the chart in the NACC manual (Petersen, 2005): 1) On review of all clinical and cognitive information, the DCC judges that a patient does not have normal cognition for age, and 2) Cognitive decline is determined to be present; and Subjects will be adjudicated as: No cognitive impairment, Mild Cognitive Impairment, , Dementia. The option of "Dementia" is included for adjudication of cognitive impairment because we may detect people with a true dementia syndrome during the course of the study.

Trauma-Related Neurodegeneration Criteria (Secondary Outcome): The determination of trauma-related neurodegeneration will be based on: 1) presence of cognitive deficits, and/or neuromotor manifestations, and/or behavioral/mood-related symptoms; 2) PET scans demonstrating abnormal amyloid and/or tau deposition; 3) quantitative MR evidence of cortical volume loss and/or white matter injury; and 4) elevated CSF tau/Aβ42 ratio. The analysis of trauma-related neurodegeneration will take into consideration history of TBI/concussion, for example utilizing the OSU-TBI survey (Corrigan, 2007).

Sample Size and Power Analysis

The sample size needed for the study assumes a study-wide type I error of 0.05, a two-sided alternative hypothesis, an equal number of exposed and non-exposed individuals, and 80% power. In addition, it is assumed that based on the available pool of former players, 150 exposed individuals will be enrolled, as well as 150 controls.

The sample size calculations are based on a chi-square test examining the relationship between exposure and mild cognitive impairment. In order to calculate the sample size, the proportion of individuals with MCI in the control population must be estimated. At this time, this proportion for the age range for subjects in this study is unknown (but can reasonably be approximated from large-scale studies in the dementia field). Based on the information described below, if the baseline MCI rate in the control group is 5%, then an increase to 14.6% of MCI in the exposed group can be detected. *Assuming 5% of controls show impairment (P2 = .05)*, a sample size of 150 former NFL players and 150 unexposed controls would be sufficient to detect a 14.6% MCI prevalence (P1 = .146) in the exposed (NFL group) (i.e. a difference between the group proportions of 0.0960). This assumes no additional covariates in the model. If the rates of MCI are maintained as above, and a one-sided alternative is used, the number of participants that need to be studied can be reduced from 300 to 236.

For two sample t-test analyzing neurodegeneration, we expect to have 150 subjects in each group for analyses. There is sufficient power (0.80) to detect differences of 0.30 (0.50, 0.40, 0.50) standard deviations or greater assuming a 2-sided significance level of 0.05 (0.0005, 0.025, 0.0007) [Bonferroni correction 0.05/96=0.0005 for PiB PET and AV-1451 PET, 0.05/2=0.025 for hippocampal and ventricular volume, 0.05/68=0.0007 for cortical thickness].

Interim Analysis and Sample Size Re-Estimation
An interim analysis will be conducted when 33% of exposed cases and controls (N=50 in each group) have been evaluated. To control for the overall type I error rate, the Lan and DeMets group sequential monitoring approach with an O'Brien and Fleming spending function will be implemented.

Typically, in the course of a randomized clinical trial, an interim analysis would be kept confidential to prevent biases impacting the remainder of the trial. In this instance, no possible biases can be identified at this time. The release of the results would not impact any exposure (since it has already occurred) or the delivery of any treatment (since none exists).

The information from the interim analysis will also be used to re-estimate the sample size. As was stated earlier, the proportion of MCI in the control group is not perfectly known. The observed rate in the first 50 control subjects will be used to calculate the size of the effect that can be detected with the available sample size.

Potential Study Limitations
Although the criteria for determination of mild cognitive impairment are broadly accepted, there is no standard definition nor diagnostic criteria for trauma-related neurodegeneration. The impact of this project is solidified by the use of mild cognitive impairment as the primary outcome measure. Nevertheless, our proposed criteria for the adjudication of trauma-related neurodegeneration may prove difficult when applied at this scale. We will rely heavily on the expertise within the Steering Committee, whose members are considered the leading experts globally for the various facets of neurodegenerative disorder assessment (clinical, radiographic, and biomarker tests). Should weaknesses in the comprehensive assessment battery be identified, we will look to the Steering Committee to recommend changes following the midpoint analysis.

Another potential risk, inherent to nearly every clinical study, is recruitment. The team physicians who have engaged in this effort have decades of experience and continuity amongst them. We believe these longstanding relationships will serve to empower success of this important effort.

Brain Health Initiative Principal Investigators

Principal Investigator: David O. Okonkwo, MD, PhD

Co-PIs: Michael W. Collins, PhD
Anthony P. Kontos, PhD
James Mountz, MD
Oscar Lopez, MD

Contact Info: Department of Neurological Surgery
University of Pittsburgh
200 Lothrop St. Suite B-400
Pittsburgh, PA 15213
OkonkwoDO@upmc.edu

Statistical Analysis Plan Overview

This document describes the analytic plan of the aims for the Prevalence of Brain Health in Former Professional Football Players (Brain Health Initiative) Study. It first details the general analytic approach for the primary aims and then provides details for each aim.

Significance Level: All statistical analyses will be two-sided and the significance level will be 0.05. Significance level will be adjusted for multiple comparison correction by using a Bonferroni correction for multiple hypothesis testing.

Missing Data: Missing data for the variables of interest will be investigated in an effort to understand how they are missing: (1) missing completely at random (MCAR), (2) missing at random (MAR) or (3) non-ignorable. Little's test will be performed and if the test indicates that the complete case data set is a random sample we will continue without imputing missing values. If Little's test indicates that the data set of complete cases is not a random sample, we will report the point estimates and their 95% confidence intervals (CI) by applying a worst/best scenario imputation for the missing values. If the worst/best case analyses allow for the same conclusion, we will not perform multiple imputations. However, if the worst/best case imputation provides different conclusions, multiple imputations will be performed, creating ten imputed data sets under the assumption of missingness at random. The result of the study will be the pooled intervention effect and 95% CI after multiple imputations.

Outcome of Interest: All participants will be discussed and classified at a Diagnostic Consensus Conference. Presence or absence of cognitive impairment (primary outcome measure) and of neurodegeneration (secondary outcome measure) will be adjudicated through blinded review of the comprehensive assessment data by a team of neurology, neuropsychiatry, neuropsychology, and radiology and experts.

> Primary Outcome: The primary outcome is mild cognitive impairment criteria. Analysis will consider mild cognitive criteria as dichotomous outcome (that is, yes/no).
>
> Secondary Outcome: The secondary outcome is trauma-related neurodegeneration criteria. Analysis will consider trauma-related neurodegeneration criteria as dichotomous outcome (that is, yes/no).

Both the Primary and Secondary Outcomes will compare the exposed group with: 1) friend and brother controls and 2) the full unexposed group (friend and brother controls, and community controls).

Descriptive Analysis: Descriptive statistics will be presented for all study participants overall and by group (exposed and unexposed). Categorical variables will be summarized by frequencies and percentages. Percentages will be calculated according to the number of patients where data are available. Where values are missing, the actual denominator will be stated. Continuous variables will be summarized using standard measures of central tendency and dispersion, using the mean and standard deviation for normally distributed continuous data or median and interquartile range for non-normally distributed continuous data. Categorical characteristics will be compared with a Chi-square test or Fisher's Exact test for categorical treatments. Continuous characteristics will be compared with a two-sample t-test or analysis of variance (or nonparametric equivalent) for categorical treatments.

Specific Aim 1: Establish the prevalence of brain health and neurodegeneration through blinded review of a comprehensive assessment of clinical (cognitive, neuromotor, behavioral function), neuroimaging (PET, MRI), and biomarker measures in former professional football players using a random sample of eligible former players.

Objective*:* The goal of this aim is to establish the true prevalence of brain health versus neurodegeneration in a sample of living, former professional football players.

Population/Subgroup: This analysis will be conducted on all former NFL players (exposed group)

Primary Outcome: Mild cognitive criteria as dichotomous outcome

Secondary Outcome: Trauma-related neurodegeneration criteria as dichotomous outcome

Statistical Analysis: The analyses will begin with basic descriptive statistics (including counts and percentages for categorical data and mean and standard deviation for normally distributed continuous data or median and interquartile range for non-normally distributed continuous data), and the examination of distributional assumptions for all study measurements by groups. These analyses will be used to identify potential influential observations and/or outliers. Tests for a priori hypothesis will be two-tailed and the significance level will be set to 0.05, unless stated otherwise.

The point prevalence of cognitive impairment (primary outcome measure) and of neurodegeneration (secondary outcome measure) will be defined as the proportion of former NFL players with cognitive impairment based on the comprehensive assessment of clinical measures. That is, point prevalence will be defined as number of former NFL players in the sample with cognitive impairment divided by the total number of former NFL players in the sample. The point prevalence of neurodegeneration (secondary outcome measure) will be similarly be defined as the proportion of former NFL players with a clinical syndrome of neurodegeneration based on the comprehensive assessment of clinical, neuroimaging and biomarker measures. We will also construct a 95% confidence interval for the estimated proportions.

Specific Aim 2: Compare the prevalence and characteristics of neurodegeneration in former professional football players to unexposed control groups [1) friend and brother controls and 2) community controls] using the comprehensive assessment.

Objective*:* The goal of this aim is to study the impact of a career in the NFL on development of neurodegeneration, compared to unexposed controls.

Population/Subgroup: This analysis will be conducted on all former NFL players (exposed group) and unexposed groups [1) friend and brother controls and 2) community controls]

Primary Outcome: Mild cognitive criteria as dichotomous outcome

Secondary Outcome: Trauma-related neurodegeneration criteria as dichotomous outcome

Statistical Analysis: Analyses will begin by comparing the characteristics of the exposed group (former NFL players) and unexposed group [1) friend and brother controls and 2) community controls] (e.g. age, race, BMI, APOE status, etc.). Two sample t-test (continuous data, e.g. age) and Chi-Square test (categorical data, e.g. race) will be used to determine differences in subject characteristics between exposed and unexposed control groups. Mann-Whitney U test will be used if the assumptions for the parametric test are not met. Fisher's exact test will be used for sparse data.

Prevalence of cognitive impairment (primary outcome measure) and of neurodegeneration (secondary outcome measure) for the unexposed control groups will be defined as in Aim 1.

Primary Analysis: Chi-Square test will be used to compare the proportion of subjects with MCI and neurodegeneration based on clinical, neuroimaging and biomarkers between the exposed (former NFL players) and unexposed control groups (defined by 1) friends and brothers 2) friends and brothers and community controls) groups. Fisher's exact test will be used for sparse data.

Secondary Analysis: Chi-Square test will be used to compare the proportion of subjects with MCI and neurodegeneration based on clinical, neuroimaging and biomarkers between the exposed (former NFL players) and unexposed control groups [1) friend and brother controls and 2) community controls]. Fisher's exact test will be used for sparse data.

For other characteristics of interest, Chi-Square test (or Fisher's exact test) will be used to compare categorical data between the two groups and two sample t-test (or Mann-Whitney U test) will be used for to compare continuous data between the two groups.

It is possible that subject characteristics may not be balanced between the two groups. To control for these potential confounding effects, a propensity score will be used. The propensity score will be developed using a logistic regression model with an indicator of an individual being in the exposed group (or not) as dependent variable, while the other imbalance characteristics of the population will be included as independent variables. Based on the model, a propensity score of being in each group will be estimated and will be used in a propensity score weighted logistic regression model (yes/no neurodegeneration outcome), main effects of exposure (former NFL players/control) and propensity score as a weight) to estimate the independent effect of exposure, controlling for the possible confounding effects. A combination of propensity score weighting and covariate adjustment (i.e. doubly robust estimation) will be performed if the propensity scores were not able to completely balance all the subject characteristics.

Specific Aim 3: Analyze factors (exposure to contact sports, concussion history, substance abuse, psychiatric disorders, cardiovascular disease, etc.) associated with increased cognitive impairment and neurodegeneration in former professional football players and unexposed control groups [1) friend and brother controls and 2) community controls]

Objective: The goal of Aim 3 is to determine whether any clinical syndrome of neurodegeneration identified in a participant is the result of trauma or some other etiology.

Population/Subgroup: This analysis will be conducted on all former NFL players (exposed group) and unexposed control groups [1) friend and brother controls and 2) community controls]

Primary Outcome: Mild cognitive criteria as dichotomous outcome

Secondary Outcome: Trauma-related neurodegeneration criteria as dichotomous outcome

Factors of Interest: exposure to contact sports, current BMI, concussion history, substance abuse, psychiatric disorders, cardiovascular disease, sleep disorders, pain, metabolic and endocrinological disorders, etc.

Statistical Analysis: The analysis will be similar to the analyses outlined in Specific Aim 2. We will begin by comparing subject characteristics between exposed group (e.g. yes to concussion history) and unexposed groups (e.g. no to concussion history) in former professional football players and unexposed control groups [1) friend and brother controls and 2) community controls]. Chi-Square test or logistic regression will be used to determine differences in neurodegeneration (yes/no) between exposed and unexposed groups based on the risk factor. Fisher's exact test (or Exact logistic regression) will be used for sparse data. If subject characteristics between the two groups are not balanced, we will compute a propensity score. Then, the propensity score will be used in a propensity score weighted logistic model (outcome mild cognitive impairment or neurodegeneration (yes/no), main effects of risk factor, and

propensity score as weight). Doubly robust estimation will be performed if the propensity scores were not able to completely balance all of the subject characteristics.

References

Corrigan, J.D., Bogner, J.A. (2007). Initial reliability and validity of the OSU TBI Identification Method. Journal of Head Trauma Rehabilitation*,* 22(6), 318-329.

DeKosky ST, Williamson JD, Fitzpatrick AL, Kronmal RA, Ives DG, Saxton JA, Lopez OL, Burke G, Carlson MC, Fried LP et al. . 2008. Ginkgo biloba for prevention of dementia: a randomized controlled trial. JAMA 300(19):2253-2262.

Lopez OL, Becker JT, Klunk W, Saxton J, Hamilton RL, Kaufer DI, Sweet RA, Cidis Meltzer C, Wisniewski S, Kamboh MI et al. . 2000a. Research evaluation and diagnosis of possible Alzheimer's disease over the last two decades: II. Neurology 55(12):1863-1869.

Lopez OL, Becker JT, Klunk W, Saxton J, Hamilton RL, Kaufer DI, Sweet RA, Cidis Meltzer C, Wisniewski S, Kamboh MI et al. . 2000b. Research evaluation and diagnosis of probable Alzheimer's disease over the last two decades: I. Neurology 55(12):1854-1862.

Lopez OL, Jagust WJ, DeKosky ST, Becker JT, Fitzpatrick A, Dulberg C, Breitner J, Lyketsos C, Jones B, Kawas C et al. . 2003a. Prevalence and classification of mild cognitive impairment in the Cardiovascular Health Study Cognition Study: part 1. Arch Neurol 60(10):1385-1389.

Lopez OL, Klunk WE, Mathis C, Coleman RL, Price J, Becker JT, Aizenstein HJ, Snitz B, Cohen A, Ikonomovic M et al. . 2014. Amyloid, neurodegeneration, and small vessel disease as predictors of dementia in the oldest-old. Neurology 83(20):1804-1811.

Lopez OL, Kuller LH, Fitzpatrick A, Ives D, Becker JT, and Beauchamp N. 2003b. Evaluation of dementia in the cardiovascular health cognition study. Neuroepidemiology 22(1):1-12.

Petersen RC. 2003. Mild cognitive impairment as an early indicator of Alzheimer's disease. Menopause-the Journal of the North American Menopause Society 10(6):562-562.

Petersen RC, and Bennett D. 2005. Mild cognitive impairment: Is it Alzheimer's disease or not? Journal of Alzheimers Disease 7(3):241-245.