IMPROVING TRAUMA TRIAGE MODELS FOR MOTOR VEHICLE CRASHES USING EVENT DATA RECORDERS AND FUNCTIONAL DATA ANALYSIS.

Yaoyuan V. Tan[1], Carol A.C. Flannagan[2], Jonathan D. Rupp[2], Michael R. Elliott[1]

[1]*Department of Biostatistics University of Michigan*
[2]*University of Michigan Transportation Research Institute*

Abstract: Quick identification of severe injury crashes can help Emergency Medical Services (EMS) better allocate their scarce resources to improve the survival of severely injured crash victims by providing them with a fast and timely response. Data broadcast from a vehicle's Event Data Recorder (EDR) provide an opportunity to capture crash information and send them to EMS near real-time. A key feature of EDR data is a longitudinal measure of crash deceleration. We used functional data analysis (FDA) to ascertain key features of the deceleration trajectories (absolute integral, absolute in- tegral of its slope, and residual variance) to develop and verify a risk predic- tion model for serious (AIS 3+) injuries. We used data from the 2002-2012
EDR reports and the National Highway and National Automotive Sampling System (NASS) Crashworthiness Data System (CDS) datasets available on the National Transportation Safety Administration (NHTSA) website. We consider a variety of approaches to model deceleration data, including non- penalized and penalized splines and a variable selection method, ultimately obtaining a model with a weighted AUC of 0.93. A novel feature of our approach is the use of residual variance as a measure of predictive risk. Our model can be viewed as an important first step towards developing a real- time prediction model capable of predicting the risk of severe injury in any motor vehicle crash.

Key words: Clinical prediction modeling, cross-validation, motor vehicle crash injuries, precision-recall curve, weighted logistic regression, weighted receiver operating characteristic curve.

1. Introduction

Vehicle crashes are responsible for over 35,000 deaths per year in the United States and are among the leading cause of death for people ages 1-44 (Murphy

et al., 2013). Among those fatally injured, most do not die immediately, resulting in the need for a fast medical response to prevent a fatal outcome (Clark et al.,

2013). The "Golden Hour" is a term used to indicate the need for fast medical response to prevent fatalities in those who have suffered an acute event. The evidence for the golden hour is mixed (Newgard et al., 2010), but an analysis of crash data by Clark et al. (2013) showed a small improvement in survival with Emergency Medical Services (EMS) response within 30 minutes and an even larger improvement with early hospital intervention.

Hence, to improve the survival of victims with severe injury from motor ve- hicle crashes, the ability to identify these crashes and relay this information to EMS quickly is very important. The capability and increasing presence of Event Data Recorders (EDRs) in vehicles provides such an opportunity. The EDR is a "black box" that records deceleration and other signals in the event of a crash. (Note: EDRs technically report acceleration. However, deceleration is more natural when describing this trajectory.) This information can be trans- mitted immediately to a dispatcher and then to EMS after a crash, potentially resulting in faster deployment of EMS vehicles.

Although transmission of the information about the crash itself can reduce delays, having a prediction model that can identify crashes with a high risk of severe injury would help EMS to allocate scarce resources more efficiently. A number of researchers have developed such algorithms to predict which crashes have a higher probability of resulting in serious injury to one or more occupants. For example, Kononen et al. (2011) describe a triage algorithm based on EDR- available time independent data elements that predicts the presence of severe injury. They also included two additional predictors obtained through contact with a vehicle occupant. Augenstein et al. (2003) also developed a triage algo- rithm to predict serious injury, but their algorithm uses a number of predictors that cannot be obtained from EDRs.

Although the algorithms proposed by Kononen et al. (2011) and Augenstein et al. (2003) were designed to be used with EDR data, they were not directly developed from such data because there was not enough publicly available EDR data at the time of development. In order to compensate for this mismatch, both algorithms used an estimate of the change in velocity of the vehicle due to the crash (delta-v) as a primary predictor of injury outcome rather than accelera- tion histories measured by the EDR. This approach is problematic for several reasons. Delta-v estimates in crash databases have been shown to underestimate EDR-based delta-v by as much as 23% (Hampton and Gabler, 2010). Second, pa- rameters derived from deceleration histories (e.g. average deceleration) have been shown to have better linear correlation with the occurrence of injury compared to the database estimated delta-v (Mizuno et al., 2014). In addition, aspects of

the deceleration history, for example the rapid changes in the deceleration, may provide additional information about the seriousness of the injury.

In this paper, we present a new triage algorithm, developed using EDR data rather than data from crash databases. We use functional data analysis (FDA) techniques (Ramsay et al., 2009) designed for analyzing time histories, to de- velop a prediction model of serious injury in crashes. We applied FDA to the EDR-reported deceleration and tested four FDA approaches to estimate this de- celeration. Because deceleration trajectories have different duration and time stamp intervals, we summarized each trajectory using a measure for overall de- celeration, a measure for the deceleration fluctuation, a measure of the residual variability in deceleration after removing the mean trend, and a measure of the length of time the force of the crash affected the vehicle. We then combined these statistics with the limited set of time-invariant variables available from the EDR to obtain our prediction model. We also considered the degree to which the use of deceleration statistics improves the prediction capability of our model compared to a model with only time-invariant variables.

The remainder of this article is organized as follows. Section 2 summarizes the crash history data and EDR data available in the National Automotive Sampling System Crashworthiness Data System (NASS-CDS). Section 3 describes the vari- ous forms of functional data analysis (FDA) we used to estimate the deceleration data and the analyses we conducted. In Section 4 we present the results of our analyses while Section 5 discusses the strengths and limitations of our model and future extensions.

2. Data

2.1. NASS-CDS data

NASS-CDS (National Highway and Traffic Safety Administration, 2008) is an annual representative three-stage probability sample of passenger vehicle crashes sponsored by the National Highway and Transportation Safety Authority (NHTSA). To be eligible for inclusion, a crash must: (1) be police reported, (2) involve a harmful event (property damage and/or personal injury) resulting from a crash, and (3) involve at least one towed passenger car or light truck or van in trans- port on a traffic way. When a crash is selected, NASS-CDS investigators obtain police reports and conduct interviews with the occupants to collect information such as driver's age and sex, vehicle curb weight, type of vehicle, severity of in- jury measured using the Abbreviated Injury Scale (AIS) score (Association for the Advancement of Automotive Medicine, 1990), and the principal direction of impact from the crash.

Because sampling in NASS-CDS was a multi-stage design with probability of selection based on model year and the severity of injury, subjects in less serious

crashes were underrepresented in the study sample. To account for this potential bias, we used case weights equal to the inverse of the probability of selection and adjusted to known crash totals to account for the oversampling of serious crashes in NASS-CDS. These weights are provided with the data. Thus for all the vari- ance estimates and tests of association reported below, Taylor series linearization estimates were used to account for the complex sample design of the NASS-CDS (Binder, 1983; Rao and Scott, 1981).

2.2. EDR data

Electronic data records (EDRs) are devices installed in motor-vehicle to continu- ously record technical vehicle and occupant information before, during, and after a crash. Examples of key variables recorded by all EDRs include vehicle longitu- dinal acceleration/deceleration, seat belt use, vehicle speed, brake use, and airbag deployment. Newer EDRs can record acceleration data at higher frequencies, ad- ditional events, acceleration and roll angle data from multiple crash directions, and activation of collision notification systems. Vehicle identification number (VIN) is also contained in EDRs. VINs can be decoded to identify standard and equipment (both safety and otherwise) in a vehicle.

A key feature of the EDR is that, in the event of a crash, airbag deployment timing and 300 ms during a crash event is locked in memory, for up to two events. These data, along with "baseline" measures of belt use and basic vehicle information obtainable from VINs, can be broadcast in real time to emergency responders.

2.3. Crash data eligible for analysis

Our analysis relies on the 2002-2012 NASS-CDS and associated EDR datasets. We first merged the NASS-CDS datasets and crash data extracted from EDR reports. Next, we removed crashes where the driver's age and sex were both missing, resulting in 5,736 crashes. We then restricted the crashes to 3,460 frontal crashes defined as those crashes where the principal direction of force was $0°$ to $40°$ and $320°$ to $350°$. We removed 3,198 unusable crashes that were missing crash pulse data or that recorded a constant crash pulse throughout the crash duration. We then removed one crash that involved a vehicle type that was neither a car, SUV, pickup, nor van. Finally, we removed 12 crashes that had missing injury outcomes (including unknown AIS scores) to arrive at 249 crashes eligible for analysis. Figure 1 shows the flowchart of our selection criteria for eligible crashes.

2.4. Variables used in analysis

Our dataset of 249 eligible crashes included seven baseline variables: maximum Abbreviated Injury Scale (AIS) in vehicle (VMAIS), sampling weight, vehicle
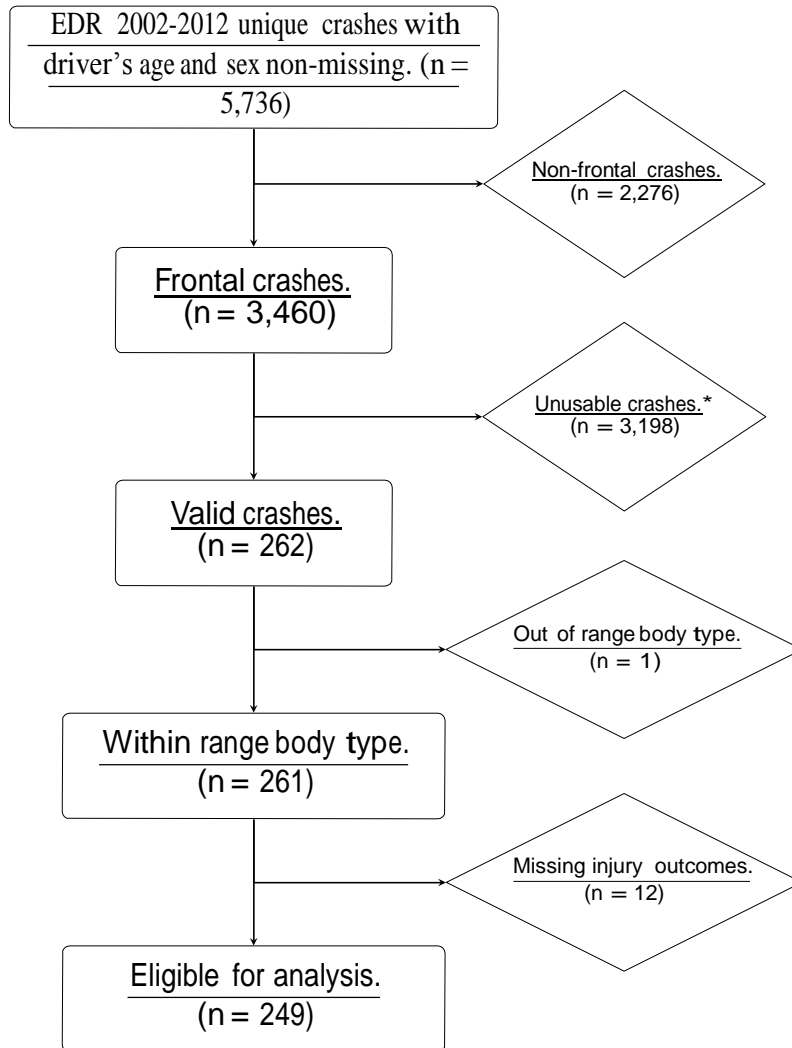
*Includes pre-crashes (time<0), constant crashes (crash pulse constant throughout), and crashes that did not record longitudinal acceleration.

Figure 1: Flow chart of crashes available for analysis.

type, curb weight (kgs), EDR driver belt status (Yes/No/Not reported), EDR front-seat passenger belt status (Yes/No/No front-seat passenger/Not reported), NASS-CDS total delta-v, multiple event indication (Yes/No), and two crash pulse variables: time (milliseconds) and longitudinal (frontal) deceleration (m/s^2). The sampling weight and VMAIS were obtained from NASS-CDS datasets while the seat-belt statuses and crash pulse were obtained from the EDR datasets. Presence of multiple events was derived from EDR reports. Although curb weight and vehicle type were obtained from NASS-CDS datasets, these two variables could, in principle, be derived easily using the VIN.

We modified two of the time-invariant variables as follows. For the vehicle type, we re-categorized the 64 different vehicle types to three new categories: cars (category 0-9), SUVs (category 14-19), and pickups or vans (category 20-33). We grouped pickups and vans together because the number of severe injuries in the pickup category was small. VMAIS was recorded from 0 (no injury) to 6(maximum injury). We considered crashes with VMAIS $\geq$ 3 as having a severe injury in the vehicle. Finally, if two events were reported in the same crash report, the multiple crash indicator was coded as "Yes".

2.5. Comparing demographics of eligible and ineligible crashes

From Figure 1, we observed that nearly 92.8% of our frontal crashes were ineligi- ble. This ineligibility is largely due to the fact that early EDRs usually provided a cruder velocity change that is not easily converted to a continuous acceleration measure; nonetheless, this high percentage of ineligible crashes could create po- tential bias in our prediction model. To verify that our prediction model would still be valid, we compared six NASS-CDS demographic variables from the eli- gible and ineligible crashes. The six variables compared were: driver's age, sex, belt use (from crash investigation), intrusion, front-seat passenger belt use (again from crash investigation), and maximum vehicle injury severity.

We observed no significant difference between eligible and ineligible crashes for the weighted mean and proportion of driver's age, sex, and belt use, front- seat passenger belt use, maximum vehicle injury severity, and sampling weights. (Note that the analysis of the sampling weights was unweighted.) The weighted proportion of driver intrusion was significantly different between eligible and inel- igible crashes, with the excluded cases having somewhat greater rates of intrusion, suggesting that our analytic dataset may contain somewhat less severe crashes than the set of frontal crashes as a whole (Table 1). We would address this issue in the discussion section.

3. Statistical methods

Table 1: Comparing demographics of eligible versus ineligible crashes.

Demographics	Eligible (sd)	Ineligible (sd)	p-value
Driver's Age	39.48 (2.72)	38.93 (1.19)	0.85
Driver's Sex			0.49
Male (%)	47 (8)	53 (3)	
Female (%)	53 (8)	47 (3)	
Driver belt use			0.71
Yes (%)	87 (4)	85 (1)	
No (%)	13 (4)	15 (1)	
Driver intrusion:			0.03
$x < 2$ cm (%)	96 (1)	89 (3)	
3 cm $\leq x < 8$ cm (%)	1 (1)	3 (1)	
8 cm $\leq x < 15$ cm (%)	2 (1)	2 (1)	
15 cm $\leq x < 30$ cm (%)	1 (1)	2 (1)	
30 cm $\leq x < 46$ cm (%)	0 (-)	1 (1)	
46 cm $\leq x < 61$ cm (%)	0 (-)	3 (3)	
61 cm $\leq x$ (%)	0 (-)	0 (-)	
Front-seat passenger belt use:			0.24
Yes (%)	12 (3)	18 (2)	
No front-seat passenger (%)	83 (4)	78 (2)	
No (%)	5 (2)	4 (1)	
Maximum vehicle injury severity:			0.56
AIS3+ (%)	2 (1)	3 (1)	
AIS≤ 2 (%)	98 (1)	97 (1)	
Sampling weight*	352.16 (868.61)	425.59 (1,149.09)	0.23

* Un-weighted

3.1. Functional Data Analysis

Figure 2 provides a typical example of the deceleration data available from an EDR report. To utilize these crash pulses in a regression model, we employed FDA. We considered the deceleration for the ith crash to be a continuous and smooth function of t, $y_i(t)$, observed with error E_{ij} at time t_{ij}, where $j = 1, \ldots, m_i$ for the m_i observations of the ith crash, $i = 1, \ldots, n$. Let y_{ij} be the observed deceleration. FDA attempts to convert $y_{i1}, y_{i2}, \ldots, y_{im_i}$ to the function $y_i(t)$. This means that for a given time t_{ij}, $y_{ij} = y_i(t_{ij}) + E_{ij}$ where $y_i(t_{ij})$ is the value of $y_i(t)$ evaluated at t_{ij} and y_{ij} is estimated using $\hat{y}_{ij} = y_i(t_{ij})$. This conversion can be accomplished using several methods (Ramsay et al., 2009). For this study, we considered four approaches: two methods using unpenalized splines, the "3-millisecond" (3mil) method and "combinatorial" (combi) method;
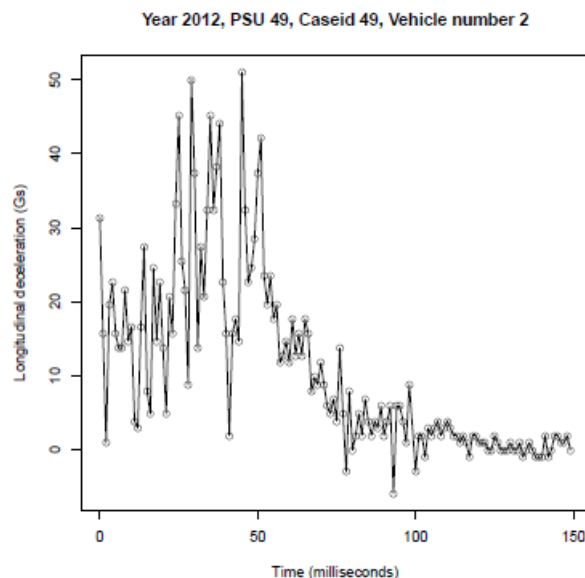
Figure 2: Example of a crash pulse.

and two   methods using penalized splines, the "penalized natural cubic spline" (PCS)
method, and "mixed-model" (MM) method.

3.1.1. 3 Millisecond Method

The first unpenalized method we used was based on basis splines (de Boor, 2001). Let

$$y_{ij} = \sum_{l=1}^{m_i} c_{il}\phi_l(t_{ij}) + \epsilon_{ij}, \epsilon_{ij} \sim N(0, \sigma_i^2)$$

(1)

where $y_i(t_{ij}) = \sum_{l=1}^{m_i} c_{il}\phi_l(t_{ij})$, $\phi_l(t_{ij}) \equiv \phi_{l,d}(t_{ij})$ are basis splines matrices of
degree d, and $\phi_{l,d}(t_{ij})$ is obtained by the recursion relation:

$$\phi_{l,d}(t_{ij}) = \frac{t_{ij} - k_{il}}{k_{i,l+d} - k_{il}}\phi_{l,d-1}(t_{ij}) + \frac{k_{i,l+1+d} - t_{ij}}{k_{i,l+1+d} - k_{i,l+1}}\phi_{l+d,d-1}(t_{ij})$$

(2)

k_{il} are the m_i internal knots of the basis splines, and

$$\phi_{l,0}(t_{ij}) = \begin{cases} 1 \text{ if } \kappa_{il} \leq t_{ij} \leq \kappa_{i,l+1} \\ 0 \text{ otherwise.} \end{cases}$$

In the absence of penalization (i.e., treating the c_{il} as fixed regression param-
eters), we obtained the least squares estimator of y_i as $\hat{y}_i = \Phi_i \hat{b}_i$, where $\hat{b}_i = (\Phi_i^T \Phi_i)^{-1} \Phi_i^T y_i$ for the $m_i \times m_i$ matrix Φ_i made up of the elements $\phi_l(t_{ij})$. This
estimator can provide a very close fit to the data, depending on the degree, d,

and choice of the κ_{il}. We chose a cubic spline $d = 4$ and defined κ_{il} at equally spaced
intervals of k. Our objective in determining k was to choose an interval so that the least squares

estimator $\hat{y}_i$ was smooth and yet followed the original crash pulse closely. Based on a visual inspection of the resulting curves for a variety of crash pulses, we chose to define k = 3 (i.e., knots at every at every 3 milliseconds) to retain as much potential information from the original crash pulse as possible while still smoothing crash pulse to some degree.

3.1.2. Combinatorial Method

As an alternative to defining κ_{il} at fixed intervals, we could fix the number of κ_{il} minus the two boundaries as $m_i \equiv o$, and allow κ_{il} to be freely placed anywhere along the interval $[t_{i2}, t_{i,m_i-1}]$. By defining o as a small value, we induce additional smoothing while hopefully still being able to find strong data signals by allowing the placement of κ_{il} to vary. To maintain a reasonable number of submodels to consider, we defined the potential interval $[t_{i2}, t_{i5}, \ldots, t_{i,m_i-1}]$, corresponding to the knots for the 3mil model, for the o κ_{ij}s to be placed. We set $o = 5$ to allow as much complexity in the fitted curve as possible while still keeping the computation manageable. We considered all possible combinations of κ_{ij} on the reduced interval $[t_{i2}, t_{i5}, \ldots, t_{i,m} \quad 1]$, and calculated $\sum_{j=1}^{m_i}(y_{ij} - \hat{y}_{ij})^2$ combination. We then chose the combination that minimized $\sum_{j=1}^{m_i}(y_{ij} - \hat{y}_{ij})^2$ for the ith crash

3.1.3. Penalized Cubic Spline Method

To reduce the risk of overfitting, we considered a penalized spline approach that downweighted values of $\hat{c}_{il}$ that lead to highly variable values of $\hat{y}_{ij}$ by introducing a penalty in the objective function to be minimized

$$\sum_j \left[y_{ij} - \sum_{l=1}^{m_i} c_{il}\phi_l(t_{ij}) \right]^2 + \lambda_i \int \sum_{i_l}^{m_i} \{c_{il}D^q[\phi_l(t)]\}^2 \, dt$$

$$(3)$$

Where $y_i(t_{ij}) = \sum_{l=1}^{m_i} c_{il}\phi_l(t_{ij})$ and $D^q[g(u)]$ is the qth derivative of function g. The first term in (3) is the objective function that minimizes the least squares estimation of c_i in (1); the second term induces a penalty for large amounts of "acceleration" or "deceleration" when estimating the function $y_i(t)$. Least-squares estimation can be adapted to accommodate (3) by rewriting as

$$\left(y_i - c_i^T\Phi_i\right)^T\left(y_i - c_i^T\Phi_i\right) - \lambda_i c_i^T R c_i \qquad (4)$$

where $R = \int D^q(\phi_L(t)) D^q(\phi_l(t))^T \, dt$ which is minimized with respect to c_i by $\hat{b}_i = (\Phi^T\Phi_i + \lambda_i R)^{-1}\Phi^T y_i$, equivalent to ridge regression for a fixed λ_i. In our analysis, we set $q = 2$ (second derivative) and obtained estimates of λ_i via generalized

cross-validation (Golub et al., 1979), an extension of standard cross-validation procedure that predicts the j^{-th} observation after leaving out the j^{-th} observation in the estimation of c_j.

3.1.4. Mixed Model Method

Noting that (4) is equivalent to the log-likelihood of a random-effects model $y_i = \Phi_i b_i + \epsilon_i$ for $b_i \sim N\left(0, \frac{\sigma_i^2}{\lambda_i} R^{-1}\right)$ and $\epsilon_i \sim N\left(0, \sigma_i^2 I\right)$ [Speed, in discussion of Robinson (1991)], it is possible to estimate the smoothing parameter λ_i using maximum likelihood. For this approach, we utilized the model proposed by Wang (1998), which we have found to be stable and provide appropriate trade-offs between smoothness and local features. It is given by

$$y_{ij} = \beta_{0i} + \beta_{1i} t_{ij} + \sum_{p=1}^{m_i-1} Z_{ijp} b_{pi} + \epsilon_{ij} \quad \epsilon_{ij} \sim N\left(0, \sigma_i^2\right)$$

(5)

Where $y_{ij}(t_{ij}) = \beta_{0i} + \beta_{1i} t_{ij} + \sum_{p=1}^{m_i-1} Z_{ijp} b_{pi}$, $t_i \in [0,1]$, $\left(b_{1i}, b_{2i}, \dots, b_{m_{i-1},i}\right)^T \sim N\left(0, \tau_i^2 I\right)$, $Z_i Z_I^T = \Omega_i$ and Ω_i is an $m_i \times m_i$ matrix with the rows and columns defined as $\Omega_{k,l} = \int_0^1 (t_k - \mu) + (t_l - \mu) + dt = \frac{1}{2}[min(t_k, t_l)]^2 \, max(t_k, t_l) - \frac{1}{6}[min(t_k, t_l)]^3$

The tuning parameter $\hat{\lambda}_i$ is estimated by $\hat{\lambda}_i = \frac{\hat{\sigma}_i^2}{m_i \hat{\tau}_i^2}$. Note that t_{ij} is a transformation from the original time $\tilde{t}_{ij}$, given by $t_{ij} = \frac{\tilde{t}_{ij} - \min_j(\tilde{t}_{ij})}{\max_j(\tilde{t}_{ij}) - \min_j(\tilde{t}_{ij})}$ so that $t_{i1} = 0 < t_{i2} < \cdots < t_{im_i} = 1$.

(Because $y_i(t_{ij}) = y_i(\tilde{t}_{ij})$, we do not need to convert $y_i(t_{ij})$ back to the $\tilde{t}_{ij}$ scale.)

We obtained Z_i by estimating the Cholesky decomposition of Ω_i using Smith's(1995) method. We then estimated y_{ij} as

$$\hat{y}_{ij} = \hat{\beta}_{0i} + \hat{\beta}_{1i} t_{ij} + \sum_{p=1}^{m_i-1} Z_{ijp} \hat{b}_{pi}$$

(6)

where, using standard mixed-model results (Harville, 1976), we have $\hat{\beta}_i = \left(X_i^T \hat{V}_i^{-1} X_i\right)^{-1} X_i^T \hat{V}_i^{-1} y_i$, $\hat{b}_i = \tau_I^2 Z_i^T \hat{V}_i^{-1}\left(y_i - X_i \hat{\beta}_i\right)$ for $\hat{V}_i = \tau_i^2 Z_i^T Z_i + \hat{\sigma}_i^2 I_i$, where τ_i^2 and $\hat{\sigma}_i^2$ are estimated using restricted maximum likelihood (REML).

3.2. Slope of $y_i(t)$ at t_{ij}

In addition to obtaining $y_i(t)$, we were interested in calculating the slope of $y_i(t)$ at t_{ij}. We hypothesized that vigorous fluctuations in the crash pulse may be an important predictor for severe injury outcomes. For the basis spline methods in 3.1.1-3.1.3, we computed the derivative of the least squares estimator as $\frac{d}{dt}\Phi_i \hat{b}_i$ where

$$\frac{d}{dt_{ij}} \phi_{l,d}(t_{ij}) = (l-1)\left(\frac{\phi_{l,d-1}(t_{ij})}{k_{i,l+d-1} - k_{il}} + \frac{\phi_{+1l,d-1}(t_{ij})}{k_{i,l+d} - k_{i,l+1}}\right)$$

(7)

For the slope of the MM method in 3.1.4, the first derivative of the equation 6 yields

$$\hat{\beta}_{1i} + \sum_{p=1}^{m_i-1} \bar{Z}_{ijp}\hat{b}_p$$

(8)

Where $\bar{Z}_{ijp}$ was the (j, p) element of $\frac{d}{dt_{ij}}Z_i$. To obtain $\bar{Z}_{ijp}$, we took the first derivative of the Ω_i matrix with respect to t_{ij}. Then using the fact that $\frac{d}{dt_{ij}}\Omega_i = \left(\frac{d}{dt_{ij}}Z_i\right)Z_i^T + \left(\frac{d}{dt_{ij}}Z_i\right)^T$ we derived $\bar{Z}_{ijp}$ from $\frac{d}{dt_{ij}}\Omega_i$. Details are provided in Appendix A .

3.3. Extracting summary measures from the crash pulse

Typically in FDA, once $y_i(t_{ij})$ has been estimated, researchers would use the coefficients c_{il} as the variables in the second-stage regression model. However, in this study we faced two difficulties. First, the duration of each crash was different. Second, the time stamp interval was different for each crash. These two difficulties implied that defining a single set of linear combination of functions for all crashes and subsequently using their coefficients would be inappropriate since each crash should have its own linear combination of basis functions $\phi_l(t_{ij})$ to define $y_i(t_{ij})$. To overcome these difficulties, we extracted four summary measures from the $y_i(t_{ij})$ of the four methods separately.

The first summary we extracted was $\hat{G}$ defined as $\hat{G} = \int_{ti}|y_i(t_i)|\,dt_i$ absolute integral of $y_i(t_i)$ over the duration t_i. We took the absolutes because we did not want $y_i(t_i)$ to cancel each other during integration. Because closed form solutions were not available, we used numerical methods (Simpson's integral) to approximate $\hat{G}_i$. $\hat{G}_i$ sould approximate the "instan taneous deceleration" measure delta-v, a good predictor of severe injury risk. Indeed $\hat{G}_i$ was correlated with the delta-v given in NASS-CDS (Kendall's correlation of 0.44, p-value<0.001)

Thus $\hat{G}_i$ should be an important predictor of injury risk. Figure 3 provides an illustration of the differences between crashes with a high $\hat{G}_i$ and a low $\hat{G}_i$.

The second summary measure that we extracted was $\hat{g}_i = \int_{ti}|y'_i(t_i)|\,dt_i$ the "absolute integral" of all the estimated slope of $y_i(t_i)$ over the duration t_i. This definition of $\hat{g}_i$ was designed to capture the extent to which the crash pulse fluc- tuated. The absolute signs prevented positive slopes from canceling out negative slopes, ensuring that we did not underestimate the fluctuations in the crash pulse. Again, we used numerical methods (Simpson's integral) to approximate $\hat{g}_i$. Fig- ure 4 illustrates the difference between a crash with higher $\hat{g}_i$ vs. a crash with a lower $\hat{g}_i$.

The third summary measure we extracted was σ_i^2 the variance of the residual. Residual variance provided an additional measure of the degree of variability in
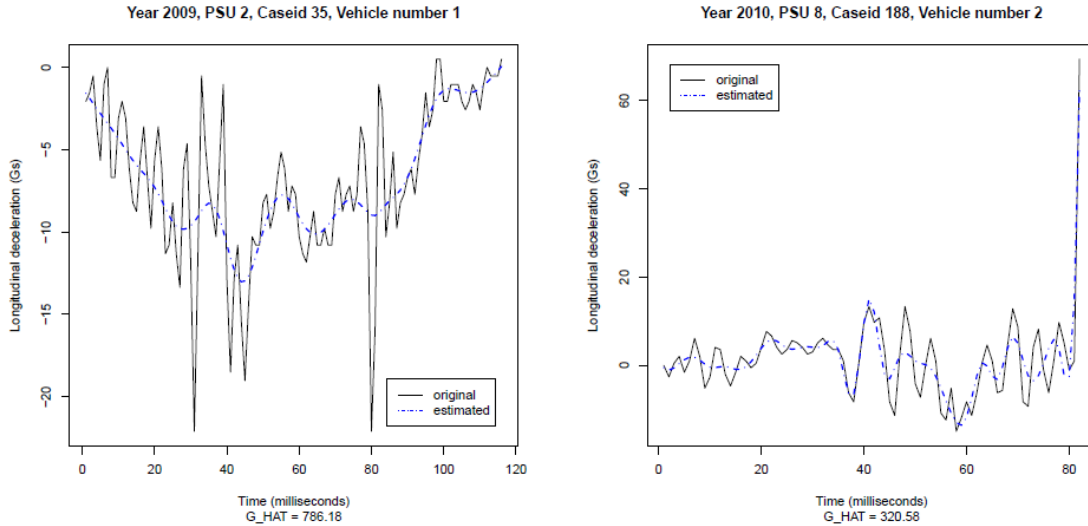
Figure 3: Original and estimated crash pulses for crashes with high and low $\widehat{G}$.

the crash pulse not explained by $\widehat{G}_i$ and $\widehat{\widehat{g}}_i$.   Figure 5 illustrates this clearly. We estimated
$\widehat{\sigma}_i^2$; for 3mil , combi , and PCS method by taking the sample variance of the residuals , while for
the MM method , we obtained the REML estimate of $\widehat{\sigma}_i^2$.

Lastly , we extracted a summary measure called time to 0.25Gs (*tt*025). Tt025 was defined as
the time the crash pulse took to eventually return to $\pm$ 0.25 G_s.
This gives a measure of how long the force from a crash affects the vehicle. We computed tt025 by
starting from the last crash pulse recorded for a given crash and moving backwards to the first
deceleration value recorded , searching for the deceleration value that first exceeded $\pm 0.25 G_s$,
$t_{0.25 G_s}$.

We obtained *tt*025 by taking the difference of $t_{0.25 G_s}$ and the time at the start of the crash pulse.
Figure 6 shows graphs of a crash pulse with a high *tt*025 vs. low *tt*025. A crash pulse with high *tt*025
tended to have their crash activity "stretched out" while crash pulses with low tt025 had their
crash activity "compressed" . Hence , we expect a crash pulse that extends over a longer time (i.e.,
high *tt*025) should result in less severe injury than one that occurs in a shorter time frame.

3.4. Predicting risk of severe injury in motor vehicle crashes

We computed our prediction model by first merging each set of summary measures with the
outcome, VMAIS and the other time-invariant variables (see subsection 2.4) to form four datasets.
Then, we expanded these four sets to a further sixteen by taking different combinations of the
following: 1. logarithm transform or no logarithm transform on the summary measures and 2.
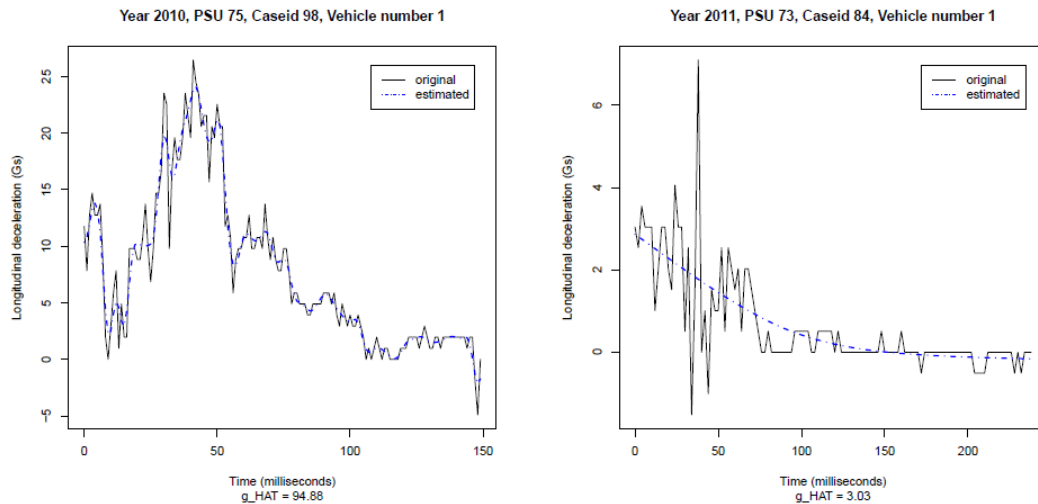scaling i.e. $\frac{\widehat{G}_i}{tt025_i}$

Figure 4: Original and estimated crash pulses for crashes with high and low $\hat{g}$;.

or no scaling of $\widehat{G}_t$. This was done to investigate whether logarithm transformation of the summary measures or whether the scaling of $\hat{G}$ improves prediction performance. As presentation of coefficients and results from sixteen different models would be cumbersome, we used the highest leave-one-out cross-validation (CV) weighted area under the curve (AUC) of the receiver operating characteristic (ROC) curve to determine which model to present. The weighted AUC was computed based on the results of a weighted logistic regression with VMAIS as the outcome and summary measures and time-invariant (baseline) variables as the covariates.

To investigate the performance of our chosen model, henceforth called the `deceleration' model, we first compared the weighted AUC of the `deceleration' model with a `baseline' model. The `baseline' model was one which contained only time-invariant EDR variables. This was to determine if adding deceleration information to our model improved prediction performance. We also measured the variability of this difference by running 1,000 bootstrap samples of the eligible crashes and calculating the weighted AUC for each of them. We then took the difference of their weighted AUC, sorted this calculated difference, and took the 25th and 975th values as the lower and upper bounds for the 95% bootstrap confidence intervals of the weighted AUC difference.

To complement the weighted AUC results, we plotted the weighted ROC and precision-recall (PR) curves (Davis and Goadrich, 2006) and constructed a false discovery rate (FDR) table for both deceleration and baseline models.
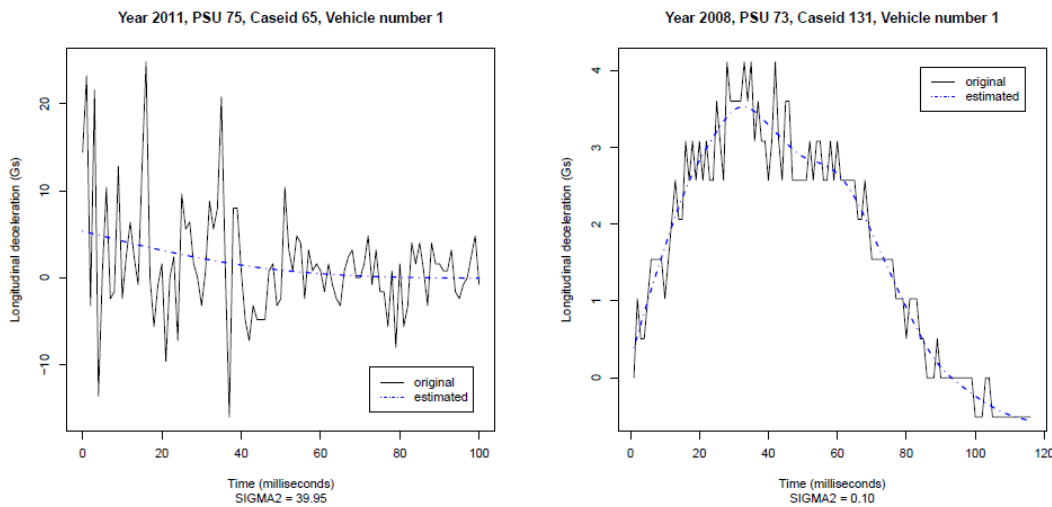
Figure 5: Original and estimated crash pulses for crashes with high and low $\hat{\sigma}_i^2$

PR curves relate the proportion of true-positives out of all positive results a model would produce (y-axis) to any level of true-positive rate (the proportion of true-positives among all cases, x-axis). More speciflcally, precision (yaxis) is defined as $\dfrac{True\ positive}{True\ positive + False\ positive}$, while recall (x-axis) is de_ned as $\dfrac{True\ positive}{True\ positive + False\ negative}$. PR curves share many properties with the ROC, but provide a focus on the trade-o_ between the fraction of AIS 3+ injuries captured at a given prediction level cutpoint versus the fraction of positive cases that will actually have an AIS 3+ injury at that level. These are particularly important measures in this setting, where the cost of missing severe injuries must be balanced against resources assigned unnecessarily to non-severe injuries. PR curves lying in the top-right imply good performance. We plotted weighted PR curves by first using the sampling weight of each crash as the number of either severe injury outcome or non-severe injury outcome a crash represents. Then, we smoothed these points using the locally weighted scatter-plot smoothing (loess) to curves the PR curves.

The FDR table we computed provides similar information as the PR curves.The difference was, instead of looking at the proportion of true-positives out of all positive results, our FDR table gave the proportion of false-positives out of all positive results. We presented this proportion in the right column of our FDR table with the predicted probability cutoff points in the left column, and the proportion of severe injuries categorized as positives in the middle column. For this study, we used predicted probability cutoff points of 1-10%, 15%, 20%, 25%,
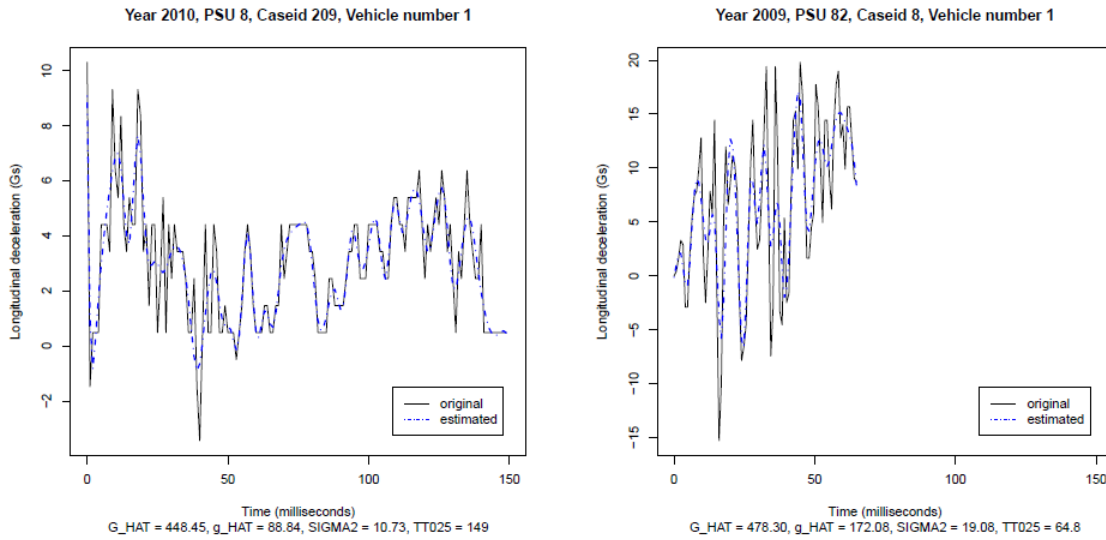
Figure 6: Original and estimated crash pulses for crashes with high and low *tt*025.and 30%.

Finally, we ran a leave-one-out CV on the deceleration and baseline model to investigate how well they will perform given new data. Similarly, we looked at their weighted ROC and PR curves, and the FDR tables using the same predicted probability cutoff points mentioned in the previous paragraph. We completed all analyses using *R*x64 3.0.3.

4. Results

To get a visual sense of how the four different FDA methods estimated the original crash pulse, we present four selected crashes with their estimated trajectories, $\hat{y}_{ij}$ in Figure 7. We observed that the 3mil and PCS methods tend to follow the original crash pulse closely while the combi method smooths almost all crash pulses. The MM method alternates between following the crash pulse when there are strong signals in the data (see the top two plots in Figure 7), or smoothing the pulses when they are \noisier" (see the bottom two plots in Figure 7).

From the results of the leave-one-out CV, we found that the model with summary measures estimated using penalized method, no logarithm transform on the summary measures, and no scaling produced the highest weight leave-one-out CV AUC, 0.75. Using this model, we next focus on the descriptive statistics of our variables (Table 2). We found that the (weighted) mean of $\hat{G}$, $\hat{g}$, $\hat{\sigma}^2$, and curb weight in severe injury crashes were higher compared to non-severe injury crashes (847 vs. 499, 87 vs. 69, 59 vs. 12, and 1,694 vs. 1,681 respectively).

However, the mean of tt025 for severe injury crashes was slightly lower compared to non-severe injury crashes (110 vs. 112). The proportion of cars and pickups or vans were higher in severe injury crashes compared to non-severe injury crashes (45% vs. 42% and 26% vs. 24% respectively) while SUVs had a lower proportion involved in severe injury crashes (29% vs. 34%). More than half (59%) of the eligible crashes did not have seat-belt statuses of the driver and front-passenger reported in their EDR reports. We observed a lower proportion of the known- belted occupants involved in severe injury crashes (12% vs. 38% and 6% vs. 7% for driver and front-seat passenger belt status respectively). The proportion of unbelted drivers involved in severe injury crashes was higher than in non-severe injury crashes (21% vs. 3%), while we found a lower rate on unbelted front-seat passengers in severe injury crashes vs. non-severe injury crashes (4% vs. 17%). Finally, we observed a higher proportion multiple crash recorded of severe injury crashes (18% vs. 6%).

Table 2: Descriptive statistics: mean (standard error in parenthesis)

Covariate (sd)	Severe Injury (n=27)	Non-severe injury (n=222)	Total
$\hat{G}$	846.86 (91.78)	498.96 (25.66)	506.05 (25.33)
$\hat{g}$	87.02 (27.64)	69.03 (8.48)	69.40 (8.34)
$\hat{\sigma}^2$	58.97 (23.17)	12.26 (2.14)	13.22 (2.20)
T $t025$	110.44 (6.16)	112.12 (3.51)	112.09 (3.44)
Body type (%):			
Car	45 (16)	42 (8)	42 (8)
SUV	29 (19)	34 (10)	34 (10)
Pickup or Van	26 (14)	24 (5)	24 (5)
Curb weight (*kgs*)	1,693.70 (142.97)	1,680.80 (33.09)	1,681.00 (32.56)
Driver belt use (%):			
Belted	12 (6)	38 (8)	37 (7)
Not reported	67 (14)	59 (8)	59 (8)
Not belted	21 (14)	3 (1)	4 (1)
Front-seat passenger belt use:			
Belted	6 (4)	7 (2)	6 (2)
No front-seat passenger	23 (14)	17 (5)	18 (5)
Not reported	67 (14)	59 (8)	59 (8)
Not belted	4 (3)	17 (5)	17 (5)
Multiple crashes? (%):			
Yes	18 (14)	6 (2)	6 (2)
No	82 (14)	94 (2)	94 (2)

Although estimated delta-v from databases like NASS-CDS underestimate actual delta-v, we still wanted to check whether our summary measures had good linear correlation to NASS-CDS delta-v because many studies have shown that NASS-CDS delta-v is a good predictor of severe injury risk. We found that all summary measures had a weak to moderately positive correlation with NASS- CDS delta-v with $\hat{G}$ having the highest linear correlation 0.44 while $tt025$ had the weakest, 0.05. $\hat{G}$ was moderately but significantly correlated with $\hat{g}$, $\hat{\sigma}^2$, and $tt025$ respectively (0.31, p-value< 0.001; 0.34, p-value< 0.001; 0.18, p-value< 0.001).

Table 3: Correlation table of $\hat{G}_i$, $\hat{g}_i$, $\hat{\sigma}_i^2$ and $tt025$. (*p*-value)

	$\hat{G}$	$\hat{g}$	$\hat{\sigma}^2$	$tt025$
CDS delta-v	0.44 (< 0.001)	0.23 (< 0.001)	0.26 (< 0.001)	0.05 (0.40)
$\hat{G}$	1	0.31 (< 0.001)	0.34 (< 0.001)	0.18 (< 0.001)
$\hat{g}$		1	0.17 (< 0.001)	-0.05 (0.36)
$\hat{\sigma}^2$			1	-0.12 (0.02)

4.1. Prediction of Injury Risk

The coefficients of our deceleration model are given in Table 4. $\hat{G}$, $\hat{\sigma}^2$, tt0.25 belted driver, and multiple crash indicator were significant, with a 10 m/s2 in- crease in $\hat{G}$ associated with a 4% increase in odds of AIS 3+ injury (95% CI 2%,6%), a 1-g increase in the error of the estimation method associated with a 2% increase in the odds of injury (95% CI 1%, 3%), a 1 millisecond increase in the duration of the crash associated with a 3% decrease in odds of AIS 3+ injury (95% CI 1%, 5%), presence of a belted driver associated with a 97% decrease in odds of AIS 3+ injury (95% CI 82%, 99%), and presence of multiple crashes associated with a 42-fold increase in the odds of AIS 3+ injury (95% CI 2-fold,780-fold). Surprisingly, the coefficient for belted front-seat passenger was positive (1.49) although not significant (*p*-value=0.43). The weighted AUC for our decel- eration model was 0.93. We also computed the weighted AUC of the deceleration model with the summary measures centered and the square of these centered summary measures added. The weighted AUC was similarly 0.93 suggesting no improvement when we include squared terms of our summary measures.

For the baseline model, only the coefficients of the belted driver and multiple crashes were significant (-3.23 and 2.04 respectively with *p*-values of < 0.001 and 0.04). The belted front-seat passenger coefficient was still positive, not significant (p-value=0.46), and the magnitude of the estimate was smaller than the deceler- ation model. The weighted AUC of our baseline model was 0.78, 0.15 less than the deceleration model. The 95% bootstrap confidence interval of the weighted AUC difference between the deceleration model and the baseline model was (0.04,

Table 4: Coefficient Estimates for deceleration, baseline, and original model with weighted AUC.

Parameter/Statistic	Deceleration Estimate (Conf. Int)		Baseline Estimate (Conf. Int)	
$\hat{G}$ (10 m/s^2)	0.04	(0.02, 0.06)***		
$\hat{g}$ (10 m/s^3)	-0.03	(-0.10, 0.06)		
$\hat{\sigma}^2$	0.02	(0.01, 0.03)**		
tt_{025}	-0.03	(-0.05, -0.01)*		
Driver belt use				
Belted	-3.35	(-5.00, -1.71)***	-3.23	(-4.79, -1.68)***
Not reported	1.00	(-1.26, 3.27)	0.25	(-1.78, 2.29)
Not belted				
Front-seat passenger belt use[†]				
Belted	1.49	(-2.19, 5.16)	1.01	(-1.66, 3.69)
No front passenger	1.58	(-0.54, 3.69)	1.58	(-0.34, 3.50)
Not belted				
Curb weight (per 1,000 kg)	0.69	(-1.06, 2.43)	-0.73	(-2.62, 1.17)
Body type				
Car	0.87	(-1.35, 3.10)	0.96	(-1.15, 3.07)
Pickup or Van	-0.80	(-3.34, 1.74)	0.31	(-1.49, 2.12)
SUV				
Multiple crashes?				
Yes	3.74	(0.81, 6.66)*	2.04	(0.11, 3.96)*
No				
AUC		0.93		0.78
ROC Deceleration - ROC Baseline	0.15 (0.04, 0.23)			

† Crashes not reporting front-seat passenger belt status were the same with driver belt status.
* $0.01 \leq p < 0.05$; ** $0.001 \leq p < 0.01$; *** $p < 0.001$.

0.23).

We next checked each observed crash for influential points using the adjusted Cook's D for survey data. We found two influential observations where their covariate pattern suggested a non-severe injury crash but were in fact severe crashes. We removed them as part of a sensitivity analysis and found the resulting coefficients, 95% confidence intervals, and AUC similar to those presented in Table 4. Hence, we choose to present the model without the removal of the two influential points.

The ROC and PR curves suggested that our deceleration model performed better than the baseline model (Figure 8). We observed that the weighted ROC for the baseline model was to the right of the ROC of the deceleration model for false positive rates between 0 and 0.8 implying better performance for the deceleration model. Similarly, we found that the loess PR curve of the baseline model was on the left of the deceleration model for all recall values indicating lower prediction performance of the baseline model.

Similar results were found in the FDR table (Table 5). The deceleration model had a large decrease in capture rate at a predicted probability cutoff of 6%, at

that cutoff, we would capture 7 in 10 of the AIS 3+ injuries, with about one in four crashes actually having an AIS 3+ injury. To capture a similar fraction of AIS 3+ crashes in the baseline model, we would have to use a cutpoint of 2%, at which point fewer than 1 in 20 crashes would have an AIS 3+ injury.

Figure 9 provides the "leave-one-out" cross-validation results for the AUC and PR plots, and Table 6 provides the equivalent for the FDR table. These results suggest less gain for the deceleration model relative to the baseline model, although a cutoff of 5% risk still captures one-third of the AIS 3+ injuries with one in ten being AIS 3+; capturing an equivalent fraction of AIS 3+ injuries using baseline data would yield a "hit rate" of less than one AIS 3+ crash in 50.

Table 5: Fraction of AIS 3+ Injury Crashes Captured and Associated Fraction of Crash that are not AIS 3+ [False Discovery Rate (FDR)] at various thresholds of injury risk cutpoints.

Cutpoints (%)	Deceleration model		Baseline model	
	Captured (%)	FDR (%)	Captured (%)	FDR (%)
1	95	94	88	97
2	75	88	54	96
3	75	83	45	94
4	75	80	40	91
5	71	77	25	87
6	71	74	17	86
7	47	79	17	83
8	46	77	17	81
9	46	76	17	81
10	46	71	17	81
15	43	66	17	68
20	42	63	17	48
25	42	58	17	48
30	39	52	2	83

5. Discussion

In this study, using only information available in EDR reports, we successfully developed a new severe injury risk prediction model able to estimate the proba- bility of a severe injury in motor vehicle crashes near real-time. Although such real-time prediction models are not new in trauma literature, our sole use of data from EDR guaranteed that our model reflects the actual situation when EDR data is received by EMS from a dispatcher. Another strength of our model lay in the use of a novel variable – the crash pulse – as the main variable in our re- gression model. To use this crash pulse, we employed a two-stage approach, with

Table 6: Fraction of AIS 3+ Injury Crashes captured and Fraction of Crashes captured that are not AIS 3+ [False Discovery Rate (FDR)] at various thresholds of cross-validation injury risk cutpoints.

Cutpoints (%)	Deceleration model		Baseline model	
	Captured (%)	FDR (%)	Captured (%)	FDR (%)
1	37	98	46	99
2	34	96	32	99
3	34	93	29	97
4	34	91	9	99
5	32	90	2	99
6	28	90	2	98
7	28	89	2	98
8	28	89	2	98
9	28	89	2	97
10	28	88	2	97
15	28	79	2	94
20	27	75	2	93
25	19	79	2	91
30	18	79	2	87

summary measures from the first-stage FDA model, as inputs for the second-stage weighted logistic regression model. We were not aware of any such applications of FDA to crash pulses for predicting severe injury risk in trauma literature. By extracting summary measures from the estimated deceleration, we were able to avoid the problem of trying to find a common set of linear combination of functions for trajectories which have different duration and time stamp interval. With the use of FDA to extract information from trajectories gaining popularity, we believe more researchers will face similar situations where application of usual FDA procedures would be inappropriate.

Our deceleration model performed well when predicting the population of eligible crashes compared to the baseline model. In addition, we found that our deceleration model was able to give predicted probability cut-offs where the capture rate of true severe injury crashes was high, while incurring a reasonable amount of false discoveries. This could be seen from Figures 8 and 9 and Tables

5 and 6. We found less optimistic results using a leave-one-out cross-validation procedure. While suggesting a similar optimal cut-off point of 6-10% predicted risk for classifying crashes, our cross-validation analysis suggested that such a cutpoint would capture slightly less than 3 in 10 of AIS 3+ crashes, but yield a false discovery rate of approximately 9 in 10.

Not surprisingly, we found that $\widehat{G}$, driver belt use, and presence of multiple crashes increased the risk of a severe injury crash significantly. However, our finding that $\hat{\sigma}^2$ and $tt025$ were significantly associated with severe injury in motor vehicle crashes was not entirely expected and could contribute to the further understanding of injury in motor vehicle crashes. A significant and positive $\hat{\sigma}^2$ value implies that differences between the estimated deceleration and the original crash pulse are associated with higher risk of severe injury. This means that rapid changes in deceleration unexplained by the estimated pulse signal causes additional injury above and beyond what we would expect if there were no rapid changes.

The second significant coefficient that we felt could further the understand- ing of severe injury crashes was $tt025$. A negative and significantly coefficient implies that as the duration of the force of the crash decreases, the risk of severe injury increases. This may seem counter-intuitive but with careful thought will be clear. Observe Figure 6. With similar force ($\widehat{G}$ of 448 [left] and 478 [right]), the deceleration duration on the left took 149 milliseconds while the deceleration on the right took 65 milliseconds. As a result, the crash pulse on the right was 'squashed' and it had to 'deliver' the similar amount of force but within a shorter duration. This most likely leads to increased risk in having a severe injury. With the understanding that longer effective duration of the force of the crash would reduce the risk of severe injury in a crash, motor vehicles which prolong the effective duration of the crash could be developed.

Our success in developing a fairly good near real-time prediction model was limited by three issues: a high proportion of unusable crashes, non-reporting of seat-belt statuses, and inflated sampling weights. In our study, we found that 3,198 out of 3,460 (about 92.4%) crashes were unusable. This was because different manufacturers present different variables in their EDR report and most reports contained velocity change instead of deceleration because earlier EDRs usually provide cruder velocity change that is not easily converted to a continuous acceleration measure. As newer vehicles equipped with newer EDRs enter the car fleet and EDR penetration increasing, we expect this ineligibility issue to subside in the future.

While we found only one statistically significant difference between the usable and unusable crashes (a tendency for usable crashes to be somewhat less severe), we did not present a post-stratified model (a model that uses weights such that the rates of intrusion between eligible and ineligible crashes would be similar) because first, although post-stratification balanced the rates of intrusion between eligible and ineligible crashes, the resulting coefficients, 95% confidence intervals, and AUC were similar to the results reported in Table 4. Second, we noted

that the intrusion rate of $46\,\mathrm{cm} \leq \mathrm{x} < 61\,\mathrm{cm}$ in ineligible crashes was 3%, an

exceedingly high proportion for such severe injury crashes in the population. We suspect this is an artificial inflation of the rates due to highly variable survey weights being allocated to these ineligible crashes, an issue we will return to in subsequent paragraphs.

The second issue we faced was that many EDRs did not report driver and front-seat passenger belt status. From Table 2, we can see that nearly 59% (weighted) of our eligible crashes did not have their seat-belt status reported. Unlike deceleration, we were unable to observe any pattern or determine the rea- son why such a large proportion of EDR reports did not report seat-belt statuses. The consequence of this was likely the biasing of seat-belt status coefficients to- wards the null. The assumption we made here was that EDR reports were equally likely to not record seat-belt statuses of belted or not belted occupants. This im- plied that the belted effect for drivers could have been more protective but for a belted front-seat passenger, the contradictory effect of a belted front-seat pas- senger being a risk for severe injury during a crash would have been increased. A second consequence of EDR reports not recording seat-belt status was likely the reduction in prediction performance of cross-validation.

Finally, the third issue we faced was highly variable sampling weights. Since the NASS-CDS sampling weights reflected how likely a crash was to be selected, we would expect crashes resulting in a severe injury or driver intrusion > 3cm to have smaller sampling weights since severe crashes were oversampled. How- ever, we found that four of our eligible severe injury crashes (about 15%) had large weights of more than 100. Similarly, we found large sampling weights for driver intrusions of > 3cm for ineligible crashes. These inflated sampling weights affected the precision of our model's coefficients by increasing their variability.

The possible remedies for these three issues we faced were limited. We could have used some imputation method to recover the crash pulses in unusable crashes or perform sensitivity analyses to determine the effects of non-reporting seat-belt statuses and inflated sampling weights on our model. However, imputation of crash pulses in unusable crashes did not seem appropriate because the proportion of crashes that did not report crash pulse was high, about 92%, and crash pulses were usually very variable, making imputation difficult and likely inaccurate. Sensitivity analysis of seat-belt status may have given us a better sense of the variability in our model but we felt that sensitivity analysis would not be helpful because we did not have a good idea of the mechanism that generated the non- reporting of seat-belt statuses in EDR reports.

From the issues presented, we can see prediction of severe injury crashes in motor vehicles is still a challenging field. The model we have developed here should not be considered as the final model to be used for all motor vehicle crashes, but we believe it can be considered a promising seminal model for real- time severe injury crash prediction. To develop a comprehensive model we suggest the following steps:

1. Develop a separate prediction model that uses information from the ad- justed velocity change instead of deceleration. This step will be essential if some EDRs do not report deceleration.

2. Add the lateral, vertical, and rollover of crash pulse and adjusted velocity components into the model described in 1.

3. Develop a joint model to compute the FDA estimates and logistic regression model in a single step. Two-stage regression models are known to suffer from bias and loss of efficiency, due to the error in the estimation procedure not be accounting for in the second stage of estimation. However, joint model estimation is currently a time-demanding procedure, one that is un- likely to be carried out in practice given the time constraints involved in responding to a crash event. Methods that can utilize aspects of a joint esti- mation procedure outside of a full model-fitting would have great potential value in this setting and are the target of future research.

References

[1] Association for the Advancement of Automotive Medicine (1990). *The Abbrevi- ated Injury Scale 1990 Revision.* Association for the Advancement of Automo- tive Medicine, Des Plaines, IL.

[2] Augenstein, J., Perdeck, E., Stratton, J., Digges, K., and Bahouth, G. (2003).

[3] Characteristics of crashes that increase the risk of serious injuries. *Annual*

[4] *Proceedings: Association for the Advancement of Automotive Medicine* **47,**

[5] 561–570.

[6] Binder, D. (1983). On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review* **51,** 279–292.

[7] Clark, D., Winchell, R., and Betensky, R. (2013). Estimating the effect of emer- gency care on early survival after traffic crashes. *Accident Analysis and Pre- vention* **60,** 141–147.

[8] Davis, J. and Goadrich, M. (2006). The Relationship Between Precision-Recall and ROC Curves. *Proceedings of the 23rd International Conference on Machine Learning* pages 233–240.

[9] de Boor, C. (2001). *A Practical Guide to Splines*. Springer-Verlag, New York, Rev. edition.

[10] Golub, G., Health, M., and Wahba, G. (1979). Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. Technometrics 21, 215–223.

[11] Hampton, C. and Gabler, H. (2010). Evaluation of the Accuracy of NASS/CDS Delta-V Estimates from the Enhanced WinSmash Algorithm. Annals of Ad- vances in Automotive Medicine 54, 241–252.

[12] Harville, D. (1976). Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects. The Annals of Statistics 4, 384–395.

[13] Kononen, D. W., Flannagan, C. A. C., and Wang, S. C. (2011). Identification and validation of a logistic regression model for predicting serious injuries associated with motor vehicle crashes. Accident Analysis and Prevention 43, 112–122.

[14] Mizuno, K., Itakura, T., Hirabayashi, S., Tanaka, E., and Ito, D. (2014). Opti- mization of Vehicle Deceleration to Reduce Occupant Injury Risks in Frontal Impact. Traffic Injury Prevention 15, 48–55.

[15] Murphy, S., Xu, J., and Kochanek, K. (2013). Deaths: final data for 2010. National Vital Statistics Reports 61, 1–118.

[16] National Highway and Traffic Safety Administration (2008). National Automotive Sampling System (NASS) Crashworthiness Data System (CDS) User's Manual 2008 File. National Center for Statistics and Analysis, NHTSA, U.S. Depart- ment of Transportation, Washington, DC.

[17] Newgard, C., Schmickler, R., Hedges, J., Trickett, J., Davis, D., Bulger, E., Aufderheide, T., Minei, J., Hata, J., Gubler, K., Brown, T., Yelle, J., Bardar- son, B., G., N., and Investigators, R. O. C. (2010). Emergency medical services intervals and survival in trauma: assessment of the "golden hour" in a North American prospective cohort. Annals of Emergency Medicine 55, 235–246.

[18] Ramsay, J., Hooker, G., and Graves, S. (2009). Functional Data Analysis with R and MATLAB. Use R. Springer Publishing Company, New York.

[19] Rao, J. and Scott, A. (1981). The analysis of categorical data from complex sample surveys: Chi-squared tests for goodness of fit and independence in two- way tables. J. Amer. Statist. Assoc. 76, 221–230.

[20] Robinson, G. (1991). That BLUP is a Good Thing: the estimation of random effects (with discussion). Statistical Science 6, 15–51.

[21] Smith, S. (1995). Differentiation of the Cholesky Algorithm. J. Comput. Graph. Stat. 4, 134–147.

[22] Wang, Y. (1998). Smoothing spline models with correlated random errors. *J. Amer. Statist. Assoc.* **93,** 341–348.

Yaoyuan V. Tan
Department of Biostatistics
School of Public Health
University of Michigan
1415 Washington Heights
Ann Arbor, MI 48109-2029

Carol A.C. Flannagan University
of Michigan Transportation
Research Institute
2901 Baxter Road
Ann Arbor, MI 48109-2150

Jonathan D. Rupp University of
Michigan Transportation Research
Institute
2901 Baxter Road
Ann Arbor, MI 48109-2150

Michael R. Elliott
Department of Biostatistics
School of Public Health
University of Michigan
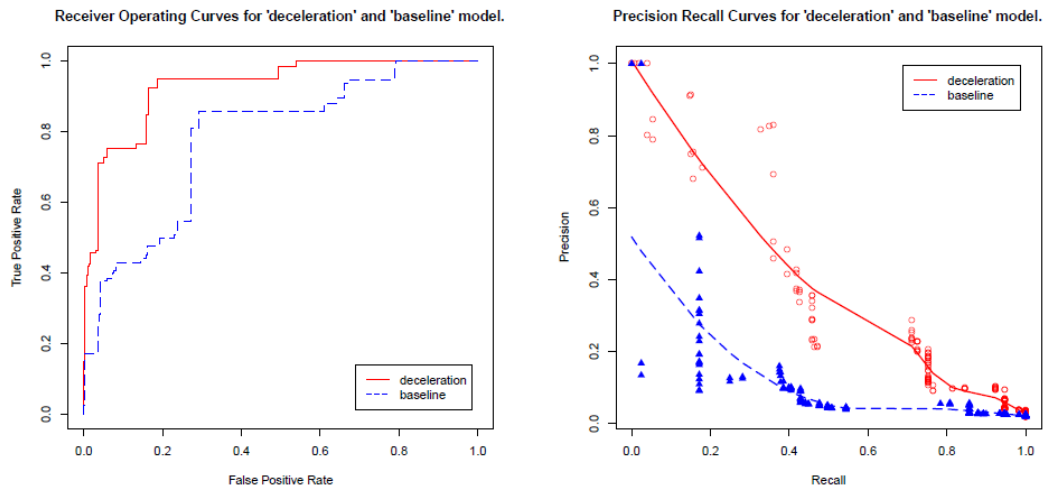1415 Washington Heights
Ann Arbor, MI 48109-2029

Figure 8: Weighted Receiver Operating Curves (ROC) and Precision-Recall (PR) Curves for the deceleration and baseline model.
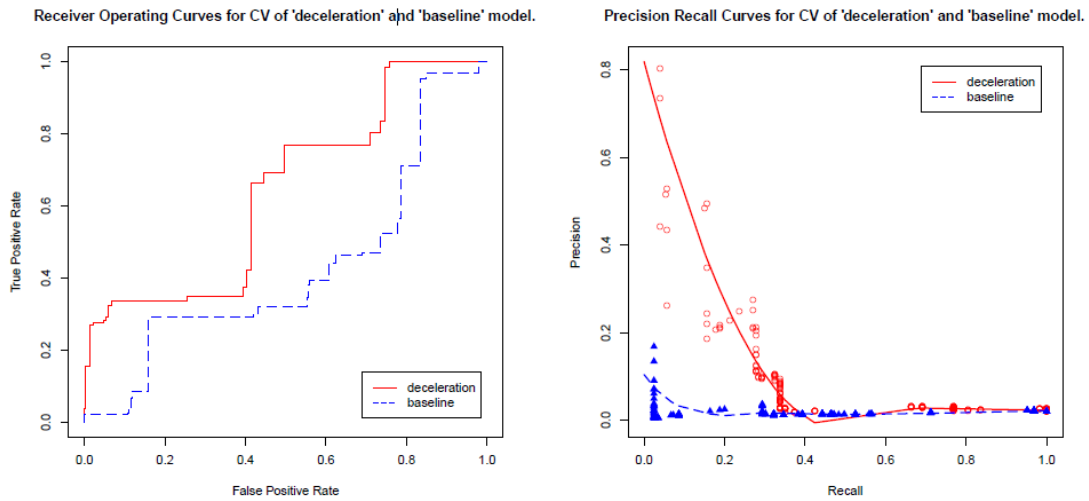

Figure 9: Cross-validation weighted Reciever Operating Curves (ROC) and Precision-Recall (PR) Curves for the deceleration and baseline model.