Abstract
In general, National Football League (NFL) players tend to live longer than the general population. However, little information exists about the long-term mortality risk in this population. Frequent, yet mild, head trauma may be associated with early mortality in this group of elite athletes. Therefore, career playing statistics can be used as a proxy for frequent head trauma. Using data from Pro Football Reference, we analyzed the association between age-at-death, position, and NFL seasons-played among 6,408 NFL players that were deceased as of July 1, 2018. The linear regression model allowing for a healthy worker effect demonstrated the best fit statistics (F-statistic = 9.95, p-value = 0.0016). The overall association of age-at-death and seasons-played is positive beginning at the 10.75 and 10.64 seasons-played point in our two models that feature seasons-played and seasons-played squared as explanatory variables. Previous research that does not account for the healthy worker effect may not adequately describe mortality risk among NFL players.

Keywords
CTE, concussions, football, gridiron football, NFL, chronic traumatic encephalopathy, sports
Corresponding author: Justin Ehrlich (jaehrlic@syr.edu)

Author roles: Ehrlich J: Conceptualization, Data Curation, Investigation, Methodology, Software, Visualization, Writing – Original Draft Preparation; Kmush B: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; Walia B: Conceptualization, Formal Analysis, Investigation, Methodology, Validation, Writing – Review & Editing; Sanders S: Conceptualization, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by funds provided by the David B. Falk College of Sport and Human Dynamics, Syracuse University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Ehrlich J et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Ehrlich J, Kmush B, Walia B and Sanders S. Mortality risk factors among National Football League players: An analysis using player career data [version 1; peer review: awaiting peer review] F1000Research 2019, 8:2022 (https://doi.org/10.12688/f1000research.21235.1)

Introduction

Very little information exists about mortality and long-term health outcomes among National Football League (NFL) players. Elite football players tend to have a lower overall mortality rate than the general population, often attributed to routine physical activity. However, this occupational group cannot be directly compared to the general population. Several studies in small numbers of NFL players have found an association between traumatic brain injuries with depression, suicide, dementia, and chronic traumatic encephalopathy. There is mounting evidence that even sub-clinical head impacts, especially when they occur frequently, can also lead to these adverse health outcomes. However, these relationships are difficult to study systematically due to few cases, challenges with diagnostics, and long lag time from the injury to symptom onset. Yet, there exists a rich repository of data surrounding NFL career playing statistics. We hypothesize that certain player career attributes, including position-of-play and seasons-played, are likely to be strong predictors for mortality from repeated, yet mild, head trauma. Here, we study the association between mortality and NFL seasons-played, while controlling for playing position.

Methods

Data was collected from Pro Football Reference, a free online database maintained by Sports Reference LLC that includes playing statistics from every player in NFL history, over 25,000 in total, with meticulously recorded data beginning in 1922. Merged and transformed data is available from an open repository. Variables of interest include birthdate, death date, position, and seasons-played. Individuals with any missing data were eliminated, leaving 24,740 players. Of those, 6,408 (25.9%) had died according to Pro Football Reference, as of July 1, 2018. Playing position was divided into three standard categories according to previous literature. Category 1: defensive back, quarterback, wide receiver, and kicker: 1,600 dead/8,415 players (19%). Category 2: running back, linebacker, tight end: 1,690 dead/7,228 players (23%). Category 3: offensive and defensive linemen: 3,118 dead/9,097 players (34%).

Statistical analysis

Expected age-at-death was calculated from the 2017 National Vital Statistics Report using average years of life remaining at 20 years of age for the decade of the 20th year plus 20. Age-at-death residuals were calculated as observed age-at-death minus expected age-at-death. This analysis was completed in Stata Version 1413, and data was visualized using R 3.6.114. Associations were assessed using linear regression models with a quadratic term for seasons-played. Specifically, we use (position) fixed-effect ordinary least squares modeling to determine whether associations exist between age-at-death residual, number of NFL seasons-played (squared), and position category fixed effects. In these models, we seek to assess whether career duration exposure relates significantly to age-at-death residual conditional on position-of-play. The healthy worker turning point was calculated using standard differential calculus techniques (i.e., calculating the minimum point of a best fit surface).

Base Model:

\[ \text{Age at Death Residual}_{i,t} = \beta_0 + \beta_1 \text{Number of Seasons Played}_{i,t} + \epsilon_{i,t} \]

Seasons-played Squared Model:

\[ \text{Age at Death Residual}_{i,t} = \beta_0 + \beta_1 \text{Number of Seasons Played}_{i,t} + \epsilon_{i,t} + \beta_2 \text{Number of Seasons Played}^2_{i,t} + \epsilon_{i,t} + \beta_3 \text{Position Category}_{i,t} + \epsilon_{i,t} \]

Position Category Fixed Effects Model

\[ \text{Age at Death Residual}_{i,t} = \beta_0 + \beta_1 \text{Number of Seasons Played}_{i,t} + \epsilon_{i,t} + \beta_2 \text{Number of Seasons Played}^2_{i,t} + \epsilon_{i,t} + \beta_3 \text{Position Category}_{i,t} + \epsilon_{i,t} \]

Results and discussion

Table 1 indicates substantial demographic sample variation between players of different position categories in height, weight, BMI, and age-at-death. Figure 1a-Figure 1b indicate a possible healthy worker effect among players of Category I and II. Certain healthy or durable players can play an increased number of seasons without a corresponding reduction in expected age-at-death as compared to players of shorter career duration.

The Seasons-played Squared and Position Category Fixed Effects models specify a quadratic term for number of NFL seasons-played. For both models, the coefficient for this variable is significant and improves the model’s explanatory power according to an Anova F-test for difference in overall model significance (F-statistic = 9.95, p-value = 0.0016; F-statistic=10.98, p-value<0.001) (Table 2). We calculate that

**Table 1. Demographics of deceased National Football League (NFL) players (1922–2018).**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total</th>
<th>Category 1 Players</th>
<th>Category 2 Players</th>
<th>Category 3 Players</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6408</td>
<td>1600</td>
<td>1690</td>
<td>3118</td>
</tr>
<tr>
<td>Average Age-at-death (sd) years</td>
<td>69.1 (15.8)</td>
<td>69.5 (15.8)</td>
<td>68.0 (16.4)</td>
<td>69.6 (15.3)</td>
</tr>
<tr>
<td>Median Seasons Played (IQR)</td>
<td>2 (3)</td>
<td>3 (4)</td>
<td>2 (4)</td>
<td>2 (3)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>27.6 (2.73)</td>
<td>25.8 (1.55)</td>
<td>27.4 (2.19)</td>
<td>28.6 (2.97)</td>
</tr>
<tr>
<td>Height (sd) (cm)</td>
<td>184 (6.04)</td>
<td>181 (5.40)</td>
<td>183 (5.68)</td>
<td>186 (5.94)</td>
</tr>
</tbody>
</table>

BMI – body mass index
overall association of age-at-death residual and seasons-played is positive beginning at 10.75 and 10.63 seasons-played for the Seasons-played Squared and Position Category Fixed Effects model, respectively. This demonstrates a healthy worker effect, where certain players are not as prone to play-related mortality risk. For this player subset, the healthy worker effect is sufficiently strong to dominate an observed mortality risk effect, where the latter effect drives the negative relationship between seasons-played and age-at-death residual for those playing fewer than 10.75 (10.63) seasons. The healthy worker effect and the mortality risk effect hold conditional upon position category control variables, where controls exhibit substantial variation in effect upon age-at-death residual as in previous literature\textsuperscript{11}. However, dividing players into three position categories may not sufficiently capture the differing on-field exposures that may contribute to mortality.

**Conclusion**

This paper finds evidence of both player health risk (in terms of age-at-death residual) for increasing NFL seasons played and
Table 2. Linear regression models predicting age-at-death Residuals among National Football League (NFL) players (1922–2018) N=6408.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Base</th>
<th>Seasons-played Squared</th>
<th>Position Category Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.402</td>
<td>0.315</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Seasons-played</td>
<td>-0.562</td>
<td>0.073</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Seasons-played Squared</td>
<td></td>
<td>0.054</td>
<td>0.017</td>
</tr>
<tr>
<td>Position Category 1</td>
<td></td>
<td>-0.042</td>
<td>0.515</td>
</tr>
<tr>
<td>Position Category 2</td>
<td></td>
<td>-2.277</td>
<td>0.504</td>
</tr>
<tr>
<td>Position Category 3</td>
<td>Reference</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Observations</td>
<td>6408</td>
<td>6408</td>
<td>6408</td>
</tr>
</tbody>
</table>

Figure 1b. Age-at-death residual versus seasons-played for category 1 and 2 deceased National Football League (NFL) players (1922–2018) N=3,290. Dots represent individual players; Solid line represents a quadratic trend.
healthy worker effects among NFL players. For Category I and II players, the latter effect dominates the former for NFL players with sufficient career survivorship. This effect holds conditional upon position-of-play control variables. Previous research not accounting for the healthy worker effect may not adequately describe mortality risk among NFL players.

**Future work**

We are pursuing additional research to examine the association of on-field playing characteristics with mortality and cause of death among NFL players.

**Ethics**

This study was determined by the Syracuse University Institutional Review Board to not be human subjects research and therefore, not to require review and oversight.

**Data availability**

**Underlying data**


This project contains the following underlying data:

- Football1922_2018.tab (Includes player name, position, year of birth, year of death, death residual, position category, and the number of seasons-played.)

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

**References**

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com