Accident Analysis and Prevention 43 (2011) 112–122

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Identification and validation of a logistic regression model for predicting serious injuries associated with motor vehicle crashes

Douglas W. Kononen[a,*], Carol A.C. Flannagan[b], Stewart C. Wang[c]

[a] General Motors Corporation, Vehicle Structure and Safety Integration, 30200 Mound Road, Warren, MI 48090, United States
[b] University of Michigan Transportation Research Institute, United States
[c] University of Michigan Medical Center, United States

ARTICLE INFO

Article history:
Received 1 April 2010
Received in revised form 26 July 2010
Accepted 29 July 2010

Keywords:
Logistic regression
Clinical prediction modeling
Motor vehicle crash injuries

ABSTRACT

A multivariate logistic regression model, based upon National Automotive Sampling System Crashworthiness Data System (NASS-CDS) data for calendar years 1999–2008, was developed to predict the probability that a crash-involved vehicle will contain one or more occupants with serious or incapacitating injuries. These vehicles were defined as containing at least one occupant coded with an Injury Severity Score (ISS) of greater than or equal to 15, in planar, non-rollover crash events involving Model Year 2000 and newer cars, light trucks, and vans. The target injury outcome measure was developed by the Centers for Disease Control and Prevention (CDC)-led National Expert Panel on Field Triage in their recent revision of the Field Triage Decision Scheme (American College of Surgeons, 2006). The parameters to be used for crash injury prediction were subsequently specified by the National Expert Panel. Model input parameters included: crash direction (front, left, right, and rear), change in velocity (delta-V), multiple vs. single impacts, belt use, presence of at least one older occupant (≥55 years old), presence of at least one female in the vehicle, and vehicle type (car, pickup truck, van, and sport utility). The model was developed using predictor variables that may be readily available, post-crash, from OnStar®-like telematics systems. Model sensitivity and specificity were 40% and 98%, respectively, using a probability cutpoint of 0.20. The area under the receiver operator characteristic (ROC) curve for the final model was 0.84. Delta-V (mph), seat belt use and crash direction were the most important predictors of serious injury. Due to the complexity of factors associated with rollover-related injuries, a separate screening algorithm is needed to model injuries associated with this crash mode.

1. Background

Models for identifying and predicting the potential severity of occupant injuries associated with highway crashes can be used to both direct appropriate first responder resources to the crash scene and provide critical information to emergency trauma centers to facilitate appropriate preparations for receipt of transported seriously injured occupants (Bahouth et al., 2004). Modern telematics communications systems available to vehicle owners (e.g., OnStar®) can provide immediate information about the nature and severity of a motor vehicle collision. This information can serve as input to predictive models designed to classify a crash as a high injury probability event or a low injury probability event. Some of the information immediately available to telematics systems from Event Data Recorders (EDRs) in modern vehicles may include: principal direction of impact force (PDOF), the total change in vehicle velocity during the event (delta-V), seat belt use status of occupants and frontal or side airbag deployment. This information, when coupled with vehicle information derived from the vehicle identification number (VIN) (e.g., vehicle type and weight) and voice contact with vehicle occupants, when available, can serve as the basis for targeted allocations of first responder services to ensure that appropriately equipped and trained Emergency Management Services (EMS) are dispatched to the crash scene.

Since automatic collision notification (ACN) is relatively new, only a few studies have been published. Augenstein et al. (2007) discuss some of the predictors of crash outcome that can be measured by the EDR, including crash direction and delta-V. They illustrate the relationship between these variables and crash outcome and discuss the benefit of both ACN and the potential for more extensive crash data from EDR reports. Rauscher et al. (2009) report on the first field experience with an ACN system that transmits geographic coordinates of the crashed vehicle. This report focuses primarily on the types of voice contact (or lack thereof) that are made after notification, but it includes some information about outcome. These data

* Corresponding author. Tel.: +1 586 899 5510.
E-mail addresses: douglas.w.kononen@gm.com, kononend@aol.com (D.W. Kononen).

0001-4575/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2010.07.018

provide a starting point for learning about the potential benefits of ACN, especially when no voice contact can be made.

In 2008, the Centers for Disease Control (CDC) published the recommendations of National Expert Panel on Field Triage. In the new Field Triage Decision Scheme (Sasser et al., 2009), "vehicle telematics" was added to the list of crash-related information to consider in determining whether to transport an occupant to a hospital. At that time, the specifics of how to use vehicle telematics were not addressed, but the expert panel made several recommendations related to future use of such information.

In this paper, we present an algorithm for predicting risk of serious injury in a vehicle as a function of crash parameters that can be obtained from the EDR or from voice communication with the vehicle. The algorithm was developed for OnStar®, and in some instances, modeling choices were made based on the capabilities of the OnStar® system. However, the model as a whole is applicable to all vehicle makes and models. Moreover, the algorithm follows the recommendations of the expert panel and is based on publically available national crash data.

The objective of this modeling effort was to develop a statistical model to predict the risk of serious injury using outcome and predictor variables identified by the CDC National Expert Panel on Field Triage based on an analysis of planar motor vehicle crashes recorded in the 1999–2008 National Automotvie Sampling System (NASS-CDS) database files for model year 2000 and newer vehicles. The model will be incorporated into the existing OnStar® response-center system to help identify crashes that have high potential to result in severe injuries.

2. Data and methods

Data from the National Automotive Sampling System (NASS) Crashworthiness Data System (CDS), years 1999–2008, were used to develop and validate a multivariate logistic regression model of serious injury as a function of those predictor variables that may be readily transmitted from Event Data Recorder (EDR) modules to the OnStar® system. The NASS-CDS database is a complex stratified sample of crashes in the United States (National Highway Traffic Safety Administration, 2007). A NASS-CDS 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 transport on a trafficway.

From a practical perspective, this modeling effort was restricted to using only those variables that can be immediately and accurately obtained via telematic transmission (i.e., EDR data) supplemented with information obtained from voice communication with crash-involved vehicle occupants, if available. This constraint on potentially significant explanatory variables was imposed to reflect the current state of information available from current vehicle telematic systems.

Most predictive models of occupant injury outcome of which we are aware focus on the occupant level (e.g., Bahouth et al., 2004; Augenstein et al., 2007). However, in the scenarios being considered here, a vehicle's telematic system would make the initial contact and Emergency Medical Services (EMS) would respond to the vehicle/crash as a whole. Thus, the models in this paper are designed to predict which *vehicles* involved in a crash are likely to contain a seriously injured occupant, and all modeling is done at the vehicle level. Occupant information available to the EDR, as well as any gathered from voice contact, is coded to the vehicle level. Methods for this are discussed in the sections describing each predictor.

Although with many systems (including OnStar at this time), EDR-based occupant information may only be available for front-seat occupants, we included the maximum injury of any passenger in a vehicle. In this way, the model represents the current state of information available to the EDR, but predicts the complete range of vehicles and their occupants. As rear-seat information (occupancy and belt use) becomes available, prediction may improve. However, it is worth noting that the most severely injured occupant is a rear-seat occupant in only 3.5% of vehicles in NASS-CDS, so future information about rear-seat occupant presence and belt use may provide limited additional predictive value.

Published models of serious injury associated with motor vehicle crashes are often based upon an injury outcome criterion such as MAIS3+ (Bahouth et al., 2004; Farmer, 2003). For the purposes of this study, Injury Severity Score (ISS) was considered a better, more clinically reliable indicator of severe injury than indices based upon the score attributed to a single (presumably most severe) coded injury using the Abbreviated Injury Scale (AIS) injury coding system (e.g., MAIS3+). Because the ISS score, defined as the sum of the squares of the AIS severity level of the three most significant coded occupant injuries, captures a larger portion of an injured occupant's "harm profile", it is thought to provide a better, more realistic assessment of occupant harm than does a simpler univariate score such as MAIS3+. While there are many ways to dichotomize the severity of injuries, the National Expert Panel chose ISS of 15 as the partition during their revision of the Field Triage Decision Scheme (Sasser et al., 2009). The National Expert Panel also chose a 20% probability of ISS 15 or greater as the threshold for considering individual triage criterion for inclusion in the Field Triage Decision Scheme. Thus, the dependent measure for this study is the binary variable, ISS 15+, indicating whether any occupant of a vehicle experienced an injury of ISS 15+ or not.

The final sets of data exclusions were related to the nature of the vehicles that contain EDRs. Although EDRs have been introduced slowly by manufacturer and model, they were generally not available before model year 2000. In addition, we limited our sample to planar collisions (i.e., excluding rollovers and the rare crashes coded with the primary general area of damage as top or bottom) and passenger vehicles. Finally, cases with weights of 5000 and up were trimmed (excluded) to improve standard errors (Little et al., 1997; Potter, 1990). Previous experience using the NASS–CDS complex survey data indicate that cases with weights greater than 5000 are usually extreme outliers that often exert a large influence on resulting model parameter estimates and their standard errors. Cases with such large weighting factors are very rare and readily identified. Our analyses with and without these influential cases demonstrated that a single case with a weight of 5000 or greater can dramatically change some model parameter estimates and their standard errors. Therefore, these cases were excluded from subsequent analyses. The resulting dataset contained 14,673 vehicles. Of these, 1212 (8.3%) contained one or more occupants with ISS 15+ injuries.

The CDC National Expert Panel on Field Triage recommended the following variables as predictors of serious occupant injury risk (ISS 15+): delta-V, principal direction of force (front, left, right, rear), seat belt restraint use (yes vs. no), vehicle type (car, sport utility, pickup, passenger van) and multiple vs. single crash events. In addition, information about occupant age and gender may be obtained if verbal contact is made by an OnStar operator or EMS dispatcher.

The following section describes each predictor variable in detail.

2.1. Delta-V

Delta-V is the change in vehicle velocity associated with the primary direction of force of the crash event. In the NASS-CDS database, delta-V is defined as the difference between Impact Velocity and Separation Velocity and is calculated by a computer model (WinSmash) based upon detailed vehicle crush measurements obtained by NASS-CDS crash investigators. Typically, the single most significant predictor of serious injury is the measure of crash energy captured by delta-V. Older EDRs can only measure longitudinal delta-V, which is of limited value. However, newer EDRs, include both lateral and longitudinal delta-V, which also allows computation of crash direction. The model we present is relevant only to bi-directional EDR data. Prior to analysis, delta-V values were converted from kph to mph.

2.2. Crash direction

Crash direction refers to the principal direction of force (PDOF) associated with the primary crash event as identified by NASS-CDS crash investigators. Crash direction was constrained to planar crashes using the "1–12 o'clock" PDOF direction variable. Frontal crashes are defined as those with PDOF of 11, 12 or 1 o'clock, right side are those from 2 to 4 o'clock, rear from 5 to 7 o'clock, and left from 8 to 10 o'clock. Similar information is available from EDR data when two axes of acceleration are captured. Of the vehicles in the study set, 69.2% were vehicles that experienced frontal crashes, 9.7% were right impacts, 9.2% were left impacts and 12.0% were rear crashes.

Previous efforts to model the injury severity of occupants using NASS-CDS data (Augenstein et al., 2007; Bahouth et al., 2004) included rollovers with planar crashes. Among the measurements provided by the EDR, delta-V is generally found to be most predictive of injury. However, in rollover crashes, delta-V is of questionable value since current automotive technologies can only reliably capture delta-V information for planar crashes. Consultations with vehicle crash engineers and biomechanical engineers indicated that delta-V information provided by current generation EDRs is expected to be of limited or no value in predicting occupant injuries associated with rollover events. Our preliminary evaluations indicated that injury risk predictions for rollover events need to be determined using a separate algorithm for rollover crashes using those variables that have previously been found to be most important in predicting occupant injury outcomes in that crash mode (i.e., belt restraint use and occupant ejection status). Thus, we did not include rollover events in this first effort.

2.3. Vehicle type

Vehicle type in this study refers to the NHTSA standard definitions of light truck and passenger cars as defined in the NASS-CDS coding and editing manual (NHTSA, 2007): passenger cars, passenger vans, pickup trucks and sport utility vehicles (SUVs). The body type variable was used with the following specific codes: 0–9 are cars, 14–19 are SUVs, 20–29 are vans, and 30–33 are pickups. All other vehicles were eliminated, but these comprise only 0.3% of the entire NASS-CDS sample. Of the vehicles included in this study, 64% were passenger cars, 18% were sport utility vehicles, 12% were pickup trucks and 6% were passenger vans or minivans.

2.4. Belt restraint use

Belt restraint use refers to NASS-CDS-identified occupant seat belt restraint use. To code at the vehicle level, we initially set up two variables, one for the driver and one for the right-front passenger. Two codes were used for the driver, based on the NASS-CDS coding manual and automatic belt-use variables: "belted" and "unbelted." For the right-front passenger, a third code of "no RFP" was added to distinguish the case where the belt buckle was not attached because there was no passenger vs. the case where a passenger was present but unbelted. In our sample, 84% of drivers were belted, 71% of vehicles had no RFP, 23% had a belted RFP, and 5% had an unbelted RFP.

After initial analyses of injury outcome using these variables, we discovered an interaction between driver and passenger belt status such that vehicles with multiple unbelted passengers did not have substantially higher risk of containing at least one seriously injured occupant compared to vehicles with one unbelted occupant. Vehicles with all front passengers belted were significantly less likely to have a seriously injured occupant in a crash. Because of this result, we simplified our belt-use code to indicate at the vehicle level: "all occupants belted" vs. "at least one occupant unbelted." In coding vehicles in this way, we did not consider the belt status of rear-seat passengers because current EDRs do not sense the presence or belt status of rear-seat occupants. Future models may incorporate this effect, but as mentioned earlier, the most injured occupant in a vehicle is rarely a rear-seat occupant, so the belt status of those in the rear is unlikely to be a strong predictor of the most serious injury in a vehicle (though it would predict the risk to the rear-seat occupant).

2.5. Multiple impacts

The occurrence of more than one significant impact to a vehicle may be an important explanatory variable in the prediction of injury outcome. Vehicles experiencing multiple vs. single impacts were identified from NASS-CDS using the accident sequence variables, which identify the two crash events that are most relevant for a given vehicle. Each crash as a whole is divided into events, which are generally separate impacts (e.g., Vehicle 1 hits Vehicle 2 in the side, sending Vehicle 2 into a tree involves two events). These impacts may affect different vehicles involved in the crash, such that some vehicles in a multiple-event crash may only experience one impact (e.g., Vehicle 1) and others may experience more than one impact (e.g., Vehicle 2). In NASS-CDS, any vehicle for which a second accident sequence was identified was coded as "multiple impacts." Vehicles with only the first accident sequence variable present were coded as "single impact." A typical EDR will record up to two impacts, as long as they are above a certain threshold for recording. Thus, the EDR is also capable of distinguishing between vehicles that experience a single impact and those that experience at least two impacts. Of the vehicles included in this study, 38% experienced multiple impacts and 62% experienced single impacts.

2.6. Age and gender

Although the EDR does not have information about age and gender of occupants, it may be possible for an operator (e.g., OnStar® advisor) to make verbal contact with a vehicle's occupants after initial notification in order to gather some additional information. Age is known to be a strong predictor of injury (MacKenzie et al., 2006; Champion et al., 1990; Grossman et al., 2002; Morris et al., 1990), and it is included in the CDC's triage rules (ACS, 2006; Sasser et al., 2009). Similarly, females have been shown to be more susceptible to some types of injury than males in motor vehicle crashes (Sampalis et al., 2009; Schiff et al., 2008; Rowe et al., 2004; Tavris et al., 2001). These are considered to be two key pieces of information that an operator might be able to obtain from a crash victim (who may be distraught or otherwise having difficulty answering simple questions).

To code these at the vehicle level, we chose an age cutoff of 55 and coded a vehicle as "55+" if anyone in the vehicle is 55 or older, and "under 55" if all occupants are under 55. This cutoff was chosen based on the National Expert Panel's decision to retain age 55 as a criterion for consideration in the Field Triage Decision Scheme (Sasser et al., 2009). Similarly, gender was coded as "female present" at the vehicle level if any occupant is female and "all male" if all occupants are male. Rear-seat occupants were included in this coding system because they would presumably be included in the answer given by a driver. Of the vehicles in the study dataset, 20% contained an occupant aged 55 or older and 58% contained a female occupant.

3. Analysis approach

3.1. Modeling

Logistic regression was conducted using SAS 9.2 PROC SURVEYLOGISTIC (SAS Institute, 2008) to account for the sample design for NASS-CDS. All analyses used weighted data, except where indicated, and weights were trimmed at 5000 (mean weight = 314.0). NASS-CDS is geographically divided into 12 strata and 27 probability sampling units (PSUs), which were accounted for in all analyses. Taylor series expansion was used to estimate standard errors.

Logistic regression is a maximum-likelihood method that has been used in hundreds of studies of crash outcome (e.g., Hours et al., 2010; Robertson and Vanlaar, 2008; Schiff et al., 2008). For a binary response variable the linear logistic regression model, expressed in terms of the logit transformation of the ith individual's response probability, p_i (e.g., probability of severe injury), is a linear function of the vector of explanatory variables or,

$$\mathrm{logit}(p_i) = \log\left[\frac{p_i}{1-p_i}\right] = b_0 + b_1x_1 + \cdots + b_jx_j + \cdots + b_nx_n \qquad (1)$$

where $j = 1$, n for n predictor variables. The predictors can be categorical dummy variables or continuous measures. The negative sign before the linear combination of predictors is arbitrary, but produces a positive relationship between the sign of the coefficient and the direction of effect on risk. In other words, a positive coefficient represents an increase in risk and a negative coefficient represents a decrease in risk.

Solving for p_i results in the individual probability of the risk of the event of interest or

$$p_i = \frac{1}{\exp[-(b_0 + b_1x_1 + \cdots + b_nx_n)]} \qquad (2)$$

3.2. Outliers

For each analysis, we computed influence statistics typically associated with logistic regression, including change in deviance and the df betas. For one case, one or more of these statistics was grossly out of line with other observations, and it was eliminated from analysis. This will be discussed in the results.

3.3. Imputation

For most variables used in this analysis, missingness is relatively minor. Of the 14,861 vehicles in the dataset, 159 were missing driver belt status. Thirty vehicles were missing occupant age information and an additional 31 included only records of occupants who were 13 or under. Two were missing occupant gender. We coded these vehicles as missing, leaving 14,673 cases. These missing cases represent such a small fraction of the total (1.4%) that the method used to handle them (elimination vs. imputation) will have little or no effect on the model estimates. In contrast, delta-V is missing in 4667 cases, or 32%. Research in the statistical literature suggests that in analyses where a substantial portion of the cases are missing a key variable, imputation will likely produce estimates with smaller standard errors (Harrell, 2001; Steyerberg, 2009). However, Harrell (2001) does include the caveat that “if the predictor of interest is the only variable having a substantial number of missing values, multiple imputation is less worthwhile, unless it corrects for a substantial bias caused by deletion of non-randomly missing data.”

Table 1
Percent missingness for levels of key variables.

Variable	Variable level	Percent missing delta-V
Delta-V		32.0%
Maximum injury level in vehicle	ISS < 15	31.4%
	ISS 15+	36.8%
Number of impacts	Single impact	29.2%
	Multiple impacts	36.1%
Principal direction of impact	Front	32.8%
	Right	26.3%
	Left	28.3%
	Rear	35.5%
Vehicle type	Car	31.4%
	SUV	30.3%
	Van	30.0%
	Pickup	37.5%

It is somewhat unclear whether analysis of NASS-CDS data will benefit from multiple imputation, and more importantly, whether the resulting coefficient of delta-V will be unbiased. Few researchers have explored this important issue, but Newgard and Haukoos (2007) conducted simulations based on NASS-CDS and concluded that multiple imputation of delta-V was beneficial. In their paper, they simulated missingness of varying degrees for delta-V as a first step in understanding the potential benefits of multiple imputation. Missing delta-V values were chosen at random with no relationship to other variables. However, in NASS-CDS, missingness of delta-V is related to outcome, as well as other variables, so it is unclear whether Newgard and Haukoos' (2007) results will extend to the more complex missingness scenario. Table 1 gives the percent of missing delta-V cases for levels of several key variables: crash direction, vehicle type, injury outcome, and number of impacts. Differences in percent missing for different levels are significant for all variables in Table 1. In particular, delta-V is missing in a higher percentage of cases with severe injury, multiple impacts, rear and frontal impacts, and pickups.

The patterns in Table 1 do not mean that relationships between these variables cannot be estimated without imputation. Breslow (1996) showed that parameter estimates other than the intercept are unbiased even when the sample is biased (as it is here). The relationships in Table 1 do indicate that imputation may be warranted but would need to account for some of the effects of not-completely-at-random missingness. Imputation would be justified if it results in a decrease in standard errors of the estimates without biasing those estimates.

Following Harrell (2001), we used multiple imputation with SAS 9.2 PROC MI and MIANALYZE (SAS Institute, 2008). Missing values of delta-V were imputed using the following variables: Maximum occupant age in vehicle (continuous), maximum ISS in vehicle (continuous), vehicle type, crash direction, vehicle-level belt use (as described earlier), number of impacts, and presence of female in vehicle. We imputed 100 datasets for each analysis (i.e., missing data were filled in 100 times to produce 100 complete datasets, which were then analyzed using standard statistical procedures (logistic regression) with the results of these analyses combined to produce results used for statistical inference).

When we compared results for imputed and non-imputed datasets, we found that the coefficient of delta-V was substantially smaller for the imputed dataset. The resulting risk predictions for most cases were lower as well. We were concerned that the imputation model may have introduced additional variance to the relationship between crash severity and injury that reduced the strength of the delta-V coefficient. As a result, we elected to present the model based on the complete-case dataset (no imputation), based in part on Harrell's (2001) caveat described above. However, we consider the question of the appropriateness of imputation of delta-V in NASS-CDS to be important and still unanswered. To further that discussion, we present our imputation results (process and model) in Appendix A for reference.

3.4. Notification-case subset

In practice, EDRs are programmed to send notification only when the crash circumstances reach a certain threshold. At present, this threshold for OnStar® is either airbag deployment or a delta-V of 15 mph or more. Other systems are likely to be similar, though not identical. When these criteria are applied to the dataset, the case count is reduced to 8679, or 59% of the original sample size. However, since the 41% of cases not reaching notification criteria are necessarily low-speed impacts, the majority (85%) of injury cases remain in the notification set. The notification cases better represent the real-world task of the model, so we limited our modeling dataset to these cases.

3.5. Validation

Validation of any model is an important step in ensuring that the model is likely to perform as expected in the field. While the split-sampling approach (fitting a model to a 'training' dataset and using the model to score a 'validation' dataset) is often used to validate predictive models, this approach does not make efficient use of all of the information contained in the dataset. Split-sample validation results in the validation of a model fit to a "training" dataset, but it does not validate the model fit to the complete dataset, the objective of a predictive model. Following Harrell (2001) we used the entire sample (n = 14,861 observations) from all years (1999–2008) to develop the predictive model of serious injury outcome. Using SAS 9.2, validation of the final model was accomplished using the BVAL SAS macro developed by Gonen (2007) for bootstrap validation of the full model receiver operating characteristic (ROC) curve. This approach gives optimism-corrected estimates of the area under the ROC curve along with other model performance statistics. Optimism is defined as true performance minus apparent performance, where true performance refers to the underlying population of crash-involved occupants, and apparent performance refers to the estimated performance in the sample (Steyerberg, 2009). Correction for optimism is important because the purpose of the model is to predict serious injury risk for new subjects from the entire population of crash-involved motor vehicle occupants. Many efforts to validate predictive logistic regression models in a clinical decision-making context rely upon simple split case analysis where the dataset is divided into two separate datasets: a model development or test dataset used to estimate model parameters and their standard errors and a validation dataset produced by scoring the validation dataset with the model developed using the test data. This does not accomplish what is needed to objectively and comprehensively validate the model since only the model fit to the test data is validated while the model fit to the complete dataset is reported as "the model". The modern statistical literature emphasizes the need for computationally intensive (e.g., bootstrap or jackknife) methods of model validation (Harrell, 2001). The SAS BVAL macro developed by Gonen (2007) is one published generic algorithm for implementing a bootstrap model validation methodology. The authors are unaware of other motor vehicle epidemiologic studies that use similar model validation techniques, although the medical statistical decision making literature contains numerous examples of bootstrap validation (e.g., Steyerberg et al., 2010; Weiser et al., 2008).

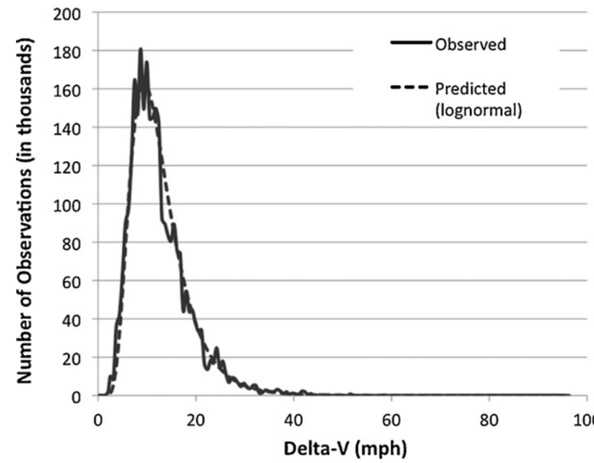

Fig. 1. Weighted distribution of delta-V (mph) for all cases. Lognormal fit is shown in dotted lines.

4. Results

4.1. Univariate relationships with injury

Before using multivariate methods, we investigated the distributions of predictors and looked at the univariate relationships between each predictor and injury outcome (serious injury in vehicle). This was done using all available cases, either in the original set or the notification subset (as indicated). NASS-CDS weights were used unless otherwise indicated.

Fig. 1 shows the distribution of delta-V for all cases, along with a lognormal fit to the distribution. The distribution is fit well by a lognormal distribution, so we used the natural log of delta-V in modeling. Fig. 2 shows the distribution of delta-V for notification cases, along with the lognormal fit. The second peak at around 15 mph reflects the shift from airbag-based notifications (with dV less than 15 mph) to delta-V-based notifications. Fig. 3 shows the relationship between delta-V and injury for all cases. Notification cases are identical from 15 mph up, so the relationship is virtually the same for these cases.

Table 2 contains the injury rates (based on weighted data) for levels of each of the categorical predictors in the model for all cases and notification cases only. Overall, the injury rate for notification cases is 2.8%. Unadjusted injury rates in Table 2 generally follow expected patterns. Higher unadjusted injury rates are associated with: multiple impacts, vehicles with any unbelted occupants, pick-

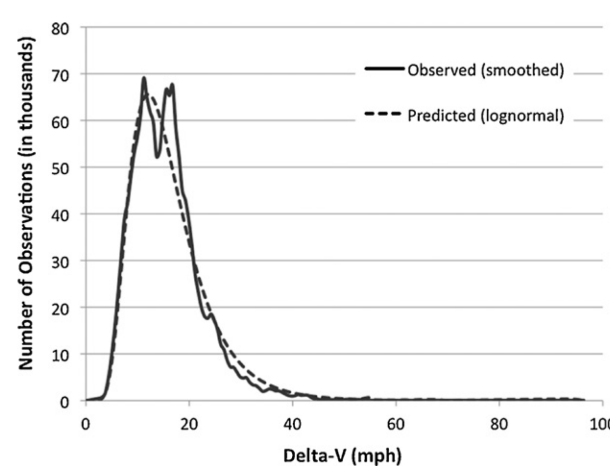

Fig. 2. Distribution of delta-V (mph) across notification cases.

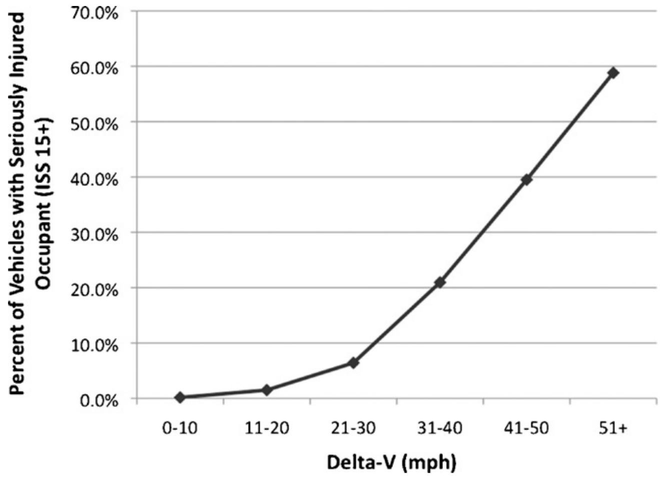

Fig. 3. Proportion of vehicles containing one or more seriously injured occupants (ISS 15+) as a function of delta-V. Only notification cases are used in this figure.

ups and cars, presence of a female occupant, vehicles with older occupant(s), and right and left impacts.

5. Modeling

5.1. Notification cases

All of the cases with missing delta-V values were deleted and the model fit to those 6625 observations that met notification criteria (delta-V $\geq$ 15 mph or airbag deployment). The Pearson residual and deviance residual outlier statistics were visually inspected, and criteria for Pearson residual, deviance residual, deviance difference, and *c*-bar were used to identify extreme outliers. One case exceeded these criteria, so the remaining number of cases used in this model was 6624. Table 3 summarizes the full model parameter estimates and associated standard errors obtained using the SAS PROC SURVEYLOGISTIC procedure. The coefficients of the vehicle-type variables were not significant ($p > 0.05$).

The following variables were used as references for their respective categories: cars for *vehicle type*, rear impacts for *direction of impact*, any unbelted for *vehicle belt use*, single for *number of events*, no for *presence of older occupants* and no for *presence of females*.

A measure of the discriminatory capability of the final model is provided by the *c*-statistic or equivalently, for binomial responses, the area under the ROC curve (AUC) as discussed in the previous section on model validation. The AUC for the model using complete case analysis is 0.85, suggesting that the model provides good discriminatory capability. To test overall goodness-of-fit, we used a variant of the Hosmer–Lemeshow test better suited for complex sample survey data (Shah and Barnwell, 2003). To do this, we divided the cases into deciles, based on predicted probability. For each decile, we computed the expected number of injury cases, based on the weighted average of predicted probability in each decile. We then compared the observed weighted number of injury cases in each decile to the expected number using a Rao–Scott chi-square, which takes into account the survey design. This resulted in a chi-square value of 12.14 (9 df), $p = 0.2060$, suggesting that the overall fit is reasonably good across the range of predicted probabilities.

Fig. 4 shows histograms of predicted probability of injury for vehicles with and without seriously injured occupants. The histograms indicate that the biggest benefit of the model is in identifying vehicles with very low probability of injury. Starting at a value of about 0.05, the percent of injury cases with predicted probability at each level is higher than the corresponding percent of non-injury cases. A vertical reference line at 0.20 is included to illustrate the Expert Panel's probability cutpoint recommendation.

Figs. 5–8 show side-by-side predicted risk for different levels of the predictors (with other predictors fixed). These graphs all use the complete-case/notification model and are meant to illustrate the nature and size of the main effects in the model. Fig. 5 illustrates the effects of age and gender on risk. As was clear from the coefficients, the age effect is substantially larger than the gender effect. For example, at 30 mph delta-V, compared to a vehicle with all young males, a vehicle with at least one female is at about 50% greater risk of having a seriously injured passenger. In contrast, a vehicle with at least one male over 55 is at 150% greater risk. Note that this risk should not be confused with the tendency of young males to be in more severe crashes. The risks of injury illustrated in Fig. 5 apply to crashes of equal severity. Table 4 summarizes the percent change in the predicted probability of severe injury risk associated with age and gender for the scenario described in Fig. 5.

Fig. 6 illustrates the effect of impact direction. The effect of direction is substantial, with left impacts producing the highest risk, followed by right, frontal, and rear impacts. It should be noted that

Table 2
Injury rates for levels of key variables.

Variable	Variable level	Percent of vehicles with seriously injured occupant(s) (notification cases only)
Number of impacts	Single impact	2.2%
	Multiple impacts	4.1%
Principal direction of impact	Front	2.1%
	Right	6.4%
	Left	8.5%
	Rear	1.6%
Vehicle type	Car	2.9%
	SUV	2.4%
	Van	1.4%
	Pickup	3.5%
Older occupant(s) (55+)	No older Occs	2.3%
	Older Occ present	5.2%
Female occupant(s)	No females	2.6%
	Female present	3.0%
Restraint use	All belted	1.9%
	Some unbelted	7.4%

Table 3
Full model results using only notification cases ($n = 6624$).

Parameter		Estimate	Standard Error	Wald chi-square	Pr > ChiSq	Odds ratio (Conf. Int)
Intercept		−15.208	0.822	341.971	<0.0001	
ln delta-V (mph)		3.603	0.329	119.718	<0.0001	36.69 (19.24, 69.95)
Direction of impact	Front impact	1.089	0.488	4.981	0.0256	2.97 (1.14, 7.73)
	Right impact	2.020	0.328	37.861	<0.0001	7.54 (3.96, 14.35)
	Left impact	2.867	0.543	27.852	<0.0001	17.58 (6.06, 50.99)
	Rear impact	0.000				
Vehicle belt use	All occupants belted	−1.450	0.227	40.975	<0.0001	0.23 (0.15, 0.37)
	Any occupants unbelted	0.000				
Vehicle type	Utility	−0.203	0.220	0.855	0.3553	0.82 (0.53, 1.26)
	Van	−1.116	0.685	2.655	0.1032	0.33 (0.09, 1.25)
	Pickup	0.167	0.411	0.166	0.6839	1.18 (0.53, 2.644)
	Car	0.000				
Number of events	Multiple	0.639	0.184	12.027	0.0005	1.90 (1.32, 2.72)
	Single	0.000				
Presence of older occupants	At least one 55+	0.991	0.200	24.523	<0.0001	2.69 (1.82, 3.99)
	No older occupants	0.000				
Presence of females	At least one female	0.450	0.174	6.715	0.0096	1.57 (1.12, 2.21)
	No females	0.000				

Table 4
Age and gender risk comparisons: Fig. 5 scenario.

Group	Compared with	% change in Pr(ISS 15+) to group
Young female	Young male	54%
Old female	Old male	49%
Old male	Young male	154%
Old female	Young female	146%

Table 5
Impact direction risk comparisons: Fig. 6 scenario.

Groups	Compared with	% change in Pr(ISS 15+) to group
Front	Rear	192%
Right	Rear	600%
Left	Rear	1375%

for a near-side occupant, right and left impacts result in similar risk. However, because this model is at the vehicle level, a left-side impact is always a near-side impact to the driver, whereas a right-side impact may or may not result in a near-side impact. From the point of view of triage, a left-side impact is a more serious crash and substantially more likely to result in serious injury. Table 5 summarizes the percentage change of the predicted probability of severe injury (ISS 15+) for frontal, left and right impacts relative to rear impacts for the crash scenario described in Fig. 6.

Fig. 7 illustrates the effect of vehicle type, which was not significant. Although the curve for vans is below the others, the effect is minor. Table 6 shows the percentage-wise changes in the predicted probability of severe injury (ISS 15+) for vans, utilities and pickups relative to cars for the crash scenario described in Fig. 7.

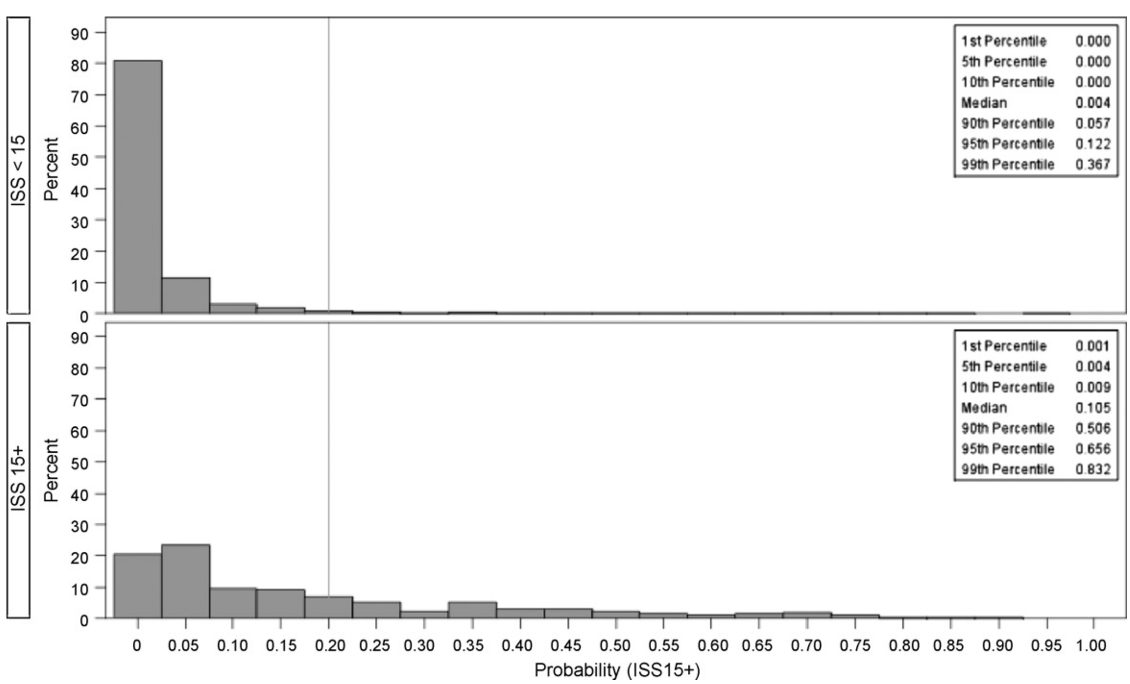

Fig. 4. Plots showing the distribution of predicted probability for vehicles without any seriously injured occupants (ISS < 15, above) and vehicles with seriously injured occupants (ISS 15+, below); $n = 6624$. A vertical line shows the recommended cutoff of 0.2.

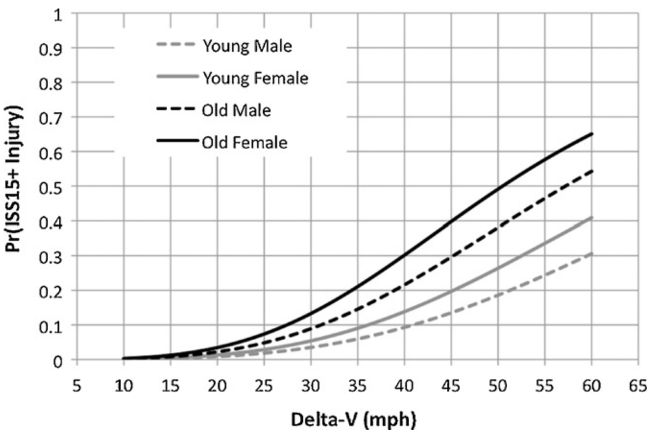

Fig. 5. Four curves showing model predictions for having an older occupant (55+) in a vehicle and having at least one female occupant on predicted risk as a function of dV (mph). For this example, other variables are fixed at: car, single impact, frontal impact, and all occupants belted.

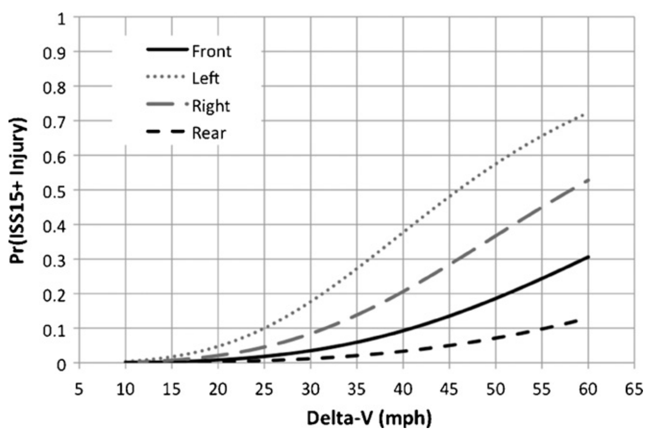

Fig. 6. Four curves showing model predictions as a function of dV (mph) by direction of impact. For this example, other variables are fixed at: car, single impact, all occupants belted, no females, and no older occupants.

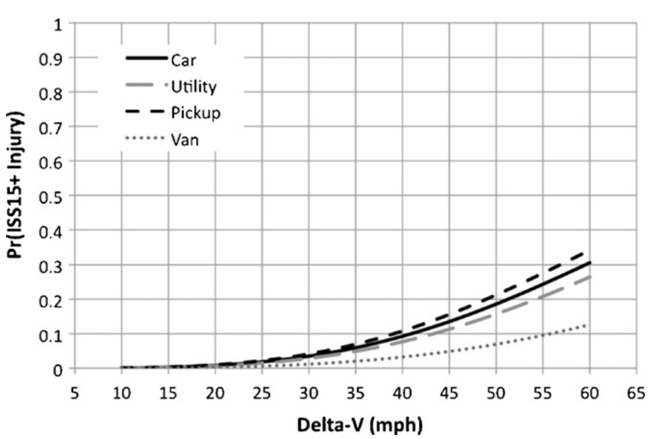

Fig. 7. Four curves showing model predictions as a function of dV (mph) by vehicle type. For this example, other variables are fixed at: single impact, front impact, all occupants belted, no females, and no older occupants.

Table 6
Vehicle-type risk comparisons: Fig. 7 scenario.

Group	Compared with	% change in Pr(ISS 15+) to group
Utility	Car	−17%
Van	Car	−66%
Pickup	Car	17%

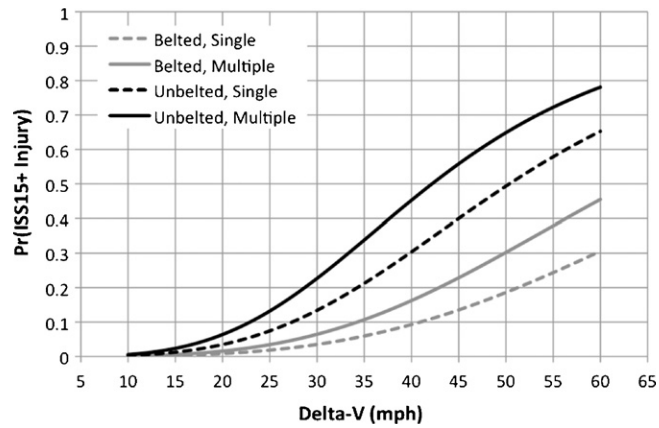

Fig. 8. Four curves showing model predictions as a function of dV (mph) by number of impacts and belt use. For this example, other variables are fixed at: car, front impact, no females, and no older occupants.

Table 7
Belt use and multiple impact risk comparisons: Fig. 8 scenario.

Group	Compared with	% change in Pr(ISS 15+) to group
Belted single	Belted multiple	−46%
Unbelted single	Unbelted multiple	−42%
Belted single	Unbelted single	−74%
Belted multiple	Unbelted multiple	−72%

Finally, Fig. 8 illustrates the effects of number of impacts and belt use in the vehicle. Although the number of impacts (single vs. multiple) is a significant predictor of injury, the presence of even a single unbelted occupant has a much greater effect on the probability that there will be a seriously injured occupant in the vehicle. For example, at 30 mph, a single impact results in 46% lower risk for a vehicle with all belted passengers relative to multiple impacts, but having all belted passengers decreases the risk of serious injury in a multiple-impact vehicle by about 72%. Table 7 summarizes the percent change in predicted probability of severe injury (ISS 15+) associated with age and gender for Fig. 8 crash scenario.

6. Validation

Since we want to maximize the model-development sample size, we chose to use bootstrap validation to avoid the problems of the split-sample approach (Gonen, 2007; Harrell, 2001). Table 8 summarizes the estimates of three common model performance statistics: the area under the receiver operating characteristic curve (AUC), the Somer's "*D*" Statistic (*DXY*), and an adjusted *R*-square measure (Nagelkerke's R Square) for logistic regression. For binary outcomes, the AUC and *DXY* statistics are related by: AUC = $DXY/2 + 0.5$ and, as such, are redundant measures of model performance. A logistic model with an optimism-corrected AUC of greater than 0.80 has a reasonable capability of discriminating between the two states of the binary outcome variable, ISS 15+ (Harrell, 2001).

7. Decision analysis and performance

Ultimately, the most important measure of the performance of this decision algorithm is how well it distinguishes between vehicles with seriously injured occupants and those without. Although the algorithm produces a continuous value for probability of injury, in practice, a cutpoint must be chosen to decide when to take action (e.g., alert EMS to high probability of injury, move triage priority higher, or transport to Level 1 trauma center). Table 9 shows the sensitivity (percent of positive tests out of all injury cases) and

Table 8
Optimism-corrected model validation statistics using) approach.

Estimate	Optimistic estimate	Optimism correction	Optimism-corrected estimate
AUC	0.8433	0.0027	0.8406
DXY	0.6865	0.0053	0.6812
Adjusted *R*-square	0.3360	0.0050	0.3310

Table 9
Sensitivity and specificity for full model by hypothetical probability cutpoints.

Probability cutpoint	Sensitivity	Specificity	True positives per 1000 notifications	False positives per 1000 notifications	False negatives per 1000 notifications
0.05	0.717	0.888	20.08	108.86	7.92
0.1	0.547	0.947	15.32	51.52	12.68
0.2	0.396	0.983	11.09	16.52	16.91
0.3	0.282	0.992	7.90	7.78	20.10
0.4	0.204	0.995	5.71	4.86	22.29
0.5	0.150	0.998	4.20	1.94	23.80

Table 10
Average maximum vehicle ISS for four decision outcome categories.

Probability cutpoint	Average ISS for true positives	Average ISS for false negatives	Average ISS for false positives	Average ISS for true negatives
0.1	31.1	27.6	3.8	1.5
0.2	31.9	28.0	4.7	1.6
0.3	33.8	27.8	4.9	1.6

specificity (percent of negative tests out of all non-injury cases) for a variety of possible cutpoints.

While sensitivity and specificity are key numbers for evaluating a decision algorithm, they do not take into account the relative number of injury and non-injury cases that will be encountered in the field. In the notification dataset, the overall probability of injury is about 2.8%. Thus, there are about 35 times as many non-injury cases as injury cases. Put another way, there are 35 times more opportunities for false positives (cases that are given unnecessary treatment) than true positives (cases that need treatment and are detected). Higher sensitivity always goes with lower specificity and results in more true positives at the expense of more false positives. The three rightmost columns put the sensitivity and specificity of each cutpoint in the context of the low base rates of injury. They indicate the number of true positives, false positives, and false negatives (missed opportunities) per 1000 vehicle notifications. These numbers should most closely reflect the expected experience in the field.

For example, out of 1000 vehicle notifications, about 2.8%, or 28 vehicles are expected to include a seriously injured occupant. With a probability cutpoint of 0.2, sensitivity is 0.396, meaning that about 40% of these, or approximately 11 will result in a predicted probability of serious injury above 0.2. These are the true positives, and the other 17 will result in false negatives. Of the same 1000 notifications, 972 will not contain a seriously injured occupant. With a cutpoint of 0.2, the specificity is 0.983, meaning that about 1.7%, or 16.52 vehicles, will result in false positives.

Another way of putting the model performance into context is to look at the average maximum ISS in a vehicle for each category of decision. The algorithm is based on a two-category concept of injury. However, a false positive that flags a vehicle containing an occupant with ISS 14 would be considered far less problematic than a false positive for a vehicle containing no occupants with any injury at all. Similarly, missing a vehicle with an occupant with ISS 30 is a “worse” miss than missing a vehicle with its worst-case ISS of 16. To illustrate this, we compared cutpoints of 0.1, 0.2 (recommended by the CDC) and 0.3.

Table 10 shows the average maximum vehicle ISS for the four categories of decision and outcome for each of three cutpoints. The key comparisons are between groups with the same outcome and different decisions. For example, true positives and false negatives encompass vehicles with one or more severely injured occupants. In Table 10, the average ISS for true positives is consistently higher than that for false negatives. In other words, the injury level of misses is lower than those that are picked up by the algorithm. Similarly, false positives and true negatives encompass vehicles without a seriously injured occupant. For these, the average maximum ISS is higher for the false positives, compared to true negatives. Although the difference between these averages (in both cases) is somewhat small (3 ISS units), the direction of effect indicates that the more serious cases get identified more readily by the algorithm and the least serious cases are more readily rejected.

8. Conclusions

We have presented results of logistic regression analyses to predict the probability of a serious injury in a crash-involved vehicle, following the approach laid out by the CDC Expert Panel on Field Triage. These analyses are based on the information that may be obtained using an EDR in a crash, or by an operator communicating with vehicle occupants immediately following a crash (age and gender).

The results of these analyses are promising for the possibility of initiating triage decisions using EDR-based crash information. The AUC of 0.84 indicates significant discrimination by the algorithm, though there is room for improvement. The key predictors in this model are delta-V (log transformed), belt use, age (presence of anyone 55 or older), and direction of impact. Additional limited predictive value comes from multiple impacts and females present in the vehicle. Vehicle type was not a significant predictor, though it was left in the model to match the predictors given by the expert panel.

8.1. Summary

We developed a triage prediction model based on NASS-CDS data from 1999 to 2008 and following the guidelines laid out by the CDC Expert Panel. Analysis was done at the vehicle level, rather than the occupant level because EMS responds to a vehicle as a unit (though treatment is ultimately at the individual level). The dependent measure was a binary variable indicating whether any occupant of a vehicle had an Injury Severity Score (ISS) of 15 or greater. Predictors included crash severity (delta-V), impact direction, vehicle type, belt use (coded as all belted vs. anyone unbelted), number of impacts (single vs. multiple), age (coded as all under 55 vs. anyone over 55), and gender (coded as all male vs. anyone female).

Logistic regression (accounting for NASS weights) was used to predict the probability of a vehicle containing an occupant with serious injury. The model fit reasonably well, and is particularly effective at weeding out many of the cases with the least potential for injury. The model itself predicts a continuous value of probability of injury, so it is up to others to choose a cutpoint that would be implemented in practice to differentiate between vehicles that meet criteria and vehicles that do not. We have presented some numbers to give context to different cutpoint choices. For example, a cutpoint of 0.1 will result in more false positives and fewer false negatives than a cutpoint of 0.3. However, the relative importance of these depends on costs that are outside the scope of this paper. Ultimately, the model shows promise for using the EDR to provide a fast first cut at identifying vehicles in injury-producing crashes.

8.2. Scope of crash problem

The selection criteria used to obtain the study vehicle population (airbag deployed or delta-V > 15, model year 2000+, planar crash) resulted in 8679 observations, of which approximately 11.8% or 1026 vehicles had occupants with a reported ISS of 15 or greater. When weighted, these represent 2.8% of the population of vehicles in crashes meeting notification criteria. Annually, approximately 1.8 million vehicles are involved in planar crashes of at least 15 mph delta-V. Thus, the problem of triage – finding the 2.8% of these that need the most help – is significant.

8.3. Non-significant predictors

Vehicle type was not a significant predictor ($p > 0.05$) of ISS 15+ injury. The presence of one or more females was a significant ($p < 0.05$) but weak predictor. Models fit without the Vehicle type and presence of one or more female variables were similar, in terms of such performance measures as AUC, sensitivity and specificity, to models that included these variables. From a statistical perspective, it makes little sense to include predictor variables that do not significantly contribute to the predicted outcome; however, these variables were left in the final predictive model to satisfy the preliminary requirements of the Expert Panel's Recommendations.

8.4. Link from WinSmash to EDR

An important, but unverified, assumption in developing our predictive injury risk model is that the delta-V values for the principal direction of force recorded in the NASS-CDS dataset correspond to delta-V values for the principal direction of force recorded by recent model-year EDR systems. While previous work to correlate EDR-recorded delta-V values with measured delta-V values obtained from full frontal barrier tests indicates general agreement between the two measurement systems (Gabler et al., 2003; Niehoff and Gabler, 2006), additional studies are needed to more fully explore the relationship between NASS-CDS and EDR derived delta-V values. Significant biases or discrepancies between the two systems would require a statistical adjustment in the delta-V parameter.

8.5. Choice of cutpoint

The CDC Expert Panel's initial recommendations for development of an injury risk algorithm specified a probability cutpoint of 0.2 to identify cases that warrant consideration as possible severe occupant injuries. This pre-specified probability cutpoint ignores both the data and the final model. Decision cutpoints should be data- and model-driven and capture the relative costs of both false positives and false negatives. An arbitrary probability cutpoint, even one specified from previous clinical experience to minimize the probability of false positives, does not make good statistical sense. Determination of the relative costs of both false positives and false negatives should be decided by expert clinicians, public health decision makers and other appropriate subject matter experts.

Our approach in presenting the results of the final predictive model is to provide a summary of the relative numbers of both types of errors associated with each of a range of possible decision cutpoints. We recommend that the final determination of which cutpoint to implement should be made by the public health decision maker community.

8.6. Future—data for recalibration

The current model, derived from NASS-CDS data, is an initial effort to predict possible occupant severe injury risk in crashes that exceed specific selection criteria (i.e., airbag deployed or delta-V greater than 15 mph). The model is contingent upon the assumption that each predictor variable in the NASS-CDS database can be captured by the vehicle EDR system and possible additional information that may be obtained via voice communication with crash-involved vehicle occupants. This initial model needs to be refined and continually recalibrated to reflect real world information. In particular, it would be useful to have EDR information fully incorporated into the NASS-CDS database. At present, unprocessed EDR reports are only available for some cases.

8.7. Future—role of ACN in triage

One of the critical issues to resolve in incorporating the EDR into the triage process is deciding what role the information and any algorithm should play. At present, information about the state of the vehicle and crash comprise the third tier of triage rules. Occupants who do not meet physiological (e.g., consciousness and blood pressure) or anatomical (e.g., broken bones) criteria for triage can be transported solely on the basis of having been in a severely damaged vehicle. These occupants may have sustained "silent/occult" internal injuries and should be evaluated by a physician as a precaution.

The EDR-based algorithm can play this same role, providing EMS with a fast, automated estimate of the potential for injury based solely on what happened to the vehicle. While this may be convenient, the greatest potential benefit of an EDR-based triage algorithm is in the speed with which the information is transmitted to dispatch. If the EDR-based algorithm is treated as Step 0, i.e., a factor in the initial allocation of emergency resources, it could provide substantial benefit in time savings and resource usage.

Appendix A. Multiple imputation

The SAS PROC MI and PROC MIANALYZE procedures were also used to derive imputed delta-V values and analyze the combined results obtained by multivariate logistic regression analysis of the imputed full and notification datasets. Four outliers were removed from the all-cases analysis and one outlier was removed from the notification analysis.

Table A1
Comparison of complete-case and MI analyses.

Parameter		Notification cases (n = 6624)		Notification cases MI (n = 8679)	
		Estimate	Standard error	Estimate	Standard error
Intercept		−15.208	0.822	−12.809	0.845
ln delta-V (mph)		3.603	0.329	3.016	0.244
Direction of impact	Front impact	1.089	0.488	0.440	0.498
	Right impact	2.020	0.328	1.451	0.468
	Left impact	2.867	0.543	2.181	0.605
Vehicle belt use	All occupants belted	−1.450	0.227	−1.395	0.155
Vehicle type	Utility	−0.203	0.220	−0.175	0.188
	Van	−1.116	0.685	−0.744	0.570
	Pickup	0.167	0.411	0.232	0.260
Number of events	Multiple	0.639	0.184	0.414	0.224
Presence of older occupants	At least one 55+	0.991	0.200	0.951	0.160
Presence of females	At least one female	0.450	0.174	0.434	0.195

Table A1 shows the comparison between complete-case analyses and analyses of the imputed dataset for notification models. Note that the MI standard errors are not smaller than the complete-case analysis on average. In addition, both MI coefficients of log delta-V are substantially lower than the complete-case coefficients.

Since multiple imputation appears to reduce estimates of predicted risk across the board, we were concerned that this might be the result of bias due to variance introduced in the imputation model. In implementation, this model would tend not to reach a given criterion as readily as the corresponding complete-case model. In addition, the notification-data model should provide more accurate estimates in the range of cases that are relevant in practice. Since nearly 40% of the full dataset falls in a delta-V range below notification, the full-data model will tend to be tuned to injury risk in that (lower) portion of the curve. As a result of these observations, we chose to focus all further investigation and validation on the notification-data/complete-case model (Table 3). However, we recommend further study of the potential benefits and consequences of using multiple imputation with NASS-CDS data.

References

American College of Surgeons, 2006. Resources for the Optimal Care of the Injured Patient. American College of Surgeons, Chicago, IL.

Augenstein, J., Digges, K., Perdeck, E., Stratton, J., Bahouth, G., Baur, P., Messner, G., 2007. Application of ACN data to improve vehicle safety and occupant care. In: Proceedings of the 20th International Technical Conference on the Enhanced Safety of Vehicles (ESV), Lyon, France, June 18–21.

Bahouth, G.T., Digges, K.H., Bedewi, N.E., Kuznetsov, A., Augenstein, J.S., Perdeck, E., 2004. Development of URGENCY 2.1 for the prediction of crash injury severity. Topics in Emergency Medicine 26 (2), 157–165.

Breslow, N.E., 1996. Statistics in epidemiology: the case–control study. Journal of American Statistical Association 91 (433), 14–28.

Champion, H.R., Copes, W.S., Sacco, W.J., Lawnick, M.M., Keast, S.L., Bain Jr., L.W., Flanagan, M.E., Frey, C.F., 1990. The Major Trauma Outcome Study: establishing national norms for trauma care. Journal of Trauma 30, 1356–1365.

Farmer, C.M., 2003. Reliability of police-reported information for determining crash and injury severity. Traffic Injury Prevention 4, 38–44.

Gabler, H.C., Hampton, C., Roston, T.A., 2003. Estimating crash severity: can event data recorders replace crash reconstruction? In: Proceedings of the 18th International Technical Conference on the Enhanced Safety of Vehicles, DOT HS 809 543, May.

Gonen, M., 2007. Analyzing Receiver Operating Characteristic Curves with SAS. SAS Institute, Inc., Cary, NC.

Grossman, M.D., Miller, D., Scaff, D.W., Arcona, S., 2002. When is an elder old? Effect of preexisting conditions on mortality in geriatric trauma. Journal of Trauma 52, 242–246.

Harrell, F.E., 2001. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer Science and Business Media, Inc., New York, NY.

Hours, M., Bernard, M., Charnay, P., et al., 2010. Functional outcome after road-crash injury: description of the ESPARR victims cohort and 6-month follow-up results. Accident Analysis and Prevention 42 (2), 412–421.

Little, R.J.A., Lewitzky, S., Heeringa, S., Lepkowski, J., Kessler, R.C., 1997. An assessment of weighting methodology for the National Comorbidity Study. American Journal of Epidemiology 146, 439–449.

MacKenzie, E.J., Rivara, F.P., Jurkovich, G.J., Nathens, A.B., Frey, K.P., Egleston, B.L., Salkever, D.S., Scharfstein, D.O., 2006. A national evaluation of the effect of trauma-center care on mortality. New England Journal of Medicine 354, 366–378.

Morris, J.A., MacKenzie, E.J., Edelstein, S.L., 1990. The effect of preexisting conditions on mortality in trauma patients. Journal of the American Medical Association 263, 1942–1946.

National Highway Traffic Safety Administration (NHTSA), 2007. National Automotive Sampling System (NASS) Crashworthiness Data System Analytical User's Manual 2007 File. National Center for Statistics and Analysis, NHTSA, U.S. Department of Transportation, Washington, DC.

Newgard, C.D., Haukoos, J.S., 2007. Advanced statistics: missing data in clinical research. Part II. Multiple imputation. Academic Emergency Medicine 14, 669–678.

Niehoff, P., Gabler, H.C., 2006. The accuracy of WINSMASH delta-V estimates: the influence of vehicle type, stiffness and impact mode. Annual Proceedings of the Association for the Advancement of Automotive Medicine 50, 73–89.

Potter, F., 1990. A study of procedures to identify and trim extreme sample weights. In: Proceedings of the Survey Research Methods Section. American Statistical Association, pp. 225–230.

Rauscher, S., Messner, G., Baur, P., Augenstein, J., Digges, K., Perdeck, E., Bahouth, G., Pieske, O., 2009. Enhanced automatic collision notification system—improved rescue care due to injury prediction—first field experience. In: Proceedings of the 21st International Technical Conference on the Enhanced Safety of Vehicles (ESV), Stuttgart, Germany, June 15–18.

Robertson, R., Vanlaar, W., 2008. Elderly drivers: future challenges? Accident Analysis and Prevention 40 (6), 1982–1986.

Rowe, S.A., Sochor, M.S., Staples, K.S., Wahl, W.L., Wang, S.C., 2004. Pelvic ring fractures: implications of vehicle design, crash type, and occupant characteristics. Surgery 136 (4), 842–847.

Sampalis, J.S., Nathanson, R., Vaillancourt, J., Nikolis, A., Liberman, M., Angelopoulos, J., Krassakopoulos, N., Longo, N., Psaradellis, E., 2009. Assessment of mortality in older trauma patients sustaining injuries from falls or motor vehicle collisions treated in regional level I trauma centers. Annals of Surgery 249 (3), 488–495.

SAS Institute Inc., 2008. SAS/STAT User's Guide, SAS 9.2. SAS Institute Inc., Cary, NC.

Sasser, S.M., Hunt, R.C., Sullivent, E.E., Wald, M.M., Mitchko, J., Jurkovich, G.J., Henry, M.C., Salomone, J.P., Wang, S.C., Galli, R.L., Cooper, A., Brown, L.H., Sattin, R.W., 2009. Guidelines for field triage of injured patients. Recommendations of the National Expert Panel on Field Triage. Morbidity and Mortality Weekly Report 58 (RR-1), 1–36.

Schiff, M.A., Tencer, A.F., Mack, C.D., 2008. Risk factors for pelvic fractures in lateral impact motor vehicle crashes. Accident Analysis and Prevention 40 (1), 387–391.

Shah, B.V., Barnwell, B.G., 2003. Hosmer–Lemeshow goodness of fit test for Survey data. In: Joint Statistical Meetings—Section on Survey Research Methods.

Steyerberg, E.W., 2009. Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating. Springer Science and Business Median, Inc., New York, NY.

Steyerberg, E.W., Vickers, A.J., Cook, N.R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M.J., Kattan, M.W., 2010. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 21 (1), 128–138.

Tavris, D.R., Kuhn, E.M., Layde, P.M., 2001. Age and gender patterns in motor vehicle crash injuries: importance of type of crash and occupant role. Accident Analysis and Prevention 33 (2), 167–172.

Weiser, M.R., Landmann, R.G., Kattan, M.W., Gonen, M., Shia, J., Chou, J., Paty, P.B., Guillem, J.G., Temple, L.K., Schrag, D., Saltz, L.B., Wong, W.D., 2008. Individualized prediction of colon cancer recurrence using a nomogram. Journal of Clinical Oncology 26 (3), 380–385.