Appendix 1 (as supplied by the authors): Additional statistical information on the risk of suicide following a concussion on weekends and weekdays

This appendix is intended to provide readers additional statistical information on the risk of suicide following a concussion on weekends and weekdays. The main sections show the (§1) Overview of Analysis, (§2) Data Extraction, (§3) Data Summary, (§4) Statistical Analysis, (§5) Testing the Proportional Hazards Assumption, (§6) Additional Patient Subgroup Analyses, (§7) Separating New from Lingering Concussions, (§8) Contrasting Concussions with Ankle Sprains, (§9) Estimating Absolute Counts, (§10) Alternative Model Validation, (§11) All cause Mortality. The content primarily addresses technical coding details as well as additional results examining robustness.

(§1) Overview of Analysis

We conducted a longitudinal cohort study to examine the long-term risk of suicide among individuals who had a concussion during the study period, April 1992 to March 2012. Patients were followed from initial concussion until the end of study or death. The objective was to assess the risk of suicide associated with concussions and distinguish concussions occurring on weekends.

The Cox proportional hazards model was used to analyze data. Since a concussion is a potentially repeated exposure, we used two variables related to patients' concussion events as time-dependent covariates: first, the cumulative exposure to concussion as measured by the number of concussions a patient had during observation; second, whether each concussion occurred on a weekend or weekday.

We used the procedure PHREG in SAS software with a counting-style process to incorporate time-dependent repeated exposures in the Cox proportional hazards model. In short, counting style is a type of data structure with each patient receiving multiple lines of data where each line represents the observation period when the covariate of interest is stable in its value.

Creating counting-style data layout in SAS has been discussed elsewhere in detail (1-2). Here, we describe how the data for our study were manipulated and analyzed in SAS.

(§2) Data Extraction

```
/* Two datasets were created. The first dataset contained multiple records for each
patient identifying all physician claims for concussion. The second dataset contained
a single record per patient identifying time-fixed patient characteristics.

The first dataset (data.all_subseq_con) contained the following variables:
IKN         Patient identifier
INDEX       First concussion date
SERVDATE    Date on which physician rendered service for concussion

The second dataset (data.thump_final) contained the following variables:
IKN         Patient identifier
INDEX       First concussion date
SUICIDE     Indicator for outcome with value 0 for being alive, 1 for death by suicide
            and 2 for death by other causes
AGE         Age in years
SEX         Sex (male or female)
DTHDATE     Date of death
*/
```

Appendix to: Fralick M, Thiruchelvam D, Tien HC, et al. Risk of suicide after a concussion.
CMAJ 2016. DOI:10.1503/cmaj.150790.

```
/* Dataset with subsequent concussions - only single physician claim per patient per
day was allowed */
proc sort data=data.all_subseq_con out=conc nodupkey;
     by ikn servdate;
run;

/* Give each subsequent concussion claim an id */
data conc2;
     set conc;
     by ikn;
     retain concid;
     if first.ikn then do;
        concid=1;
     end;
     else do;
     concid+1;
     end;
run;


proc sort data=conc2;
     by ikn concid;
run;


/* All subsequent physician claims for a concussion diagnosis were included in this
dataset.  We then applied a 28 day rule to distinguish concussions that were new from
those that were lingering in a series of physician claims.  If 28 days or more
elapsed between the previous concussion claim and the next concussion claim, then the
current claim was considered a separate concussion.  Otherwise, the claim was
considered part of the previous concussion and not a new exposure.

The following variables were created:
1) lagdt  = previous concussion date
2) difdt  = number of days between current and previous concussion dates
3) concno = number of additional concussions (by 28 day rule) */

data A;
     set conc2;
     by ikn concid;
     retain concno;
     lagdt = lag(servdate);
     difdt = dif(servdate);
     if first.ikn then do;
        lagdt=index;
        difdt=servdate-index;
        concno=1;
     end;
     if difdt>=28 then concno=concno+1;
     format lagdt date9.;
run;


proc sort data=A;
     by ikn concno servdate;
```

```
run;

/* Identify the earliest date patient visited for subsequent concussion (first
physician claim for each additional concussion) */
data B;
     set A;
     by ikn concno;
     if first.concno;
     keep ikn index concno servdate;
run;


/* Keep records with only new subsequent concussions (ie. Concno>1) */
data subseqconc;
     set B;
     if concno>1;
     length conc_wk $8;
          if weekday(servdate) in (1,7)         then conc_wk='Weekend';
     else if weekday(servdate) in (2,3,4,5,6)   then conc_wk='Weekday';
     keep ikn servdate conc_wk concno;
run;



/* Create the records for the first / index concussion from single record per patient
dataset */
proc sort data=data.thump_final out=final;
     by ikn;
run;


data index(rename=(index=servdate));
     set final;
     length conc_wk $8;
          if weekday(index) in (1,7)         then conc_wk='Weekend';
     else if weekday(index) in (2,3,4,5,6)   then conc_wk='Weekday';
     concno=1;
     keep ikn index conc_wk concno;
run;


/* Append all concussions (index plus subsequent concussions) */
data allconc;
     set index subseqconc;
run;


proc sort data=allconc;
     by ikn concno servdate;
run;


/* Copy time-invariant covariates such as age and sex for patients with multiple
records.  This merging also adds back the patients who had no additional concussions
as well */
data final3;
```

```
     merge allconc(in=a)
           final(in=b);
     by ikn;
     if a or b;
run;


/* Create the final dataset (one record for each concussion) */
proc sort data=final3 out=data.thump_final2;
     by ikn servdate;
run;


/* Prepare the dataset for start-stop counting process to capture time-dependent
measures*/
proc sort data=data.thump_final2 out=A;
     by ikn descending servdate;
run;



/* Patients in the risk set were re-evaluated at each concussion time for their
cumulative number of concussions at that time and whether the given concussion
occurred on a weekday or weekend.

Define the start and stop times for each time-dependent concussion event as follows
start date = OHIP claim date with concussion diagnosis (servdate);
stop date  = next concussion date(if there is another concussion) or end of follow-up
date (eof)
The dataset below is processed by descending concussion date so that the next
concussion date can be identified using SAS lag function and to re-code suicide
(outcome) variable so that the outcome changes only during the last start/stop
period.
After sorting, the first record on the dataset is the patient’s last observation
period in which the outcome occurs */

data B;
     set A;
     by ikn descending servdate;
     lagdt    = lag(servdate);
     suicidex = 0;
     if first.ikn then do;
        lagdt    = eof;
        suicidex = suicide;
     end;
     format lagdt date9.;
run;

/* Sort back by ascending concussion dates */
proc sort data=B;
     by ikn servdate;
run;


/* Calculate number of days since index date (same as the first concussion date)
and keep only the first four concussions */
```

```
data data.thump_final3;
     set B;
     start=servdate-index;
     stop =lagdt-index;
     if concno>=4 then do;
        stop=timeto;
        suicidex=suicide;
     end;
      *This step ensures that the counting process yields results identical to
      programming method;
     if start=stop then stop=start+0.0001;
     if concno>4 then delete;
run;
```

(§3) Data Summary

```
/* Here are the number of patients along with the number corresponding concussions */
```

Count of concussions	Count of patients	Count of records
1	210364	210364
2	17422	34844
3	3958	11874
4	3366	13464
Total	**235110**	**270546**

(§4) Statistical Analysis

Long-term risk of suicide was examined in relation to concussion-related characteristics, while accommodating for death by other causes as a competing risk. Proportional cause-specific hazards models were used to examine the effect of concussion-related time-dependent covariates on death by suicides (primary endpoint) accounting for death by other causes (competing endpoint).

```
/* Cox proportional hazard model for primary end point of suicides */
Proc phreg data=data.thump_final3;
      Class conc_wk(param=ref ref='Weekday');
      Model (start,stop)*suicidex(0,2)=conc_wk concno / ties=efron rl;
Run;
```

```
/* The results for the above model are displayed below */

                              The PHREG Procedure

                               Model Information

                      Data Set                  WORK.COUNT
                      Dependent Variable        start
                      Dependent Variable        stop
                      Censoring Variable        suicidex
                      Censoring Value(s)        0  2
```

Ties Handling	EFRON

Number of Observations Read	270546
Number of Observations Used	270546

Class Level Information

Class	Value	Design Variables
conc_wk	Weekday	0
	Weekend	1

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
270546	667	269879	99.75

Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	15836.005	15817.346
AIC	15836.005	15821.346
SBC	15836.005	15830.351

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	18.6592	2	<.0001
Score	20.8550	2	<.0001
Wald	20.5894	2	<.0001

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
concno	1	12.5106	0.0004
conc_wk	1	10.0387	0.0015

```
                    Analysis of Maximum Likelihood Estimates

                             Parameter      Standard
Parameter            DF       Estimate         Error    Chi-Square    Pr > ChiSq

concno                1        0.26308       0.07438       12.5106        0.0004
conc_wk   Weekend     1        0.30109       0.09503       10.0387        0.0015

                    Analysis of Maximum Likelihood Estimates

                           Hazard      95% Hazard Ratio
   Parameter                Ratio      Confidence Limits
   concno                   1.301       1.124       1.505
   conc_wk   Weekend        1.351       1.122       1.628
```

These final results lead to the following two summary interpretations related to how the risk of suicide is associated with the number of concussions and associated with whether the concussion occurs on a weekend compared to a weekday. First, each additional concussion significantly increases the risk of suicide by about 30% (hazard ratio = 1.301, confidence interval: 1.124 - 1.505). Second, having a concussion on a weekend significantly increases the risk of suicide by about 35% (hazard ratio = 1.351, confidence interval: 1.122-1.628). Together, these two interpretations are clinically relevant and the main findings from of the Cox proportional hazard model.

(§5) Testing the Proportional Hazards Assumption

One assumption of a Cox proportional hazard model is that risks are proportional and the proportion is constant over time. We chose the classic approach to test the proportionality assumption by adding a time-dependent variable to the model representing the product of survival time with the covariate (weekend vs weekday concussion). The following section provides the codes and results for the basic model with the time-fixed covariate (whether the first concussion occurred on a weekend or weekday) along with the interaction term (whether proportional hazard change with time):

```
Proc phreg data=data.thump_final3;
      Class case(param=ref ref='Weekday');
      Model timeto*suicide(0,2)=case / ties=efron rl;
      tcase=case*timeto;
Run;


/* The results for the above model are displayed below */

                                       The PHREG Procedure

                                    Model Information

                   Data Set                  DATA.THUMP_FINAL
                   Dependent Variable        timeto
                   Censoring Variable        suicide
                   Censoring Value(s)        0  2
```

Ties Handling	EFRON

Number of Observations Read	235110
Number of Observations Used	235110

Class Level Information

Class	Value	Design Variables
case	Weekday	0
	Weekend	1

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
235110	667	234443	99.72

Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	15836.005	15824.870
AIC	15836.005	15828.870
SBC	15836.005	15837.875

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.1351	2	0.0038
Score	11.9833	2	0.0025
Wald	11.8997	2	0.0026

The PHREG Procedure

Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq

```
                    case            1          7.1363          0.0076
                    tcase           1          0.4538          0.5006


                     Analysis of Maximum Likelihood Estimates

                                  Parameter      Standard
Parameter              DF      Estimate         Error    Chi-Square    Pr > ChiSq

case       Weekend      1       0.39146       0.14654        7.1363        0.0076
tcase                   1    -0.0000368     0.0000546        0.4538        0.5006

                     Analysis of Maximum Likelihood Estimates

                             Hazard      95% Hazard Ratio
     Parameter                Ratio      Confidence Limits     Label

     case      Weekend        1.479       1.110       1.971    case Weekend
     tcase                    1.000       1.000       1.000
```

Since the interaction of the main covariate (case) on time was not significant, we conclude that the proportionality assumption holds for the weekend vs weekday covariate during the course of the study. Interestingly, the actual coefficient estimating the increased risk following a weekend concussion is somewhat greater in this model that contains the (non-significant) interaction term.

(§6) Additional Patient Subgroup Analyses

We also explored a different approach by avoiding the proportional hazards model and using subgroup analyses to assess the robustness of results. Doing so was based on calculating absolute risk of suicide as the total number of suicides divided by the number of patient-years at risk. Specific subgroups were then generated from baseline characteristics with particular attention to whether the first concussion occurred on a weekend or a weekday. These subgroups appear below and examine demographic characteristics and clinical characteristics.

```
/* Here are the estimates of absolute risks and relative risks */
```

Demographic Subgroup	Total Events	Weekend Event Rate	Weekday Event Rate	Relative Risk	Confidence Interval
Full Cohort	667	39	29	1.36	1.14 – 1.64
Age ≤ 29y	194	27	22	1.24	0.90 - 1.74
Age 30-44y	279	58	41	1.39	1.05 – 1.86
Age 45-59y	137	51	34	1.49	1.00 – 2.29
Age ≥ 60y	57	31	16	2.00	1.14 – 3.65
Sex= male	481	48	40	1.22	0.99 – 1.51
Sex= female	186	26	17	1.57	1.12 – 2.23
Income = 1	102	41	23	1.81	1.18 – 2.83
Income = 2	91	31	20	1.57	0.99 – 2.56
Income = 3	123	35	28	1.28	0.85 – 2.00
Income = 4	139	41	29	1.40	0.96 – 2.10

Income = 5	212	46	42	1.10	0.79 - 1.56
Urban	573	41	29	1.43	1.21 - 1.78
Rural	94	30	30	1.01	0.62 - 1.74

Clinical Subgroup	Total Events	Weekend Event Rate	Weekday Event Rate	Relative Risk	Confidence Interval
Full Cohort	667	39	29	1.36	1.14 - 1.64
Enroll remote	487	39	31	1.28	1.04 - 1.60
Enroll recent	180	39	24	1.59	1.14 - 2.25
Image= yes	146	43	39	1.09	0.75 - 1.61
Image= no	521	38	27	1.43	1.16 - 1.76
Psych= yes	411	102	68	1.50	1.19 - 1.90
Psych= no	256	20	15	1.35	1.02 - 1.82
Hosp= yes	26	191	60	3.19	1.48 - 7.21
Hosp= no	641	37	28	1.32	1.10 - 1.60
Attempt= yes	17	1543	1149	1.34	0.52 - 3.89
Attempt= no	650	38	28	1.35	1.12 - 1.63
Nil subgroup	387	28	23	1.20	0.94 - 1.54

The results show reasonable consistency across different subgroups with almost all point estimates of relative risk overlapping the primary analysis. The subgroup with prior hospitalization is the high outlier and the subgroup with a rural home location is the low outlier. The consistency in relative risk estimates is distinctive given that absolute risks for some subgroups vary by almost two orders of magnitude. The consistency in relative risk estimates also extends to the subgroup with no past psychiatric diagnosis, hospitalization, or suicide attempt (Nil subgroup). The main interpretation is that the increase in risk of suicide following a weekend concussion is unlikely to be an artifact of the proportional hazards assumption.

(§7) Separating New from Lingering Concussions

For curiosity we also explored different thresholds for defining separate concussions. Our primary analysis applied a 28 day rule to distinguish concussions that were new from those that were lingering in a series of physician claims. We next applied a shorter threshold (14 days) and a longer threshold (56 days) and repeated our entire time-dependent repeated exposures proportional hazards model. The results appear in the table below.

	Definition	Dose Gradient	Confidence Interval	Weekend Gradient	Confidence Interval
Primary Analysis	28 days	1.301	1.124-1.505	1.351	1.122-1.628
Shorter Threshold	14 days	1.198	1.043-1.376	1.355	1.125-1.633

Longer Threshold	56 days	1.392	1.186-1.635	1.353	1.124-1.629

The results show reasonable consistency across different thresholds for defining separate concussions. As expected, apparent dose gradient correlating concussion count with subsequent risk of suicide is intensified with a longer elapsed time for defining a new concussion. As expected, the apparent weekend gradient correlating weekend events with subsequent risk of suicide is relatively consistent regardless of different thresholds. Together, these findings suggest that are main conclusions are not an artifact of uncertainty about when a concussion resolves.

(§8) Contrasting Concussions with Ankle Sprains

One more test of robustness involved switching the analysis to examining patients following an ankle sprain (code 845) instead of a concussion (code 850). The core rationale was that ankle sprains tend to heal and are less prone to cause lasting consequences compared to concussions. The purpose, therefore, was to change the data extraction code (switching the inclusion criteria from a diagnosis of concussion to a diagnosis of ankle sprain) and rerun the analysis otherwise unchanged. The hypothesis was that the increased risk of suicide following a concussion was distinct and different than following an ankle sprain. This analysis also provided a method for checking for programming errors and several other potential biases. The results appear below.

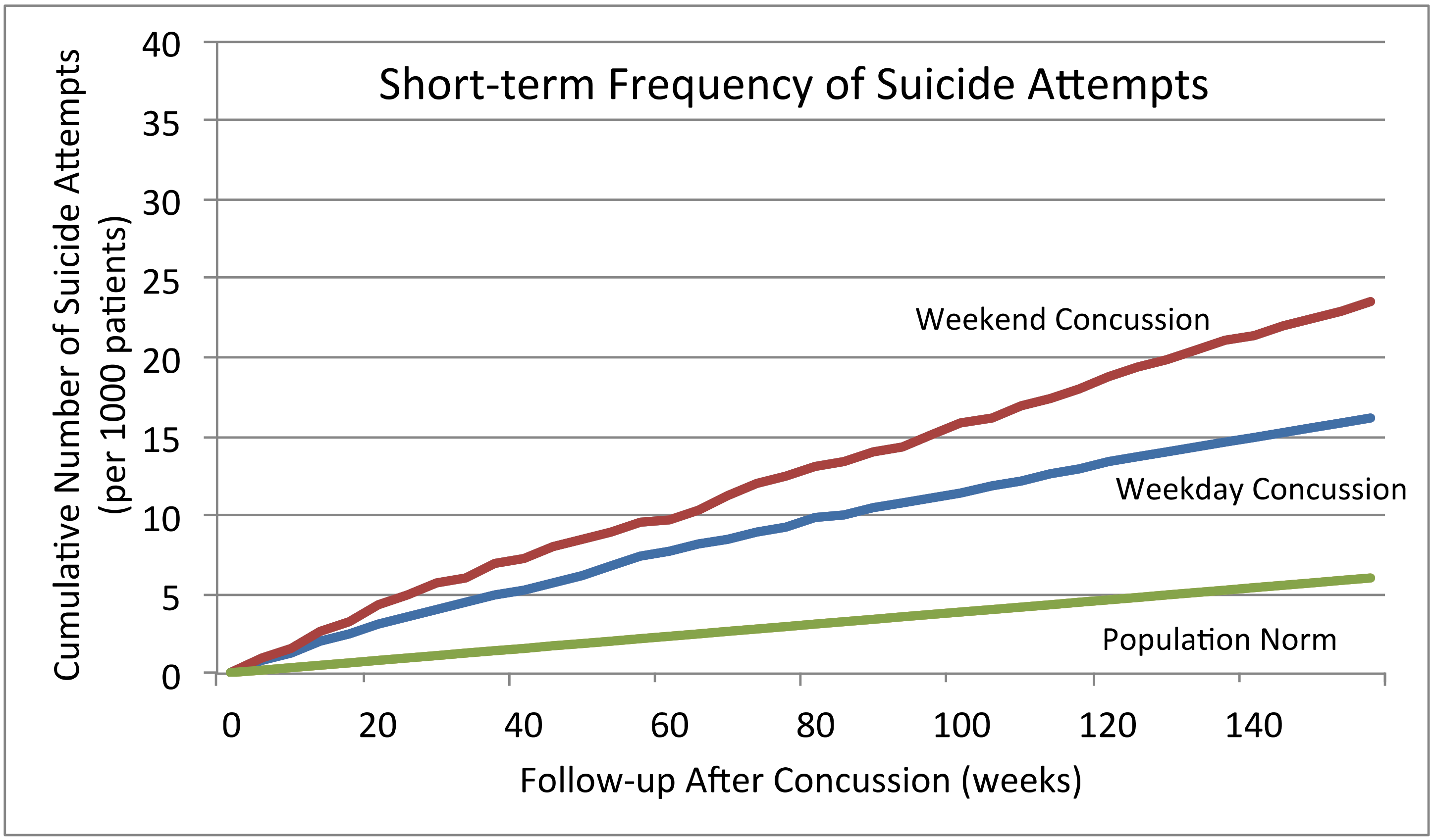
Analysis	Total Patients	Weekend Event Rate	Weekday Event Rate	Absolute Difference	Confidence Interval
Concussion	235,110	39.2	28.8	10.5	4.1 – 17.7
Ankle Sprain	2,397,192	16.5	14.8	1.7	0.4 – 2.8

These findings show that ankle sprains are more common than concussions. The absolute risk of suicide following an ankle sprain is modest (averaging 15 per 100,000 patient-years), much lower than following a concussion (averaging 31 per 100,000 patient-years), and near the population norm (averaging 9 per 100,000 patient-years). The day of the ankle sprain predicted long-term risk of suicide to only a marginal degree (perhaps due to shared factors such as alcohol). Together, the findings suggest the association of concussions with subsequent risk of suicide is not easily explained as an artifact of statistical analysis and, instead, is a reflection of the underlying data. Of course, some degree of residual uncertainty is still possible due to unmeasured confounders.

(§9) Estimating Absolute Counts

The longitudinal cohort design of our analysis also yields estimates of total counts represented by the observed increased absolute and relative risk of suicide. The baseline population average risk of suicide is about 9 per 100,000 person-years in Ontario. Overall, the study cohort encompassed 235,110 individuals followed for an average of 9.3 years, amounting to a total 2,181,635 person-years of observation. The expected number of suicide deaths was about 196 (9 x 21.81635). The observed number of suicide deaths was 667. The absolute difference was 471 (667 – 196). This is a substantial count explaining about 6% of total population deaths from suicide. The main limitation of this analysis is the possible losses to follow-up that cause our numerical evaluation to underestimate the absolute magnitude of risks.

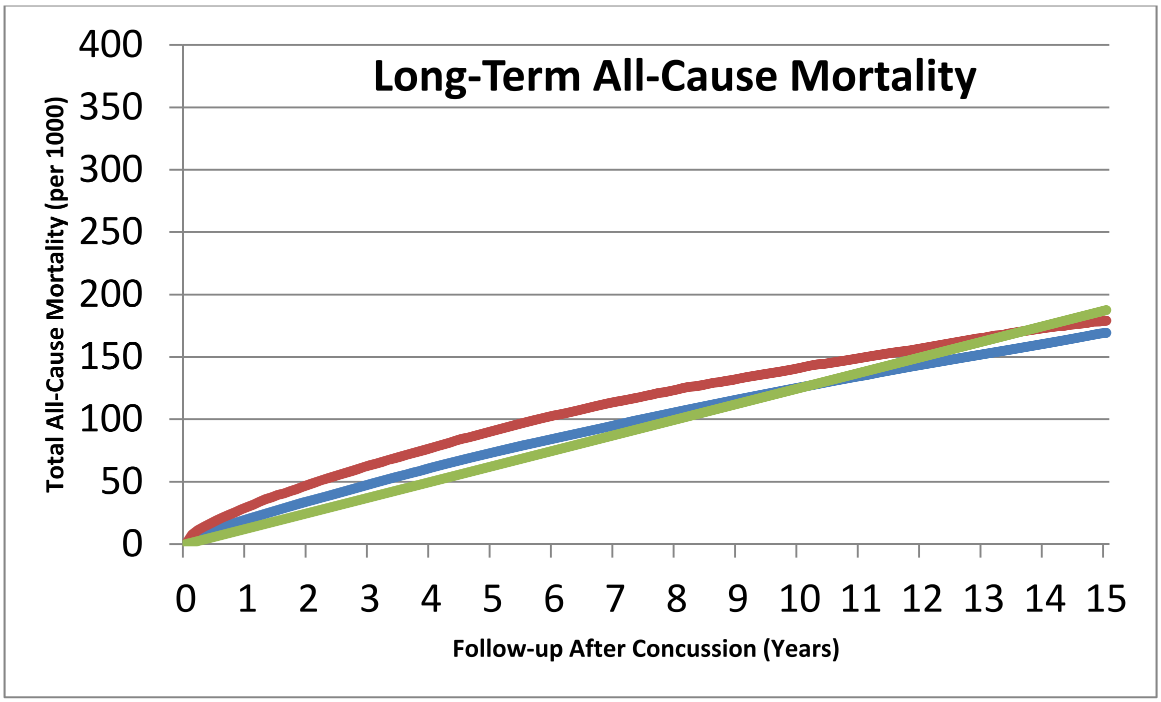
(§10) Alternative Model Validation
An additional check on the analysis applied repeated events analysis (rather than time to event analysis) and focused on suicide attempts (rather than suicide deaths). The core rationale was that suicide attempts are sometimes a harbinger of suicide deaths and yet are more readily analyzed due to higher relative frequency. The purpose, therefore, was to explore the correlation of concussion with suicide by examining suicide attempts as an intervening variable. To do so, we identified the subgroup of our cohort with first concussion between April 1, 2005 and March 31, 2012 (to guarantee a 3 year subsequent interval following the concussion for each patient). We defined segments of 28 days (to guarantee to same number of weekdays and weekends each time). We then extracted counts of emergency department visits for each patient diagnosed as suicide attempts (X60 - X84, Y10 - Y32, Y34). The cumulative results appear below.



Overall, the subgroup amounted to 91,733 patients. The total number suicide attempts was 1589, equal to an overall frequency of 5.77 attempts per 1000 patient-years. This was about three times the prevailing population norm of 2 attempts per 1000 patient-years. Patients with weekday concussions (n = 76,937) accounted for 1240 suicide attempts, equal to an overall frequency of 5.37 attempts per 1000 patient-years. Patients with weekend concussions (n = 14,796) accounted for 349 suicide attempts, equal to an overall frequency of 7.86 attempts per 1000 patient-years. The differential equaled a 46% relative increase in frequency following a concussion on weekends compared to weekdays (95% confidence interval 1.30 to 1.65). Together, the data corroborate the primary analysis and suggest that adults diagnosed with a concussion have an increased frequency of subsequent suicide attempts, particularly following concussions on weekends.

(§11) All Cause Mortality

A further check on the analysis returned to time-to-event analysis, focused on the original concussion cohort, and examined all-cause mortality (rather than suicide deaths). The rationale was to indirectly test for baseline imbalances in unmeasured factors (such as genetics, lifestyle, environment, health care) that might correlate with both the likelihood of a concussion and the probability of subsequent death. The purpose, therefore, was to change one subsection of computer code (switching the outcome from suicide death to death from any other cause) and rerun the analysis otherwise unchanged. The hypothesis was that an increased risk of suicide following a concussion was distinct and different than the risk of death due to other common diseases prevalent in the community. This analysis also provided another method for checking for programing errors (none detected). The results appear in the cumulative incidence graph below.



These findings show that all-cause mortality is more common than suicide deaths, averaging about 1,300 per 100,000 patient-years (whereas the suicide rate averaged about 31 per 100,000 patient-years). The all-cause mortality rate following a concussion was near the population norm (green line). The all-cause mortality rate following a weekend concussion (red line) was marginally higher than the all-cause mortality rate following a weekday concussion (blue line). The contrast between the analysis of all-cause mortality (shown here) and suicide deaths (shown in manuscript) is striking and suggests that adults diagnosed with a concussion have a distinctly increased long-term risk of suicide, particularly following concussions on weekends, that is not fully explained by baseline imbalances in other factors commonly contributing to mortality. A partial contribution from unmeasured factors, of course, remains possible.

References

1. Carpenter A, Ake C. Extending the use of proc phreg in survival analysis. Proceedings of the 11th Annual Western Users of SAS Software, Inc. Users Group Conference. SAS Institute Inc. Cary, NC., 2003.

2. Powell T, Bagnell M. Your “Survival” Guide to Using Time-Dependent Covariates. SAS Global Forum. 2012.