PREDICTING SEVERE INJURY IN MOTOR VEHICLE CRASHES

Jacob Enriquez
Ellen Lee
National Highway Traffic Safety Administration
U.S.A.

Paper Number 17-0406

ABSTRACT

Research Question/Objective:
The National Highway Traffic Safety Administration (NHTSA) is actively studying the implementation of Advanced Automatic Collision Notification (AACN) systems in motor vehicles. This technology allows motor vehicles to notify a Public Safety Answering Point (PSAP), such as a 911 call center, in the event of a severe crash. The system provides crash location, vehicle identification information, as well as a prediction of severe injury to occupants in the motor vehicle. This paper describes the development of a statistical model that predicts the presence of severely injured and fatal occupants in a motor vehicle involved in a crash.

Methods and Data Source:
A logistic regression model was developed using data from the 1999 – 2015 Crashworthiness Data System (CDS) of the National Automotive Sampling System (NASS). The binary response variable indicates whether or not a crashed vehicle contains a severely injured occupant or a fatally injured occupant, defined by an Injury Severity Score (ISS) of 16 or greater. The predictors are those recommended by the Centers for Disease Control and Prevention (CDC) National Expert Panel on Field Triage, which are delta-V, vehicle body type, multiple vs. single impact, seat belt usage, and principal direction of force. The final dataset is at the vehicle level.

Results:
The area under the receiver operator characteristic curve (AUC) was 0.843, indicating that the model was able to discriminate between vehicles with and without severely injured occupants. At the CDC recommended 0.20 risk threshold, the model produced a sensitivity rate of 26%, a specificity rate of 99%, and identified 41% of vehicles with a fatally injured occupant.

Conclusion:
The sensitivity rate at the CDC recommended 0.20 risk threshold missed 59% of vehicles with a fatally injured occupant. A preliminary cost-benefit analysis showed that the optimal threshold was close to 0.008 after considering the cost of lives saved versus the cost of overtriaging minor injured people using the AACN algorithm. At the 0.008 threshold, 92% of fatal occupants are predicted, the sensitivity is 91%, and the specificity is 60%, which comes close to the recommended levels by the American College of Surgeons.

Limitations:
An AACN system uses data from the event data recorder (EDR) of a vehicle; however, the model developed in this paper was trained with data collected from crash investigations, which may differ from EDR data. Also, this paper only considered the logistic regression model, whereas other data mining classifiers which may produce better results. The initial set of predictors was limited to those selected by the CDC Expert Panel.

INTRODUCTION

In the event of a crash, an Advanced Automatic Collision Notification (AACN) system makes an emergency wireless call to a telematics service provider to send the vehicle's GPS location and crash-related data, and establishes a voice communications channel to the emergency call center. AACN differs from its predecessor, the Automatic Collision Notification (ACN) by including crash severity data as well as a prediction of severe injury. The prediction of severe injury is recommended to be used as part of the Emergency Medical Services (EMS) triage protocol (National Center for Injury Prevention and Control, 2008) to determine which facility to transport an injured patient to (e.g. a local hospital or a trauma center that has additional experience and equipment for treating severely injured people).

The purpose of this paper is to develop a logistic regression model that predicts the presence of severely injured and fatal occupants in a crashed motor vehicle. Published injury severity predictive algorithms were examined in preparation for this paper: an algorithm developed by Kononen et al. (2011) for GM OnStar®, an algorithm developed by Bahouth et al. (2012) for BMW, and an algorithm developed by Stitzel et al. (2016) for Toyota. Similar to the model developed by Kononen et al. (2011), the model developed for this paper follows the approach laid out by the Centers for Disease Control and Prevention (CDC) Expert Panel on Field Triage. In 2008, CDC assembled a panel of experts from various fields such as emergency medicine, trauma surgery, public health, vehicle telematics, and vehicle safety. The panel's purpose was to "*develop a medical protocol for utilization of AACN data from crashes to better predict severity of injury and use this information to improve the ability to respond to crashes and appropriately triage crash victims*." They made several recommendations including a list of predictor variables, criteria for severe injury, and a choice of risk threshold. We used these recommendations as a starting point in this study, and performed tests to assess their validity.

DATA

The predictive model was developed using the Crashworthiness Data System (CDS). It is the only source of data that provides detailed information on injuries as well as crash severity. CDS is a nationally representative probability sample survey whose target population is police reported motor vehicle crashes on a trafficway involving at least one passenger car, pickup, van, or SUV that was towed from the scene due to damage. Crash investigators visit an annual sample of about 5,000 crashes to conduct a vehicle and scene inspection. The CDS three stage sample design and weight computation are described by Zhang and Chen (2013).

This study used CDS years 1999 – 2015 and applied the following filter criteria:

1. Passenger vehicles only (passenger cars, SUVs, vans, and pickups).
2. Deformation locations are front, right, left, and back only (no top or under).
3. Direction of force is between impact points 1 o'clock and 12 o'clock.
4. Vehicle model years 2000 – 2016.
5. Front row passengers only.
6. Passenger ages 0 – 97.
7. Planar crashes (no rollovers).

In addition to these filters, each record (vehicle) must also meet the crash conditions required for the AACN system to make a notification call. Kononen et al. (2011) used the condition of delta-V ≥ 15 mph or airbag deployment, which is also applied in this study. After removing observations with missing data, the final data set has 13,146 records, with a weighted total of 4,206,182. Each record represents a vehicle.

Response variable

The binary response variable, y_i, indicates whether or not a crashed vehicle contains a severely injured occupant or a fatally injured occupant. A value of 1 was assigned to a vehicle if any of its occupants experienced an Injury Severity Score (ISS) of 16 or higher, and zero otherwise. The weighted rate of occurrence of severe injury is 2% (Table 1).

$$y_i = \begin{Bmatrix} 1, & if\ any\ occupant\ with\ ISS \geq 16 \\ 0, & otherwise \end{Bmatrix}$$

An ISS of 16 or greater was used to indicate severe injury and is an anatomic scoring system based on the individual's three highest Abbreviated Injury Scale (AIS) values in different body regions. This was the outcome of interest specified by the 2008 CDC Expert Panel, when they defined severe injury in the context of vehicle telematics. The American College of Surgeons (ACS) periodically publishes a document titled "Resources for Optimal Care of the Injured Patient", which represents the ACS Committee on Trauma's guidelines and recommendations for all aspects of trauma care, including pre-hospital care. In the 2014 version, the

ACS also recommended an ISS of 16+ be used to define major trauma patients. Therefore, this paper focuses on ISS of 16 or greater as the indicator for severe injury.

Table 1 Distribution of the Response Variable

y_i	Frequency	Weighted Frequency	Percent
0	11,984	4,123,989	98
1	1,162	82,194	2
Total	13,146	4,206,182	100

VARIABLE SELECTION

An important step in building a statistical model is determining which variables should be included in the model. For this study, the initial set of predictors were those recommended by the CDC Expert Panel (Table 2). These variables can be electronically transmitted by the vehicle to the AACN providers in the event of a crash.

Table 2. Selected Predictors and Their Descriptions

Variable name	Type	Values	Description
LN_DVMPH	Continuous	0 – 100	Change in the vehicle velocity. Log of delta-V.
DOF1	Categorical	Front, Left, Right, Rear	Direction of force.
CBELT	Categorical	Yes, No	Seat belt usage. Yes = all occupants belted. No = at least one occupant unbelted.
BODY	Categorical	Car, SUV, Pickup, Passenger van	Type of vehicle.
ACCSEQ	Categorical	Multiple, Single	Number of significant impacts to a vehicle.

Note: The variable names are specific to this study and are not the same as in CDS.

A univariate analysis was conducted to determine whether each predictor is "significantly" related to the response variable. This was done using the likelihood ratio chi-squared test and the Wald test.

The likelihood ratio chi-squared test was used to test the null hypothesis of statistical independence between the response variable and each predictor. The p-value for this test is less than 0.0001 for each predictor (Table 3), which provides evidence of an association. The large chi-squared statistics may be heavily influenced by the large weighted sample size.

Table 3. Likelihood Ratio Chi-Square Test

Predictor	Likelihood Ratio Chi-Square Statistic	DF	Probability
ACCSEQ	12,384	1	< 0.0001
BODY	6,257	3	< 0.0001
CBELT	31,948	1	< 0.0001
DOF1	27,091	3	< 0.0001
LN_DVMPH	131,696	1	< 0.0001

A univariate logistic regression model was fit for each predictor to test for the significance of the coefficient using the Wald Chi-Square test statistic,

$$W = \frac{\left(\hat{\beta} - \beta_1\right)^2}{\widehat{Var}\left(\hat{\beta}\right)}$$

Under the null hypothesis that β_1 is equal to zero, the statistic *W* follows a chi-square distribution with 1 degree of freedom. All the predictors and their design variables had p-values less than 0.05, except for the Pickups design variable for the predictor BODY (vehicle body type). It had a p-value of 0.6852 (Table 4). However, the Type 3 multivariate Wald test for the BODY variable, which tests all its design variables simultaneously, has a p-value of <0.0001.

Table 4. Wald Chi-Square Test for the BODY Variable

Param	DF	Est.	Std Err	Wald Chi-Sq	Pr > ChiSq
Intercept	1	-3.766	0.1579	568.3079	<.0001
BODY (Pickups)	1	-0.126	0.3116	0.1643	0.6852
BODY (SUV)	1	-0.751	0.1891	15.7759	<.0001
BODY (Vans)	1	-0.757	0.3521	4.6181	0.0316

In addition to the univariate analysis, a stepwise procedure was conducted. This procedure systematically checks for the "importance" of variables, and either includes or excludes them in the model depending on a decision rule. The procedure starts off with no predictors in the model. In each step, the predictor with the largest Score chi-square statistic that meets the $p < 0.01$ level is included in the model; while the predictor considered least significant according to the Wald test and does not meet the $p < 0.01$ level is removed from the model. The process terminates if no further predictor can be added or if the current model is identical to a previously visited model. Results of this method (Table 5) show that all predictors entered the model, and none were removed. The first variable to enter was delta-V and the last to enter was ACCSEQ (number of impacts to a vehicle). The large Score test statistic values may be heavily influenced by the large weighted sample size.

Table 5. Summary of Stepwise Selection

Step	Var Entered	Var Rem-oved	DF	Score Chi-Sq	Pr>Chi Sq
1	LN_DVMPH	-	1	123,576	<.0001
2	DOF1	-	3	37,454	<.0001
3	CBELT	-	1	37,650	<.0001
4	BODY	-	3	4,913	<.0001
5	ACCSEQ	-	1	4,880	<.0001

Results from the likelihood ratio chi-square test, the Wald test, and stepwise procedure show that it is reasonable to use all the CDC recommended predictors for the multivariate model.

MODELING

Logistic regression was used to estimate the probability that a crashed vehicle contained a seriously injured or fatal occupant, conditional on the values of the predictor variables. The logistic regression model is,

$$P(Y = 1|\boldsymbol{x}) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p}}, \quad \text{Eq. (1)}$$

where $p = 9$ (total predictor variables), $\mathbf{x}' = (x_1, x_2, \dots, x_\mathrm{p})$ is a vector of predictor variables, and $\beta_0, \dots, \beta_p$ are parameters. There are now nine independent variables instead of the initial five since design variables were created for the BODY (vehicle body type) and DOF1 (direction of force) variables.

The model was fit using the maximum likelihood method, which produces an estimate for the parameters that maximizes the probability of obtaining the observed set of data. The SURVEYLOGISTIC procedure in SAS was used to incorporate the CDS survey design by specifying the primary sampling unit (PSU), the PSU stratum, and weight variables. Results of fitting the multiple logistic regression model are given in Table 6.

Testing for the significance of the model

To assess the overall significance of the coefficients for the predictor variables in the model, the likelihood ratio test was used with the null hypothesis that all coefficients in the model are equal to zero. The p-value for the test is < 0.0001 (Table 7), rejecting the null hypothesis, and conclude that at least one coefficient was different from zero.

Table 6. Maximum Likelihood Estimates

Parameter		DF	Estimate	Std. Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	95% Confidence Limits
Intercept		1	-14.4707	0.9508	231.6557	<.0001		(-16.3341, -12.6072)
ACCSEQ	Multiple	1	0.5392	0.1734	9.6657	0.0019	2.3948	(0.1993, 0.8792)
BODY	Pickups	1	-0.5337	0.2015	7.0141	0.0081	-1.4104	(-0.9287, -0.1387)
BODY	SUV	1	-0.7507	0.2056	13.3315	0.0003	-2.8966	(-1.1537, -0.3477)
BODY	Vans	1	-0.4891	0.4739	1.0654	0.3020	-1.1083	(-1.4179, 0.4397)
CBELT	All Belted	1	-1.4283	0.1182	145.9042	<.0001	-5.1421	(-1.6601, -1.1966)
DOF1	Front	1	1.0557	0.3984	7.0230	0.0080	4.0478	(0.2749, 1.8366)
DOF1	Left	1	2.6775	0.4530	34.9351	<.0001	6.0612	(1.7897, 3.5654)
DOF1	Right	1	1.7839	0.4048	19.4198	<.0001	4.3774	(0.9905, 2.5774)
LN_DVMPH		1	3.5073	0.2376	217.8784	<.0001	13.6964	(3.0416, 3.9730

Note: The column between Parameter and DF specifies the comparison group. For example, Multiple is indicated for the variable ACCSEQ because the estimate corresponds to that of multiple event crashes in reference to single event crashes.

Table 7. Testing Global Null Hypothesis: Beta=0

Likelihood Ratio Test	DF	Pr > ChiSq
201,000.246	9	<.0001

The Vans design variable for BODY (vehicle body type) is not significant with a p-value greater than 0.05 for the univariate Wald test, and a confidence interval that includes zero (Table 6). However, the Type 3 multivariate Wald test, which tests the null hypothesis that all the coefficients of the design variables for BODY are simultaneously zero, has a p-value of <0.0001 (Table 8). Hence the BODY variable is not excluded from the model.

Table 8. Type 3 Analysis of Effects

Effect	DF	Wald Chi-Sq	Pr > ChiSq
ACCSEQ	1	9.6657	0.0019
BODY	3	22.1161	<.0001
CBELT	1	145.9042	<.0001
DOF1	3	91.1701	<.0001
LN_DVMPH	1	217.8784	<.0001

Ranking the predictors

A standardized coefficient indicates how many standard deviations of change in the respondent variable are associated with a one standard deviation increase in the predictor variable. Shown in the 8th column of Table 6, the highest standardized coefficient (absolute value) belongs to the LN_DVMPH (log of delta-V) predictor followed by CBELT (all occupants belted or not) and the DOF1 (direction of force) design variables. The lowest standard coefficient belongs to the BODY (vehicle body type) design variables and ACCSEQ (number of significant impacts to the vehicle). This coincides with the order in which the variables entered the stepwise method (Table 5).

Interaction effects

Two-way interaction effects were entered into the main effects model, one at a time, and checked for statistical significance. All interaction effects either had p-values > 0.05 for the univariate Wald test, or did not make scientific sense. Hence no interactions terms were included in the model.

Distribution of the estimated probabilities

The estimated probabilities produced by the model are very low, with a median of only 0.0059 (Table 9). This is due to the rarity of the occurrence of severe injury, with only 2% of the vehicles having at least one occupant with an ISS of 16 or greater (Table 1).

Table 9. Weighted Quantiles of the Estimated Probabilities

Quantile	Estimate
100% Max	0.950031471
99%	0.221031351
95%	0.085091246
90%	0.043765134
75% Q3	0.015382169
50% Median	0.005946187
25% Q1	0.002076919
10%	0.000794507
5%	0.000375181
1%	0.000155431
0% Min	0.000004069

ASESSING THE PREDICTIVE ACCURACY OF THE MODEL

To assess the predictive accuracy of the model, the k-fold cross-validation method was used. In this method the data was split into k = 10 equal-sized subsets. One of the subsets was chosen for testing the model, while the remaining nine subsets were used for training the model. This was repeated k = 10 times so that each record was used for training exactly nine times and testing exactly once. The resulting estimated probability of each record was used to assess the discrimination and accuracy of the model.

Area under the curve

Discrimination refers to the model's ability to distinguish low from high risk vehicles. This means vehicles with y = 1 should have higher probability estimates than vehicles with y = 0. Discrimination can be quantified by the area under the receiver operating characteristic curve (AUC), which is a curve constructed by plotting sensitivity against 1-specificity for different cut-offs. An intuitive explanation of the AUC is that if each vehicle with y = 1 is paired with each vehicle with y = 0, then the AUC is the proportion of the pairings where the vehicle with y = 1 has a higher estimated probability than the vehicle with y = 0. The AUC for this model is 0.843, which is considered excellent discrimination according to Hosmer and Lemeshow (2000).

Classification table

A classification table cross-classifies the binary response variable with the prediction of the model (1 or 0). The estimated probabilities are converted to predictions by first selecting some risk threshold, *t*, where 0 < *t* < 1 (statistical texts refer to this as the cutpoint or the cut-off). If the estimated probability is greater than or equal to *t*, then set the prediction equal to 1; otherwise set the prediction equal to 0. Table 10 shows the classification table, with weighted counts, for a threshold of 0.20, which is the recommended threshold by the CDC Expert Panel. The overall rate of correct classification is estimated as (4,093,805 + 21,688)/ 4,206,182 = 98%, with 26% (21,688/82,194) of the y = 1 group (sensitivity) and 99% (4,093,805/4,123,989) of the y = 0 group (specificity) being correctly classified.

Table 10. Classification Table (Weighted) Using a Threshold of 0.20.

	Response		
Predicted	**0**	**1**	**Total**
0	4,093,805	60,506	4,154,311
1	30,184	21,688	51,872
Total	4,123,989	82,194	4,206,182

Aside from sensitivity and specificity, the model was also assessed in how well it identified vehicles with a fatally injured occupant, referred to as fatal vehicles. Fatal vehicles are a subset of the y = 1 group, and should have a prediction of 1. The proportion of fatal vehicles identified by the model (having a predicted value of 1) was 41%, using the 0.20 threshold. Figure 1 plots the sensitivity, specificity, and percent of fatal vehicles identified by the model at different thresholds.

Relating sensitivity and specificity to undertriage and overtriage, and their recommended levels

The 2014 edition of the American College of Surgeons (ACS) Resources for Optimal Care defines undertriage as severely injured patients transported to lower-level trauma centers or other facilities, and overtriage as minimally injured patients transported to higher-level trauma centers. The ACS gives higher priority to reduction of undertriage, because undertriage may result in preventable mortality or morbidity from delays in definitive care. The recommended level for undertriage is 5%. Overtriage may result in higher costs and also increase the burden for higher-level trauma centers because resources needed for more severely injured patients are unnecessarily being used for minimally injured patients. Acceptable rates for overtriage are in the range of 25-35% according to the ACS.

In the context of the injury prediction algorithm developed here, the sensitivity of the algorithm is equal to 100% minus the undertriage rate (i.e. a sensitivity of 95% will result in 95% of seriously injured occupants being *correctly identified* as seriously injured, and 5% being *undertriaged, or incorrectly identified* as not seriously injured). Specificity, or the true negative rate (proportion of occupants with ISS < 16 who are correctly identified by the algorithm as having a low risk of injury), is

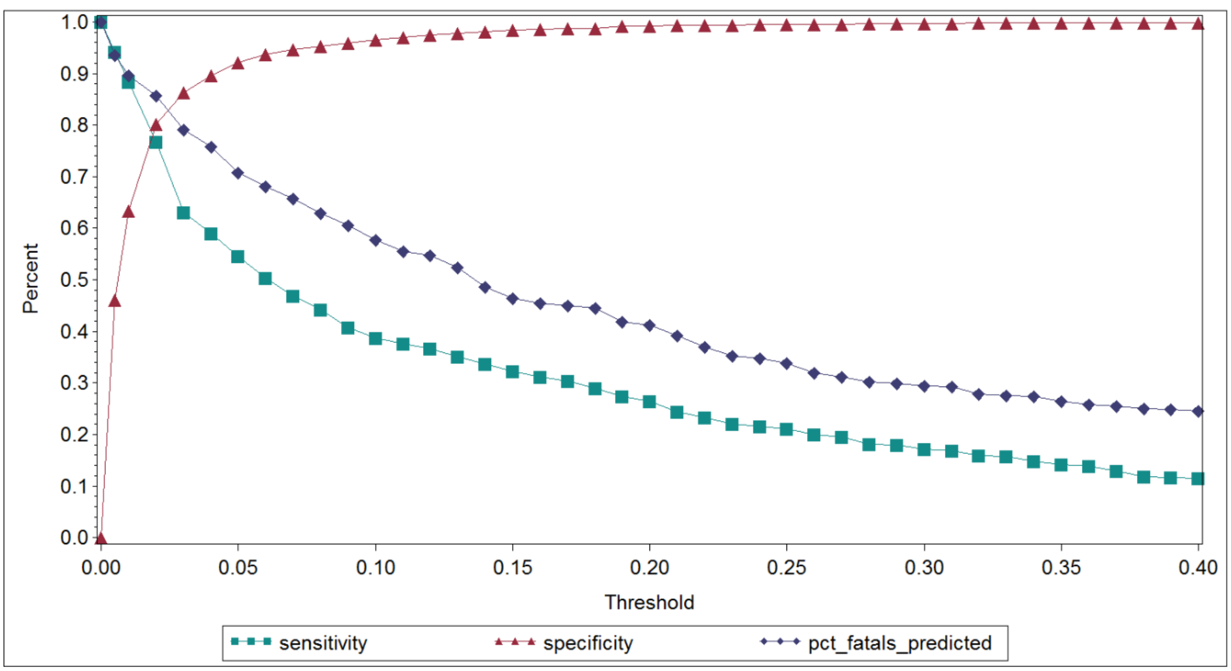

Figure 1. Plot of Sensitivity, Specificity, and Percent Fatal Vehicles Identified by Threshold.

equal to 100% minus the overtriage rate. It is noteworthy that at the CDC recommended 0.20 risk threshold, the prediction algorithm falls far short of the recommended 5% undertriage rate (instead resulting in a 74% undertriage rate), while it far exceeds the recommendations for overtriage (predicting only 1% overtriage, rather than the ACS recommended 25-35%).

In order to meet the 5% undertriage rate, the threshold needs to be lowered from the 0.20 threshold. As shown in Figure 1, lowering the threshold increases both sensitivity and percent of fatal vehicles identified, but it also lowers the specificity. Lowering the specificity is equivalent to increasing the rate of false positives (false alarms), which results in overtriage costs. Finding the right balance of increasing the percent of fatal vehicles identified by the model while minimizing the rate of false positives is addressed in the next section.

FINDING AN OPTIMAL THRESHOLD

As demonstrated above, at the CDC recommended 0.20 risk threshold, the prediction algorithm falls far short of the undertriage rates recommended by the ACS. To provide a basis for choosing an optimal threshold that deviates from the CDC recommendation, the costs of under- and overtriage were evaluated. For a preliminary determination of an ideal threshold for the model, the benefit of true positives was weighed against the cost of false positives at thresholds below 0.20. The benefit of true positives is the economic savings from those that would have died but were saved due to AACN. The cost of false positives comes from overtriage, which is transporting occupants without serious injuries to major trauma centers.

Benefits

The benefits at a specific threshold is the number of lives saved by AACN multiplied by the dollar amount saved per fatality prevented. Lee et al. (2017) estimated the number of lives saved by AACN to be, at most, 721 per year. This number assumes the predictive model identifies 90% of the fatal occupants. This percentage is replaced with the appropriate percentage at each threshold. As for the economic savings, Blincoe et al. (2015) estimates the comprehensive fatality injury cost to be $9,129,066.[1] Since a fatality prevented by AACN cannot be considered to be uninjured, it is assumed that the saved occupant will still have a maximum AIS (MAIS) 4 injury level with a comprehensive injury cost of $2,414,252.[1] The cost savings of preventing a fatality is the difference between these two injury costs which is $6,714,814. The benefit at a particular threshold, *t*, can now be expressed as,

$$Benefit(t) = \frac{721}{0.90} \times \%FatalsPred(t) \times \$6{,}714{,}814$$

Costs

The cost at a specific threshold is the number of minor injured occupants (ISS < 16) unnecessarily treated at a trauma center multiplied by the cost of overtriage per patient. The number of occupants with ISS < 16 is estimated to be around 4 million annually, using CDS 2006-2008. Since not all of these occupants will be sent to a trauma center as a direct result of AACN, the following reduction factors were applied:

1. % overtriage NOT identified by steps 1 and 2 of the triage protocol = 78%.[2]
2. % of occupants with ISS < 16 that were in a crashed vehicle that met the conditions for the AACN system to make a call (i.e. delta-V ≥ 15 or airbag deployment) = 60%.
3. % access to trauma center = 80% (NHTSA, 2012).

Applying the reduction rates to the 4 million occupants produces 1,497,600 which is then applied the rate of false positives at a specific threshold. The rate of false positives is equal to one minus the specificity computed at the occupant level.

The cost of minor injured occupants treated at a trauma center is approximately $5,000 - $10,000 according to Newgard et al. (2013) and Faul et al. (2012). Using the midpoint of this range, the cost at a particular threshold is,

$$Cost(t) = 1{,}497{,}600 \times (1 - specificity(t)) \times \$7{,}500$$

Computed values for benefits, costs, and their difference are shown in Table 11 and plotted in Figure 2. At the CDC recommended threshold of 0.20, benefits exceed costs by about $2.18 billion. As the threshold is lowered, benefits continue to be

[1] This equals comprehensive costs less congestion costs and property damage costs. Comprehensive costs consist of tangible losses (such as property damage, medical care, insurance costs, legal costs, etc.) plus costs associated with lost quality of life.

[2] Newgard et al. (2011) and Brown et al. (2011) show 14% - 22% overtriage using just steps 1 and 2 of the triage protocol.

greater than costs. Around the 0.06 threshold, costs start to climb at a higher rate than benefits, and eventually the two become equal somewhere between the 0.008 and 0.007 thresholds. After this point costs exceed benefits.

Since lowering the threshold results in more lives saved, then 0.008 is the threshold where the maximum number of lives can be saved without costs exceeding benefits. This number seems to be the logical choice as the optimal threshold. At this threshold, 92% of fatal occupants are predicted, the sensitivity is 91% (undertriage rate of 9%), and the specificity is 60% (overtriage rate of 40%). These results are approximately consistent with the ACS recommended under- and overtriage levels of 5% and 25-35% respectively.

The small threshold of 0.008 may seem to suggest that the model will predict nearly all vehicles that meet the AACN crash criteria to have a severely injured occupant. This is not the case. According to the distribution of the estimated probabilities (Table 9), among vehicles that meet the AACN crash criteria, the proportion having an estimated probability greater than 0.008 is around 40%.

Although other published logistic regression models (e.g. Bahouth et al. 2012; Stitzel et al. 2015) did not consider the economic costs and benefits of under- and overtriage, their optimal predictive performance occurred at thresholds lower than the CDC recommended 0.20, similar to the findings of the current study.

CONCLUSION

The purpose of this paper was to develop a logistic regression model that predicts the presence of severely injured and fatal occupants in a crashed motor vehicle. The model was trained using 1999-2015 CDS data, accounting for its sample design. The binary response variable indicates whether or not a crashed vehicle contains a severely injured occupant or a fatally injured occupant. The predictors are those recommended by the CDC Expert Panel on Field Triage, which are delta-V, direction of force, vehicle body type, seat belt use, and number of crash events (multiple or single). The most significant predictor is delta-V followed by seat belt use and direction of force. At the CDC recommended threshold of 0.20, the model produces an AUC of 0.843, a sensitivity of 26%, a specificity of 99%, and predicts 41% of the fatal vehicles (Figure 1). Based on a preliminary cost-benefit analysis considering the cost of lives saved versus the cost of overtriaging minor injured people using the AACN algorithm, the study showed that the optimal threshold was close to 0.008. At this threshold, 92% of fatal occupants are predicted, the sensitivity is 91%, and the specificity is 60%.

LIMITATIONS AND FUTURE CONSIDERATIONS

1. An AACN system uses data from the event data recorder (EDR) of a vehicle. The model developed in this paper, as well as others, were not trained with EDR data but with data collected from crash investigations (CDS). There may be differences between these two data sources, particularly the WinSmash delta-V estimates in CDS that have been found to underestimate EDR delta-V by as much as 23%. This study attempted to use EDR data but found it to be quite incomplete.

2. This study only considered predictors recommended by the CDC Expert Panel. While these variables were approved by subject matters experts, this study did not consider all possible predictors in CDS and other data sets.

3. This paper only considered one statistical model, logistic regression, among many classifiers that may produce better results. Kusano and Gabler (2014) compared several competing classification algorithms for predicting injured occupants in vehicle crashes and concluded that logistic regression slightly outperformed the machine learning algorithms based on sensitivity and specificity of the models.

Table 11. Benefits and Costs at Different Threshold Levels Below 0.20

Threshold	Specificity (occupants)	% Fatals predicted (occupants)	Benefits	Costs	Benefits minus Costs
0.000	0	1	$5,379,312,104	$11,232,000,000	($5,852,687,896)
0.005	0.45261	0.93756	$5,043,416,563	$6,148,264,446	($1,104,847,883)
0.007	0.54937	0.92921	$4,998,535,709	$5,061,452,716	($62,917,007)
0.008	0.57975	0.91807	$4,938,578,607	$4,720,220,174	$218,358,433
0.010	0.62122	0.89965	$4,839,495,036	$4,254,452,851	$585,042,185
0.020	0.79856	0.86168	$4,635,224,187	$2,262,520,828	$2,372,703,359
0.030	0.85933	0.79801	$4,292,737,805	$1,579,995,498	$2,712,742,307
0.040	0.89118	0.76739	$4,128,026,556	$1,222,219,386	$2,905,807,171
0.050	0.91967	0.71522	$3,847,380,225	$902,317,804	$2,945,062,421
0.060	0.93615	0.68907	$3,706,736,390	$717,113,769	$2,989,622,621
0.070	0.94585	0.66701	$3,588,052,336	$608,160,059	$2,979,892,276
0.080	0.95181	0.64044	$3,445,146,443	$541,242,997	$2,903,903,446
0.090	0.95953	0.61765	$3,322,509,772	$454,586,639	$2,867,923,132
0.100	0.96538	0.58943	$3,170,754,006	$388,886,788	$2,781,867,218
0.150	0.98345	0.47427	$2,551,270,193	$185,916,103	$2,365,354,090
0.200	0.99227	0.42096	$2,264,466,194	$86,810,724	$2,177,655,470

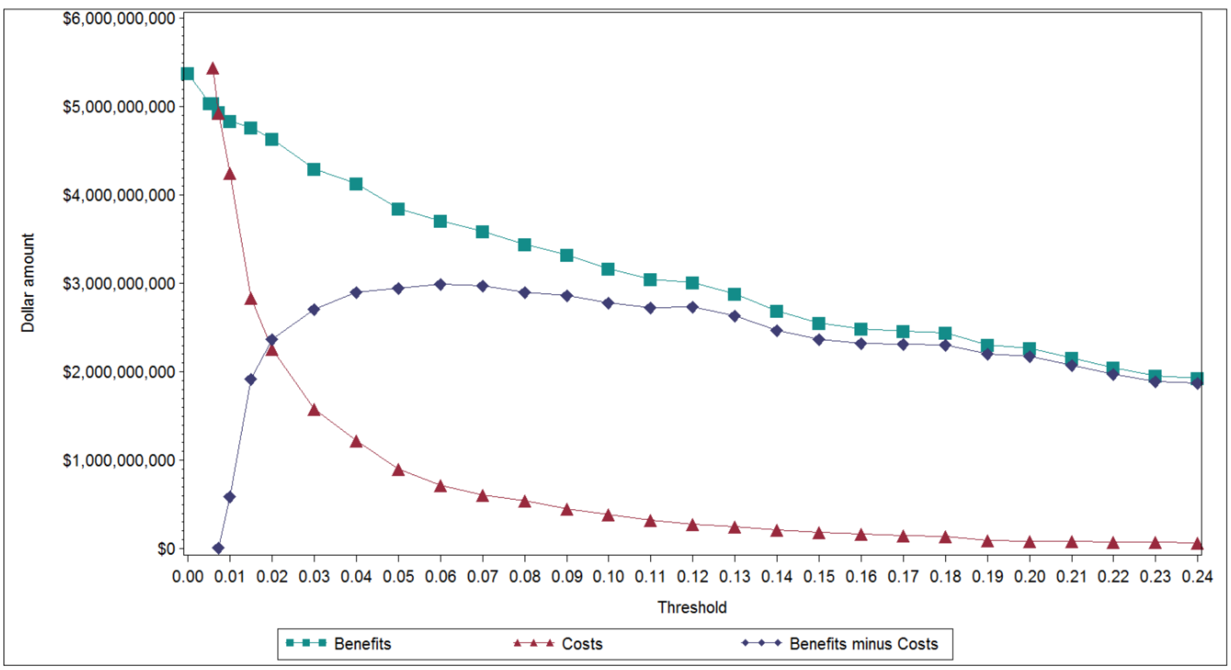

Figure 2. Difference between benefits and costs by threshold levels.

REFERENCES

[1] Agresti, A. 2002. Categorical Data Analysis, Second Edition, New York: John Wiley & Sons.

[2] Bahouth, G., Digges, K. and Schulman, C. 2012. Influence of injury risk thresholds on the performance of an algorithm to predict crashes with serious injuries. In Annals of Advances in Automotive Medicine/Annual Scientific Conference (Vol. 56, p. 223). Association for the Advancement of Automotive Medicine.

[3] Blincoe, L. J., Miller, T. R., Zaloshnja, E., and Lawrence, B. A. 2015. May. The economic and societal impact of motor vehicle crashes, 2010. (Revised) (Report No. DOT HS 812 013). Washington, DC: National Highway Traffic Safety Administration.

[4] Brown, J.B., Stassen, N.A., Bankey, P.E., Sangosanya, A.T., Cheng, J.D. and Gestring, M.L. 2011. Mechanism of injury and special consideration criteria still matter: an evaluation of the National Trauma Triage Protocol. Journal of Trauma and Acute Care Surgery, 70(1), pp.38-45.

[5] Faul M, Wald MM, Sullivent EE, Sasser SM, Kapil V, Lerner EB, Hunt RC. 2012. Large cost savings realized from the 2006 Field Triage Guideline: reduction in overtriage in U.S. trauma centers. Prehosp Emerg Care 16:222-229.

[6] Hastie, T., Tibshirani, R., Friedman, J. H. 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York, Springer.

[7] Hosmer, D. W., Jr. and Lemeshow, S. 2000. Applied Logistic Regression, Second Edition, New York: John Wiley & Sons.

[8] Kononen, D.W., Flannagan, C.A. and Wang, S.C. 2011. Identification and validation of a logistic regression model for predicting serious injuries associated with motor vehicle crashes. Accident Analysis & Prevention,43(1), pp.112-122.

[9] Kusano, K., Gabler, H. 2014. Comparison and Validation of Injury Risk Classifiers for Advanced Automated Crash Notification Systems. Traffic Injury Prevention. 2014. 15, S126–S133.

[10] Lee, E., Wu, J., Kang, T. and Craig, M. 2017. Estimate of Safety Benefit with Full Implementation of Advanced Automatic Collision Notification. ESV (Enhanced Safety of Vehicles Conference). Detroit, USA.

[11] National Center for Injury Prevention and Control. 2008. “Recommendations from the Expert Panel: Advanced Automatic Collision Notification and Triage of the Injured Patient.” Atlanta (GA): Centers for Disease Control and Prevention.

[12] Newgard, C.D., Zive, D., Holmes, J.F., Bulger, E.M., Staudenmayer, K., Liao, M., Rea, T., Hsia, R.Y., Wang, N.E., Fleischman, R. and Jui, J. 2011. A multi-site assessment of the ACSCOT field triage decision scheme for identifying seriously injured children and adults. Journal of the American College of Surgeons, 213(6), p.709.

[13] Newgard CD, Staudenmayer K, Hsia RY, Mann NC, Bulger EM, Holmes JF, et al. The cost of overtriage: more than one-third of low-risk injured patients were taken to major trauma centers. Health Aff (Millwood). 2013;32:1591–1599. doi: 10.1377/hlthaff.2012.1142.

[14] NHTSA. 2012. The association between crash proximity to level 1 and 2 trauma centers and crash scene mortality of drivers injured in fatal crashes. DOT HS 811 599.

[15] NHTSA. 2012. Analysis of the proximity of fatal motor vehicle crash locations to the availability of helicopter emergency medical service response. DOT HS 811 542.

[16] Sasser, S.M., Hunt, R.C., Faul, M., Sugerman, D., Pearson, W.S., Dulski, T., Wald, M.M., Jurkovich, G.J., Newgard, C., Lerner, E.B. and Cooper, A. 2012. Guidelines for field triage of injured patients recommendations of the national expert panel on field triage, 2011. Morbidity and Mortality Weekly Report, 61(RR-1), pp.1-23.

[17] Stitzel, J.D., Weaver, A.A., Talton, J.W., Barnard, R.T., Schoell, S.L., Doud, A.N., Martin, R.S. and Meredith, J.W. 2016. An injury severity-, time sensitivity-, and predictability-based advanced automatic crash notification algorithm improves motor vehicle crash occupant triage. Journal of the American College of Surgeons, 222(6), pp.1211-1219.

[18] Tan, P., Steinbach, M., Kumar V. 2005. Introduction to Data Mining. Addison-Wesley Longman Publishing.

[19] Zhang, F., & Chen, C-L. 2013. NASS-CDS: Sample Design and Weights. (Report No. DOT HS 811 807). Washington, DC: National Highway Traffic Safety Administration.