PROPABILISTIC LINKAGE AND ANALYSIS OF ARIZONA'S CRASH DATA AND TRAUMA REGISTRY

Site: Arizona Department of Health Services

Student Name: Erin Campbell

Completion Date: May 1, 2013

I do ☒ do not☐ give permission for this report to be made available for review.

Student: Erin Campbell

Signature: _____

APPROVAL BY INTERNSHIP COMMITTEE CHAIR

This internship report has been approved on the date shown below:

NAME OF INTERNSHIP COMMITTEE CHAIR: Dr. Denise Roe DATE:

_____

SIGNATURE OF INTERNSHIP COMMITTEE CHAIR

THE UNIVERSITY OF ARIZONA®
MEL AND ENID ZUCKERMAN ARIZONA COLLEGE OF PUBLIC HEALTH

STATEMENT BY AUTHOR

This internship report has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the Office of Student Services and Alumni Affairs at the Mel and Enid Zuckerman College of Public Health to be made available to borrowers under the rule of the College.

In submitting this report, I do give permission for this report to be read by others.

Brief quotations of this internship report are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission of extended quotation from or reproduction of this manuscript in whole or in part may be granted by the author.

STUDENT NAME: Erin Campbell

STUDENT SIGNATURE: _____

Acknowledgements

I would like to thank my committee, Dr. Denise Roe and Dr. Angelika Gruessner. I would also like to thank Dr. David Harden and the rest of the faculty and staff at the Arizona Department of Health Services. I would also like to thank the staff at the Arizona Department of Transportation.

Table of Contents

Abstract - Page 7

Background - Page 9

Part 1: Probabilistic Linkage

Introduction - Page 11

Methods - Page 14

Results - Page 16

Discussion – Page 18

Part 2: Trauma versus Non-Trauma Crashes

Introduction - Page 20

Methods - Page 21

Results - Page 27

Discussion - Page 36

Recommendations and Future Study – Page 38

Conclusion – Page 41

Appendix I – Page 43

Appendix II – Page 48

Appendix III – Page 52

References – Page 56

List of Figures and Tables

Percent Missing - Page 17

Number of Matches and Non-Matches – Page 17

Distribution of Scores – Page 17

Airbag Usage (Front Row Passengers Only) – Page 25

Drug and Alcohol Use (Drivers Only)- Page 26

Variable Characteristics- Page 27

Variable Characteristics by Outcome for Passenger Vehicles- Page 29

Trauma vs Non-Trauma by Collision Manner- Page 30

Logistic Regression Model 1 – Page 31

Variable Characteristics by Outcome for Motorcycles – Page 32

Logistic Regression Model 2- Page 33

ROC Curve for Model 1- Page 34

ROC Curve for Model 2- Page 34

Body Area Injured by Collision Manner- Page 35

% of people with Multiple Body Part Injuries- Page 35

Abstract

Motor vehicle collisions (MVC) are responsible for billions of dollars in damages and medical expenses every year. In Arizona, information on the nature of the crash and resulting injuries are captured by two separate non-compatible databases, the Accident Location Identification Surveillance System (ALISS) and the Arizona State Trauma Registry (ASTR) managed independently by the Arizona Department of Transportation (ADOT) and the Arizona Department of Health Services (ADHS), respectively. The information contained within these two databases refers to a singular event.

This internship had two goals, to probabilistically link the two databases and to look at the factors of a MVC that contribute to passengers being traumatically injured. The two databases were probabilistically linked utilizing age, gender, date and time as linking variables. 60.6% of the records in ASTR were successfully matched to ALISS reports. Logistic regression was used since the data was not measured over time and the outcome was binary; whether the persons involved in the MVC will suffer a traumatic injury or not. The exposure variables used were collision manner, estimated speed, seat belt usage, seat position and whether a person had to be extricated from their vehicle. The results showed that not wearing a seat belt and being extricated from a vehicle increases one's risk of being traumatically injured. Not wearing a helmet increases the odds of a person on a motorcycle being injured when involved in a MVC. Injuries were classified into 6 major categories head, face, chest, abdomen, extremity and external. It was shown that

regardless of collision type, external injuries were the most common injury type for those that were traumatically injured.

Background

A study done in 2010 showed that motor vehicle collisions (MVCs) nationwide were responsible for nearly $100 billion in wage and productivity losses, medical expenses, administrative expenses, motor vehicle damage, and employer costs in 2005, the latest year that data was available.[1] Arizona had losses of almost $2.7 billion in 2010 and almost $2.9 billion in 2011.[2,3] Information on the nature of the crash and resulting injuries has been captured for years by law enforcement and public health agencies, respectively. Often times the two agencies maintain and manage their records in two separate non-compatible databases.

An initial study by the National Highway Traffic Safety Administration (NHTSA) found that crash data alone provided insufficient information on medical and financial outcomes of MVCs and led a push for states to develop Crash Data Outcome Evaluation Systems (CODES).[4] Starting in 1992, NHTSA began funding states to conduct probabilistic linkages in order to better study patient outcomes in relation to motor vehicle crashes.[5] Previously with funding from the NHTSA and CODES, Arizona linked MVC reports with hospital discharge and emergency department records.[6]

Probabilistic linkage has become an increasingly prevalent tool to link two separate but related data files. By combining two separate data files, one is able to obtain more information about a singular event. Probabilistic linkage works by comparing variables that are common in both databases. The linking procedure is most successful when there are variables present that uniquely differentiate one individual from another, such as name, Social Security number or date of birth. More general variables, such as age and

gender may also be used. [7] The relative success of the linking procedure is dependent upon the uniqueness of the variables being used and also on the number of errors present in the data fields. If insufficient information is present, the probabilistic linkage will result in linking errors. These errors may result in the prevalence of the event under investigation to be overestimated or underestimated. [8] Because of the *Health Insurance Portability and Accountability Act (HIPAA) of 1996*, as amended, access to unique identifiers such as names may be restricted. It is possible that enough unique information may be present in the database without the addition of names to confidently link the two records.

The goal is to merge a database of traffic-related crashes with the record of Emergency Medical Services (EMS) and trauma care. The former is managed by the Arizona Department of Transportation (ADOT) and the latter is managed by the Bureau of EMS and Trauma Systems (BEMSTS) within the Arizona Department of Health Services (ADHS). The two databases contain records referring to singular events. By establishing a method to link specific data elements of the two databases, the two agencies will be able to study MVCs where injuries or loss of life have occurred and identify new associations between the data that will prompt better or new preventative measures and treatment modalities that help reduce mortality and morbidity. This linkage may also provide measureable data to support new or revisions of state and/or local highway safety laws.

Part 1: Probabilistic Linkage of the Arizona State Trauma Registry (ASTR) and the Accident Location Identification Surveillance System (ALISS)

Introduction

In calendar years 2010 and 2011 more than 580,000 individuals occupying 406,789 vehicles involved in 210,382 crashes were reported to ADOT as crash reports. These crash reports generate a database called the Accident Location Identification Surveillance System (ALISS). Of the 580,000 individuals involved in MVCs, only 129,761 were documented as having a possible injury. These injured individuals were in 107,328 vehicles involved in 85,623 crashes.

In order to have accurate and consistent linkage with the Arizona State Trauma Registry (ASTR), the completeness and accuracy of the ALISS database are important. Law enforcement officers complete the Arizona Crash Report (ACR), which is then submitted to ADOT for entry into the ALISS database. The 170 data elements entered into the ALISS database are only a sample of what is included on the ACR. However, unique identifiers such as name or driver license number are not input into the database. Birth dates are input but the output only provides the age in years.

After a MVC occurs, a 911 call is placed and law enforcement and potentially EMS are dispatched to the scene. Any accident occurring on public roads where a police report is completed is entered into the ALISS database. Information on the crash is collected at the crash scene on the Arizona Crash Report (ACR) (Appendix 1). The ACR contains information about crash victim demographics, information on whether injuries are

involved, information about the vehicles and the nature of the collision. The ACR is a paper form that is manually entered into ALISS by ADOT staff, with the exception of ACRs submitted by the Arizona Department of Public Safety (DPS), which are submitted electronically to the ALISS database.

If EMS is not initially sent to the scene, they may still be called if after arrival law enforcement deems their response necessary. EMS assesses and treats any occupants of the vehicles who are injured. Medically-related details about injured persons are documented on a patient care report (PCR) by EMS crews at the crash scene. Individuals qualifying as major trauma patients due to the severity of injuries and/or the nature of the crash (e.g., rollover, compartment intrusion of 12 inches or greater) are transported by EMS to a state trauma center.

The trauma center emergency department (ED) generates a separate trauma record documenting patient medical history, injuries sustained in the crash, treatment administered and vital signs obtained by paramedics on scene, test results, treatment, and final outcome. Trauma center registrars enter trauma data into their respective hospital's trauma registry. These trauma records are subsequently transmitted electronically to the ASTR on a quarterly basis.

To qualify for entry into the ASTR, individuals must meet inclusion criteria specified in Title 9, Ch.25, Article 14, § R9-25-1402(A) of the Arizona Administrative Code (A.A.C.). The ASTR captures all persons that enter the emergency department that meet the inclusion criteria regardless of mechanism of injury.

To best facilitate matching, only ASTR records that indicate a motor vehicle incident as the mechanism of injury were selected. This resulted in 23,787 records for 2010 and 2011 combined. The ASTR captures a large number of fields and only those fields of interest were selected. Unlike ALISS, names and dates of birth are recorded for all trauma patients in the ASTR. Due to the limitations on the unique identifiers from the crash data, these unique identifiers were not included in the data set.

Methods

There are two methods that can be employed when trying to match records within two different databases, deterministic and probabilistic. In order to use deterministic, a verifiable unique identifier must be available for every record.[9] Probabilistic matching, however, is utilized when a unique identifier is not available or when the reliability of a unique identifier is not well recorded. In the case of the ALISS and ASTR databases, there is only a unique identifier present in the latter; therefore deterministic matching is not currently an option.

In order to utilize probabilistic matching, it was necessary to determine the variables that would be used to match the records, the probability weights of those matches and ultimately the linkage decision. To facilitate the speed of the matching, variables were selected as blocking variables so that pairs of records had an initial amount of similarity before determining the probability of a match. Given the limited number of common fields, the two records were blocked on age and gender. In order to allow for some errors in recording, multiple passes were made using each blocking variable. This was to minimize the number of true links that may have gone unmatched.

In order to determine if the record pairs represent a match, it is necessary to calculate the probability weights based on the founding principles of frequency ratios.[10] These scores are calculated via a computer algorithm, in this case by the Centers for Disease Control and Prevention (CDC) program Link Plus. Each variable designated as a part of the matching algorithm is given a weighted score that represents the probability that the records belong to the same person. This is achieved through the ‘m probability’ and the

‘u probability’. The ‘m probability’ is a measure of the probability that an identifying variable will agree for a pair of records that truly belong to the same person. The probability that a variable will randomly match across two records is called the ‘u probability’ and is based on the prevalence of a value in each variable.

A variety of positive and negative scores will be obtained from the varying levels of agreement and disagreement across two records. The set of weights for a record is summed to give an overall weight for the record pairing. This ‘score’ is used to determine whether or not the records represent a link or a non-link. The distribution of the probability weights is normally bimodal.[11] The first peak reflects pairings that reflect incorrect matches while the second peak reflects correct matches. The space between the two peaks is a grey area containing matches and non-matches that need to be manually reviewed to determine their true identity. Manual review is necessary in order to determine the cut off weights for matches and non matches. The goal of the cut off points is to maximize sensitivity and specificity.

After utilizing Link Plus to calculate the ‘m’ and ‘u’ probabilities and the weighted scores, all other data manipulations were done using SAS 9.3 (SAS Institute Inc., Cary, NC).

Results

The ALISS and the ASTR databases were matched on date of incident, age, gender and hour of injury utilizing Link Plus. There were a number of records from each data set that were missing variables. (Table 1) A disagreement because of a missing score resulted in a weighted score of zero for that variable as opposed to a positive score for a match and a negative score for a mismatch. A plot of the weighted scores for all linked pairs follows the previously described distribution of a probabilistic match with two peaks. (Figure 1) There is a large peak at 9.6 and a smaller peak at 11.9. The cut point for matches was set at 11.8 because there were no scores between 11.1and 11.7. This resulted in 10,982 records being linked. These records were manually checked for accuracy. Of the 10,982 records, 729 records were considered non-matches because the dates did not match on year. This left 10,253 records as linked, matched pairs before de-duplication.

The remaining 12,805 unlinked records were analyzed for additional probable matches as a result of the high number of missing variables throughout both data sets. There is also the possibility of the hour of injury may be slightly skewed as a result of travel time for the scene of the accident to the admitting trauma center. Therefore a ± 1 hour buffer was allowed as match for the hour of injury. For records that were below the 11.8 cutoff, records that were considered unlinked but had a missing age, gender or time variable in the pair, they were considered matches. This resulted in an additional 4,967 records being considered matches.

Overall 15,220 pairs of records were matched. After de-duplicating the matches, retaining the match with the higher weighted score, the final number of matched pairs was 14,412.

(Table 2) This represents 60.6% of the trauma records that indicated motor vehicle incident as the mechanism of injury. These results yield a positive predictive value (PPV) of 91.6% while the negative predictive value (NPV) is 66%. The specificity of the matching definitions is 90.2% while the sensitivity of the methodology in this case is 69.8%.

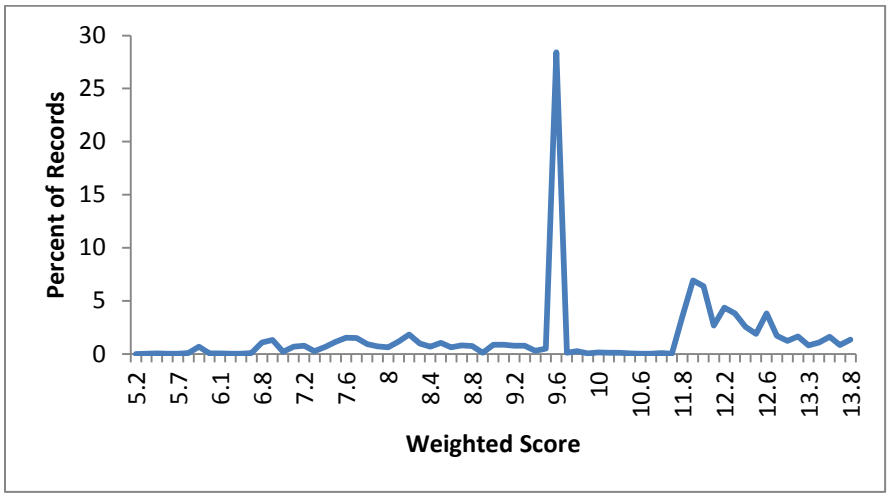
Table 1: Percent Missing

	ALISS (N=129,761)	ASTR (N=23,787)
Date	0 (0%)	0 (0%)
Age	13,462 (10.4%)	1 (<0.001%)
Gender	14,005 (10.8%)	0 (0%)
Hour of Injury	33,191 (25.6%)	7,101 (29.9%)

Table 2: Number of Matches and Non-Matches

	Matches	Non Matches
Linked	10,060	922
Unlinked	4,352	8,453
Total	14,412	9,375

Figure 1: Distribution of Scores



Discussion

Ideally, all trauma patients injured in motor vehicle incidents would be successfully matched to a crash report. The lack of unique identifiers in the crash data makes this a difficult goal to achieve. The specificity and the sensitivity of the matching are fairly high representing a successful matching schematic. However without unique identifiers such as names or date of birth, confidence in the accuracy of the matching is not high. The addition of a unique identifier is needed to fully validate the linkage. At present, names are physically recorded on the ACR but are not input into the ALISS database. Birth dates are entered but only the age is reported back out. Both of these fields would allow for matching at a higher efficiency and confidence. These fields are currently not available in ALISS because of confidentiality laws.

Merging crash records (ALISS) with the trauma records (ASTR), the respective start and end points of a particular crash incident, is a shared goal of ADHS and ADOT. The middle point in a crash incident is the EMS agency creating a PCR and transporting the injured from the crash site to a hospital for treatment. By adding EMS PCR information to the merged files, it may be possible to fill in the gaps of missing information and allow for more complete records. All trauma records indicate the reporting facility. In the ALISS database, there is a “transported to” field. This field is either missing data or incorrectly populated. Previously, it was instructed that this field was to be completed with the facility ID. This initiative has not been consistently followed, resulting in this field containing names and addresses in addition to facility ID. By improving the data entry consistency and accuracy of the transport to field, an additional field may be added

to the linkage that may improve the number of records linked together. (A more detailed look at missing data from ALISS is provided in Appendix II.)

Part 2: Trauma causing crashes versus non-trauma causing crashes

Introduction

By linking the records contained in the Accident Location Identification Surveillance System (ALISS) managed by the Arizona Department of Transportation (ADOT) with the records maintained in the Arizona State Trauma Registry (ASTR) managed by the Bureau of Emergency Medical Services and Trauma Systems (BEMSTS) within the Arizona Department of Health Services (ADHS) more information is available on the nature of motor vehicle collisions (MVC) and safety devices utilized by the MVC occupants injured severely enough to qualify for inclusion in the ASTR. This study is a first look at the differences between crashes in Arizona that result in injuries qualifying as major trauma and crashes that do not. This project independently looks at such crashes for both persons in passenger vehicles and those on motorcycles. The types of injuries by manner of collision were also calculated.

Methods

Data Collection

The data for this analysis came from ALISS and the ASTR, individually managed by ADOT and BEMSTS, respectively, and were probabilistically linked and met inclusion criteria as described previously in the introduction section of part I. Only those records missing important values for the variables of interest were excluded from this study. For the purposes of this report, "traumas" refer to individuals with records that were matched across the two databases, and "non-traumas" refer to individuals who were not injured in the crash or did not sustain injuries severe enough to qualify for ASTR inclusion.

Statistical Analysis

The statistical analysis was done using SAS 9.3. Exposure variables of interest were manner of collision, safety devices used, and estimated speed at the time of the incident. The manner of collision was organized into 11 different categories. Due to the small number of "rear-to–side" and "rear-to–rear" accidents, these categories were combined. A number of accidents that were already categorized as "Other," and along with the ambiguity present in the actual chain of events, were combined with "unknown manner of collision" to minimize the number of records omitted from the model. Safety devices documented as being used included seat belts, car seats, helmets and airbags. Estimated speed was measured in miles per hour and reported on the ACR if available. A log transformation of the estimated speed was used to normalize the data distribution. Potential confounders assessed were age and gender, with age measured in years.

Missing Data

As previously noted, missing data was a problem in matching the ASTR and ALISS databases. It presented a further issue in attempting to form a predictive model. To prevent loss of observations from the predictive model, missing information for age, speed, and seatbelt usage was imputed. These variables were determined to be missing at random. Age was imputed with the average age based upon seat position. Speed was similarly imputed using the average reported speed for traumas and non-traumas, respectively. If seatbelt usage was not reported, it was assumed that the individual was not wearing their seatbelt. From the information that was provided in ALISS, 81.2% of individuals were reported to have been wearing a seatbelt. This value is consistent with data available for the state of Arizona show that a seatbelt usage percentage of approximately 82% for years 2010 and 2011.[12] Seat position and gender also had missing information, and were left as missing because there was no way to determine where in the vehicle the individual was sitting at the time of the incident or their gender.

Only crashes that occur on public roads are entered into the database, which may lead to an underreporting of the number of persons involved and potentially injured in crashes. If law enforcement is not called to the scene of minor accidents then there is no record of this accident occurring. Seatbelt usage is often self-reported so there is potential for that percentage to be higher than reality.[13]

In looking at the subset of crashes involving motorcycles 30% of individuals had a safety device documented as “not applicable.” This makes it impossible to ascertain if the individual was wearing a helmet or not. For this reason, these individuals were excluded

from the model. It is interesting to note that the 2,115 records missing a documentation of helmet use or non-use were all for cases that did not result in a trauma. This is in contrast to other variables that have missing data as they have a similar number of missing records.

ROC Curves

Receiver Operating Characteristic (ROC) curves measure the accuracy of a test to discriminate "diseased" cases from "normal" cases, which in this case is the ability to discriminate persons with injuries from MVC that qualify as a trauma from those that were not traumas. There will not be perfect separation between these two groups. Based on the available characteristics, there will be some individuals who are classified as traumas that will be predicted to have not suffered a trauma; this population makes up the false negative fraction of the population (over-triage). Likewise, persons who did not suffer a trauma but are characterized to have suffered a traumatic injury are the false negative cases (under-triage).

Sensitivity is calculated as the number of true positives over the sum of the true positives and false negatives. Specificity is the number of true negatives over the sum of true negatives and false positives.

The ROC curve is a plot of the false positive rate (1-specificity) and the true positive rate (sensitivity). A test with perfect prediction will have a curve that passes through the upper left corner of the plot. The area under the curve indicates the diagnostic discrimination of the test. If the area under the curve is 0.5, the curve will lie directly on

the diagonal and indicates random predictions. It is generally accepted that an area under the curve of greater than 0.8 is valuable in predicting the responses of the individual subjects.[14]

Limitations

There are a number of limitations within the data. According to the ASTR, there were 23,787 patient records which indicated motor vehicle accident as the mechanism of injury. Of these only 14,412 were matched to their respective crash records. This leaves 9,375 records potentially incorrectly classified as non-trauma. These unmatched records probably do not exert a strong influence on the ratios as the number of non-trauma cases is over 500,000. This number of unmatched records from the trauma registry will hopefully decrease in the future as access to unique identifiers becomes available from ALISS.

Airbag deployment has been shown to decrease the number of MVC mortalities.[15] There have been conflicting studies done on whether airbags are responsible for an increase in risk of upper and lower extremity injuries when airbags deploy.[16] The increased presence of injuries has been shown in lab tests to be a result of the transfer of energy that occurs during collisions.[17] Studies have also found that the lack of seatbelt usage with airbag deployment is a contributing factor to injuries sustained in a crash.[15]Airbag deployment is noted in both the ALISS and the ASTR, but is not well populated in either database. The information on airbag deployment from the ALISS and the ASTR are independently shown in Table 3. In cross referencing the airbag field in both databases, it is possible to fill in some of the missing information on airbag use but this is only for trauma patients

which make up a small subset of the entire data set. Without reliable data from either source, it is impossible to include this as a predictor of trauma incidence or use it in looking at trauma severity.

Table 3: Airbag Usage (Front Row Passengers Only, N=496,059)

	Total	Trauma (N=9,963)	Non-Trauma (N=486,096)
Not Applicable	22,252 (4.5%)	319 (3.2%)	21,933 (4.5%)
Not Deployed	101,444 (20.5)	1,072 (10.8)	100,372 (20.7)
Deployed	27,249 (5.5)	2,098 (21.1)	25,151 (5.2)
Unknown	345,114 (69.6)	6,474 (65.0)	338,640 (69.7)

Another important factor that may contribute to car crashes and in turn injuries sustained is drug and alcohol use. It is possible to cross reference this information for the records that were successfully matched as it is reported from both data sources. It is better detailed in the ASTR as trauma physicians have the time and capability to accurately test for usage. In looking at drivers, only 263 of the 9,382 drivers with injuries meeting ASTR inclusion criteria were noted as being suspected of alcohol use on the ACR. From the ASTR, 1,354 drivers were reported as having alcohol limits above Arizona's legal limit. This apparent discrepancy carries over to drug use, as only 224 drivers were suspected of drug use from the crash data when 1,207 tested positive for some form of legal or illegal substance. (Table 4) Given that the apparent alcohol and drug use levels are poorly reported in the crash data when cross referenced with the trauma data, this information is excluded from any potential predictive model.

Table 4: Drug and Alcohol Use (Drivers Only , N=9,382)

Alcohol use		Drug Use	
Suspected from ALISS	263 (2.8%)	**Suspected from ALISS**	224 (2.4%)
Confirmed Above Legal Limit from ASTR:	1,354 (14.4)	**From ASTR:** **Confirmed Legal use:** **Confirmed Legal and Illegal use:** **Confirmed Illegal use:**	 164 113 930
		Total	1207(12.9)

Results

Demographic information for all persons entered into the ALISS regardless of injury status is provided in Table 5. The continuous variables such as age and estimated speed are given with their means and standard deviations. Estimated speed is highly skewed and a log transformation was used to help normalize the data.

Table 5: Variables Characteristics (N=580,561)

	Mean ± SD
Age	34.9 ± 18.6
Estimated Speed	18.8 ± 18.9
Log(Estimated Speed)	2.17 ± 1.55
	Number (Percent)
Gender	
Male	296,596 (51.1%)
Female	260,783 (44.9)
Missing	*37,594 (6.5)*
Seat Belt	
Yes	471,240 (81.2)
No	18,806 (3.2)
Missing	*90,605(15.6)*
Seat Position	
Driver	381,422 (65.7)
Front Row Other	91,314 (15.7)
2nd Row	63,701 (11.0)
3rd Row	7,047 (1.2)
External	189(0.03)
Missing	*36,978 (6.4)*
Extraction	
Yes	5,815 (1.0)
No	574,836 (99.0)
Collision Manner	
Single Vehicle	55,031 (9.5)
Angle	92,003(15.8)
Left Turn	61,698 (10.6)
Rear End	247,124 (42.6)
Head on	10,435 (1.8)
Sideswipe same direction	71,790 (12.4)
Sideswipe opposite direction	8,008 (1.4)

	Mean ± SD
Rear to Side/ Rear to Rear	7,526 (1.3)
Other	21,975 (3.8)
Missing	*5,061(0.87)*
Car Seat	
Yes	21,382 (3.7)
No	559,269 (96.3)
Motorcycle	
Yes	7,045 (1.2)
No	573,516 (98.8)
Helmet (Motorcycle Only)	
Yes	2,924 (41.5)
No	4,120 (58.5)

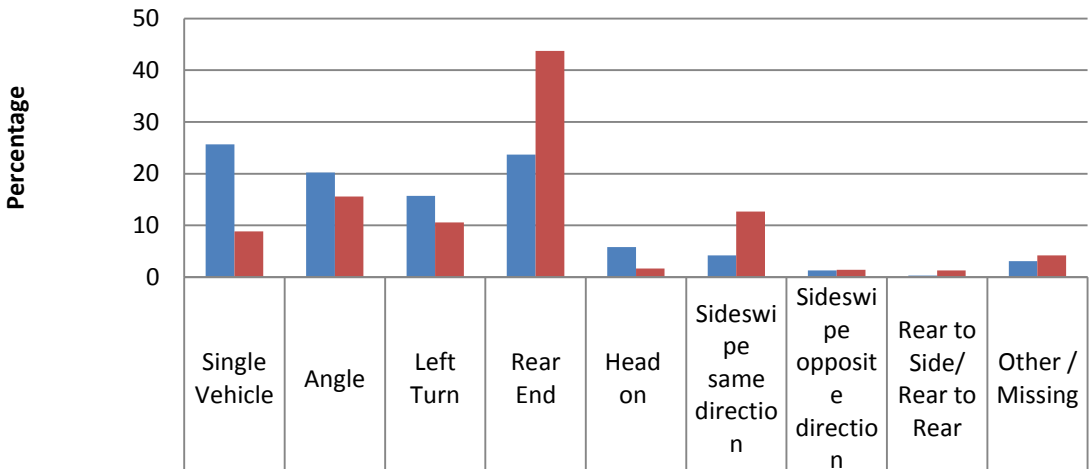
Table 6 is the same demographic information stratified by trauma after removing individuals involved in MVC but not within an automobile. The removals include motorcycle drivers and their passengers, individuals on bikes and pedestrians who may have been injured in the course of the MVC. Figure 2 shows the percentages of trauma versus non-trauma by collision type. There are visible differences in the outcomes based on collision manner. Each categorical variable was tested using a Chi-square test while the means of the two continuous variables were tested using a two sample independent t-test. Both tests were done to check for a statistically significant difference between the two groups. There was a statistically significant difference between the two groups for all variables included in Table 6. These variables were carried forward into the model. Previous research into car accidents has shown that there is interaction between seatbelt usage and gender and also an interaction between seatbelt usage and seat position.[13] Because both pieces of the potential interaction term are significant they will be tested in

the model. The interaction between seatbelt usage and seat position was not statistically significant so it was excluded. The odds ratios for Model 1 are shown in Table 7.

Table 6: Variable Characteristics by Outcome for Persons in Motor Vehicles at the Time of Accident

	Total	Trauma	Non-Trauma	P-value
N	566,559	11,185 (2.0%)	555,374 (98%)	
Age	35.05 ± 18.7	36.98 ± 20.3	35.01 ± 18.7	<0.0001
Estimated Speed	18.96 ± 18.9	28.21 ± 23.2	18.78 ± 18.8	
Log(Estimated Speed)	2.09 ± 1.57	2.52 ± 1.7	2.08 ± 1.6	<0.0001
Gender				**Chi-Square Test p-value**
Male	285,801	5,575 (49.84)	280,226 (50.5)	<0.0001
Female	257,700	5,610 (50.16)	252,090 (45.4)	
Seat Belt				
Yes	471,085	7,688 (68.7)	463,397 (84.4)	<0.0001
No	15,180	1,613 (14.4)	13,567 (2.4)	
Seat Position				
Driver	394,060	7,661 (68.5)	386,399 (69.6)	<0.0001
Front Row Other	91,180	2,098 (18.8)	89,082 (16.0)	
2nd Row	63,357	1,068 (9.5)	62,289 (11.2)	
3rd Row	7,027	143 (1.3)	6,884 (1.2)	
External	116	11 (0.1)	105 (0.02)	
Extraction				
Yes	5,755	1,152 (10.3)	4,603 (0.8)	<0.0001
No	560,804	10,033 (89.7)	550,771 (99.2)	
Collision Manner				
Single Vehicle	52,045	2,871 (25.7)	49,174 (8.85)	<0.0001
Angle	88,605	2,263 (20.2)	86,342 (15.6)	
Left Turn	60,573	1,756 (15.7)	58,817 (10.6)	
Rear-end	245,519	2,645 (23.7)	242,874 (43.7)	
Head-on	10,061	645 (5.8)	9,416 (1.7)	
Sideswipe same direction	70,856	474 (4.2)	70,382 (12.7)	
Sideswipe opposite direction	7,796	150 (1.3)	7,646 (1.4)	
Rear to Side/ Rear to Rear	7,459	38 (0.3)	7,421 (1.3)	
Other/Missing	23,645	343 (3.1)	23,302 (4.2)	
Car Seat				
Yes	21,374	259 (2.3)	21,115 (3.8)	<0.0001
No	545,185	10,926 (97.7)	534,259 (96.2)	

Figure 2: Trauma vs Non-Trauma by Collision Manner



	Single Vehicle	Angle	Left Turn	Rear End	Head on	Sideswipe same direction	Sideswipe opposite direction	Rear to Side/ Rear to Rear	Other / Missing
Trauma	25.7	20.2	15.7	23.7	5.8	4.2	1.3	0.3	3.1
Non-Trauma	8.85	15.6	10.6	43.7	1.7	12.7	1.4	1.3	4.2

As seen in Table 7, the odds of an individual suffering a traumatic injury in a car crash are 8.4 times higher if they have to be extricated from the vehicle by EMS or law enforcement. Extraction indicates that the accident occurred with significant forces to either seriously damage the frame of the vehicle or that the injuries sustained are such that the person cannot exit the vehicle under their own power. In comparison to rear-end collisions, persons traveling in vehicles that are sideswiped from the same direction in which they are traveling and those that are involved in rear-to-rear or rear-to-side accidents are less likely to suffer traumatic injuries. Logically this makes sense as forces are traveling in the same direction in the sideswipe incident and the majority of rear-to-rear and rear-to-side accidents occur at low rates of speed. The odds of a person suffering

a traumatic injury when failing to wear a seatbelt or be restrained in a car seat is 3.98 times higher than when one is properly restrained. An ROC curve was used to test the goodness of fit for the model (Figure 3). The area under the curve for Model 1 is 0.7523. This shows that this model is good but does not quite reach the accepted value of 0.8 for a good predictive model.

Table 7: Logistic Regression Model 1

	Odds Ratio	95% Wald Confidence Limits	P-value
Age	1.007	(1.006, 1.008)	<0.0001
Log(Estimated Speed)	1.095	(1.079, 1.110)	<0.0001
Gender			
Male	1.000	Reference	
Female	1.297	(1.239, 1.357)	0.0004
Seat Restraint (Belt or Carseat)			
Yes	1.000	Reference	
No	3.984	(3.703, 4.288)	<0.0001
Seat Position			
Driver	1.000	Reference	
Front Row Other	1.079	(1.025, 1.135)	0.0034
2nd Row	0.911	(0.851, 0.979)	0.0107
3rd Row	1.056	(0.892, 1.259)	0.5093
External	1.281	(0.678, 2.429)	0.4429
Extraction			
No	1.000	Reference	
Yes	8.667	(8.057, 9.327)	<0.0001
Collision Manner			
Rear-End	1.00	Reference	
Single Vehicle	3.631	(3.419, 3.846)	<0.0001
Angle	2.029	(1.913, 2.152)	<0.0001
Left Turn	2.396	(2.251, 2.553)	<0.0001
Head-on	4.705	(4.285, 5.167)	<0.0001
Sideswipe same direction	0.566	(0.512, 0.626)	<0.0001
Sideswipe opposite direction	1.586	(1.339, 1.879)	<0.0001
Rear to Side/ Rear to Rear	0.428	(0.307, 0.596)	<0.0001
Other/Unknown	1.250	(1.113, 1.404)	0.0002
Gender*Seat Restraint	0.886	(0.810, 0.969)	0.0083

(33,499 omitted for missing gender and/or seat position- 5.9%)

Motorcycles prove to be a unique set of motor vehicle accidents. These incidents are often looked at separately due to the unique nature of the safety devices utilized and disproportionate amount of injuries and fatalities that result.[18] The demographic information of this sub-population is given in Table 8, both as a whole and stratified by injury outcome. Similar to Table 6, the variables were tested using t-tests (continuous variables) and chi-square tests (categorical variables). Unlike the first model, gender does not differ significantly by trauma outcome and is not included in the model. The odds ratios for Model 2 are shown in Table 9.

Table 8: Variable Characteristics by Outcome for Persons on Motorcycles at the Time of Accident

	Total	Trauma	Non-Trauma	P-value
N	7,045	1,734 (24.6%)	5,311 (75.4%)	
Age	39.65 ± 15.8	41.36 ± 15.4	39.09 ± 15.9	<0.0001
Log(est. speed)	2.65 ± 1.4	2.91 ± 1.4	2.57 ± 1.4	<0.0001
Gender (101 missing)				**Chi-Square Test P-value**
Male	5,890	1,487 (85.8)	4,403 (82.9)	0.2108
Female	1,054	247 (14.2)	807 (15.2)	
Helmet				
Yes	2,924	693 (53.1)	2,231 (61.5)	<0.0001
No	2,006	611 (46.9)	1,395(38.5)	
Collision Manner				
Single Vehicle	2,986	764 (44.1)	2,222 (41.8)	<0.0001
Angle	852	253 (14.6)	599 (11.3)	
Left Turn	609	210 (12.1)	399 (7.5)	
Rear-End	1,410	249 (14.4)	1,161 (21.9)	
Head-on	152	47 (2.7)	105 (2.0)	
Sideswipe same direction	579	114 (6.6)	465 (8.8)	
Sideswipe opposite direction	95	17(0.98)	78 (1.5)	

	Total	Trauma	Non-Trauma	P-value
Rear to Side/ Rear to Rear	43	4 (0.23)	39 (0.7)	
Other/Unknown	319	76 (4.4)	243 (4.6)	

As seen in Table 9, accidents occurring while making left turns and head-on collisions have the highest odds of resulting in a traumatic injury in comparison to rear-end accidents. Not wearing a helmet yields odds of traumatic injury 1.371 times higher than those wearing a helmet. Again an ROC curve is used to test the goodness of fit for the model. (Figure 4) The area under the curve in this case is 0.6224. This model is not as good at predicting the outcome as Model 1.

Table 9: Logistic Regression Model 2

	Odds Ratio	95% Wald Confidence Limits	P-value
Age	1.010	(1.007, 1.014)	<0.0001
Log(Estimated Speed)	1.170	(1.122, 1.221)	<0.0001
Helmet			
Yes	1.000	Reference	
No	1.371	(1.204,1.561)	<0.0001
Collision Manner			
Rear End	1.000	Reference	
Single Vehicle	1.350	(1.117, 1.631)	0.0019
Angle	1.780	(1.397, 2.268)	<0.0001
Left Turn	2.623	(2.038, 3.376)	<0.0001
Head on	2.077	(1.335, 3.230)	0.0012
Sideswipe same direction	0.973	(0.722, 1.312)	0.8589
Sideswipe opposite direction	1.021	(0.516, 2.022)	0.9524
Rear to Side/ Rear to Rear	0.377	(0.087, 1.626)	0.1908
Other/Unknown	1.379	(0.968, 1.963)	0.0748

(2,129 omitted because of missing gender and/or helmet usage- 30.2%)

Figure 3: ROC Curve for Model 1

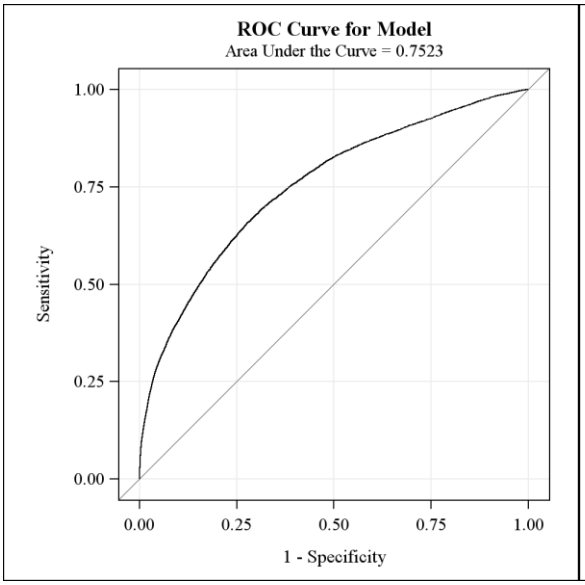
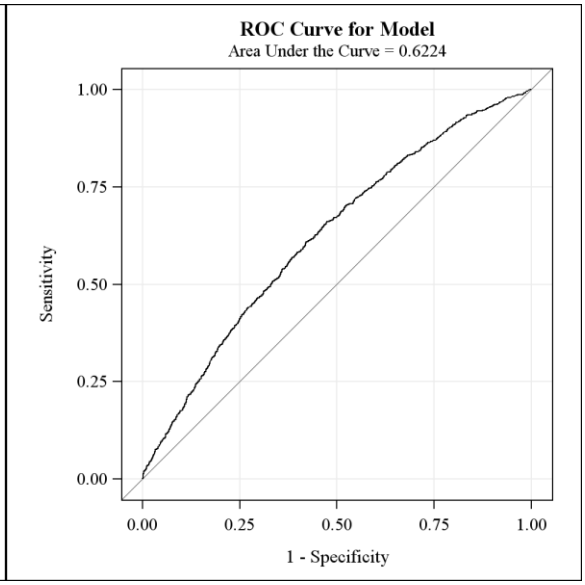


Figure 4: ROC Curve for Model 2



In addition to utilizing the data linkage procedure to assess differences between accidents that cause injuries that are classified as traumas and those that are not, the linkage allows the types of injuries (Figure 5) and the number of body areas injured (Figure 6) to be assessed based on collision manner. In this case other and missing manner of collision are not included in the data analysis. As seen in Figure 5, the majority of individuals have at least some sort of external injury, such as bruises or abrasions. Injuries to the head and the extremities are also highly prevalent. Figure 6 shows that over two-thirds of individuals injured in accidents only have injuries to 1 or 2 major body areas.

Figure 5: Body Area Injured by Collision Manner

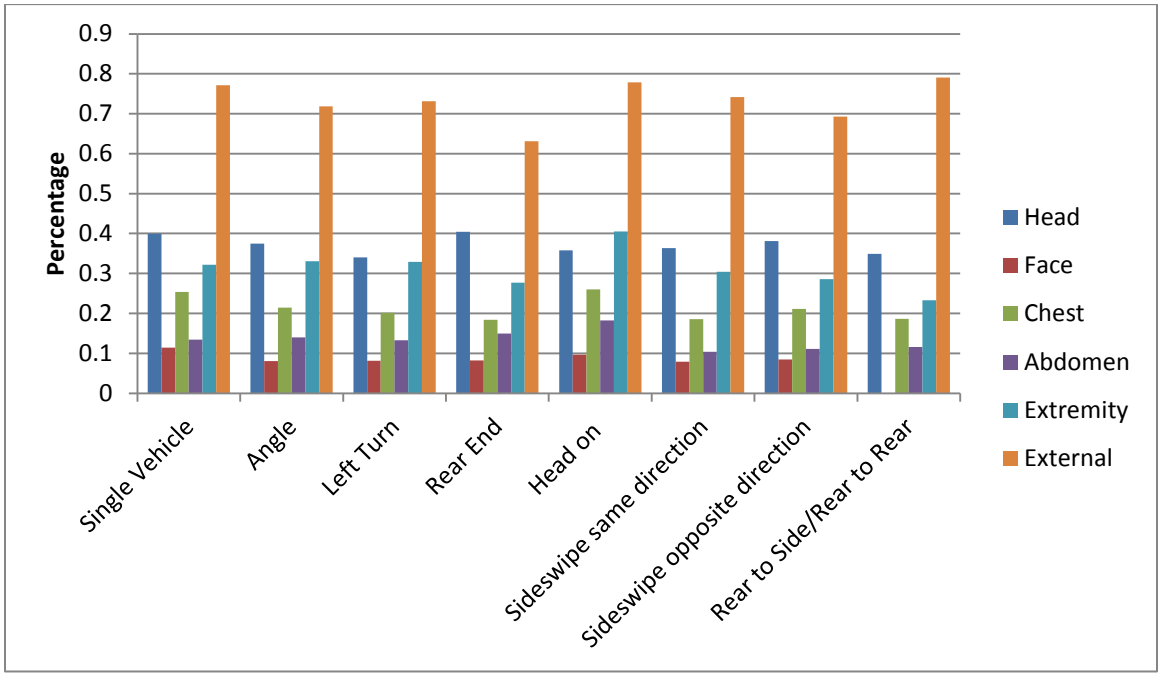
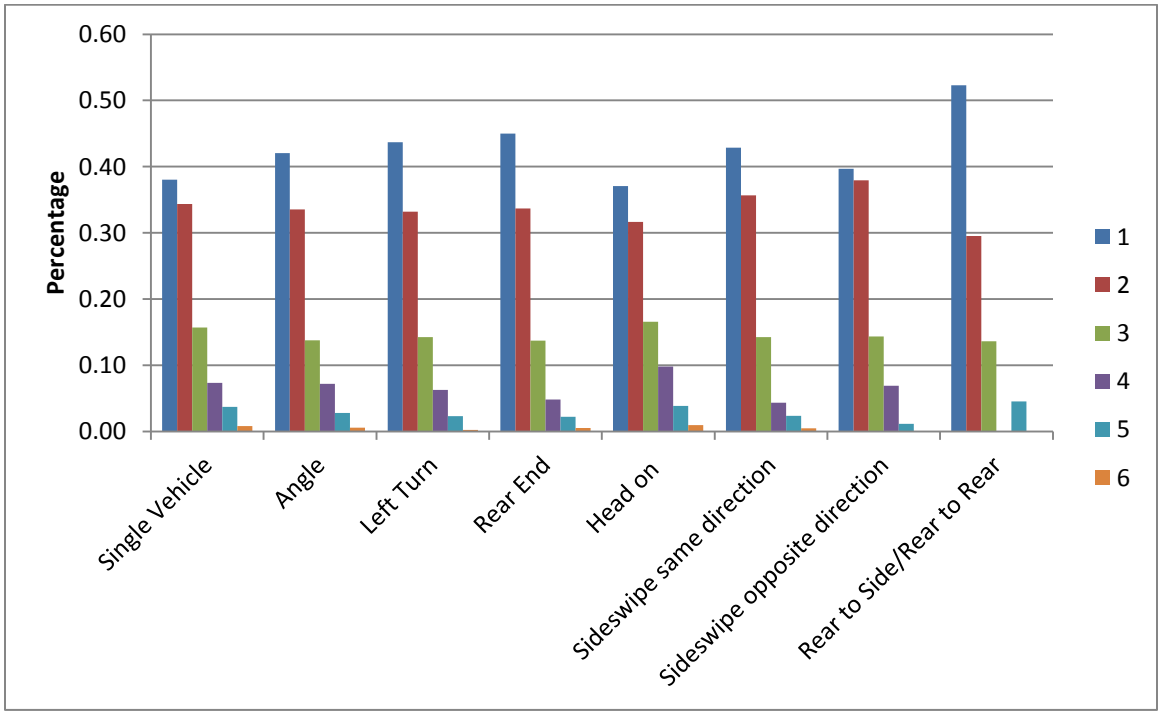


Figure 6: % of people with Multiple Body Part Injuries by Collision Manner



Discussion

The current study looks at the difference in injury outcomes for all individuals involved in MVCs in Arizona for 2010 and 2011. The initial model for predicting the occurrence of an injury does not meet the generally accepted value for a valuable predictive model at 0.7523. This, however, is to be expected because not all car crashes are created equal. Additionally, there are a number of factors that were not looked at in this study. The type of vehicle and weather conditions were omitted and not included in the study. Even though a majority of Arizona does not experience a lot of icy roads, conditions like wet roads, blinding sunlight and dust storms are still very much present and may affect the severity of a crash. No two car crashes are the same, meaning that a perfect predictive model does not exist.

The analysis confirms some previous studies on the effect of seatbelt usage. Seatbelt usage was shown to have a significant decrease in the odds of a person sustaining a serious injury. Individuals not wearing seatbelts were 3.98 times more likely to suffer a trauma than those who were wearing a seatbelt. Seatbelt usage may be associated with better overall driver safety and consequently a decrease in crash fatality rates.[19] Arizona does not have a primary seatbelt law. Only minors and all people in the front seat are required to wear seatbelts. Vehicles cannot be pulled over for other occupants not wearing seatbelts without some other primary violation (ARS Title 28-909). Individuals can be cited for not wearing a seatbelt if the car is pulled over for another violation. The data in this study showed that higher seatbelt usage was reported for minors, with 89% reported wearing a seatbelt or being secured in a car seat. This was an increase from adults where 83% were wearing a seatbelt at the time of the accident. Individuals in the

front row of the vehicle had a seatbelt usage of 88%. The interaction between seat position and seatbelt usage was not significant but seatbelt usage was higher in the first row. Arizona also lags behind the national average for seatbelt usage. The national average for seat belt usage is 86%. From a geographic standpoint the western United States has an average seatbelt usage of 94%.[20] There have been recent efforts across the state to increase seatbelt usage. The fact that wearing a seatbelt decreases the chances of serious injury from an accident is not new information, yet close to 18% of Arizona's population still chooses to not wear their seatbelts.

The motorcycle model has a lower goodness of fit score than that of the initial model for all other vehicles. Almost a third of the observations for this model were omitted due to missing helmet use documentation. There is only a motorcycle helmet law for minors in Arizona. More information about the events of collisions that involve motorcycles may help predict trauma incidents. The type of vehicle may play an important role in motorcycle incidents due to the large range of size discrepancies between a motorcycle and other motor vehicles.

Recommendations and Future Study

Safety

The information available shows that seatbelt usage is an important factor in preventing traumatic injuries in passenger vehicle collisions. This is the only factor looked at that can be directly changed by the population. Speed is also an important factor in the severity of crashes. However, the reporting of estimated speed is just that, an estimate. The model does show that as speed increases so does the risk of being severely injured. This is consistent with other models that have assessed speed and injury severity. Speed has also been shown to increase the risk of being involved in a car accident.[21]

In this study airbag deployment was not adequately reported, so it was not included in the model. It is not possible to say what impact airbag deployment is having on the injury report. It would be interesting to see what the impact is in Arizona, as there are disagreeing view points on whether airbags, while preventing mortality, may be resulting in an increase in injuries to the extremities.[22] In order to fully explore the effect of airbag deployment, it is important to keep the make and model of the vehicle in mind as airbag technology is constantly improving. Airbags have only been required in cars since 1998. Side impact air bags are not currently required but their presence is increasing.

Future

If ADOT and ADHS stay committed to working together, the data reported from both departments should improve. Discussions have already led ADOT to begin to set up a process to electronically submit to ADHS crash data containing unique identifiers. Unique identifiers will be crucial for improving the accuracy and validity of the linkage.

With each passing year, it will be possible to identify trends and areas of concern that need to be addressed. This could be reductions in seatbelt usage or increases in trauma-causing accidents. Being able to track trends would allow Arizona to identify problem areas and weaknesses in the current system.

The addition of an EMS database will fill in the gap of data between the incident (the crash) and treatment in the emergency department. This database may also be able to fill in fields that are missing from some of the crash records. By establishing a complete chain of events, Arizona will have a better idea of how best to distribute resources to best serve the public.

This study did not look specifically at crashes resulting in fatalities. There were 226 records in the ASTR of patients who died in the emergency department. This study also does not include information on deaths that occur at the scene. These crashes are captured by ADOT and submitted to the national FARS (Fatality Analysis Reporting System) database managed by the National Highway Transportation Safety Administration (NHTSA).

This study did not include the first harmful event as a predictive factor as there are over 40 categories for this field. It would have been interesting to see how the outcomes were affected by the rate of speed the vehicle was traveling in comparison to the posted speed limit. The posted speed field had a number of missing values and a large proportion completed with a zero, so it was not utilized in this study. It may be possible through the use of geographic information systems to obtain valid information for this field as the location of the crash is well documented.

This project has applicability to many areas of public health. Of the three core functions, this project most specifically dealt with assessment and assurance. By using the ASTR and ALISS databases, it is possible to monitor the health status of crash patients based on crash type and other contributing factors.

A number of the ten essential public health services are also covered within this internship. ASTR identifies and monitors crash frequency and contributing factors to injury severity. This function will monitor the health status of the community to identify community health problems. Although not addressed in this project, the linkage of ASTR and ALISS will allow officials to investigate potential community health hazards. The information used for this project is collected as part of a joint effort between ADHS, ADOT, and the Governor's Office of Highway Safety (GOHS). This collaboration meets another essential public health service as the community partnerships are used to identify and solve health problems. The linkage outcome of this project will provide these organizations data that can be further analyzed to facilitate policies and future plans. Information on crash-related traumatic injuries will ensure that public health and health care workers are adequately trained to provide appropriate care. Further data analysis may also identify high frequency crash areas to help facilitate research to result in crash mitigation.

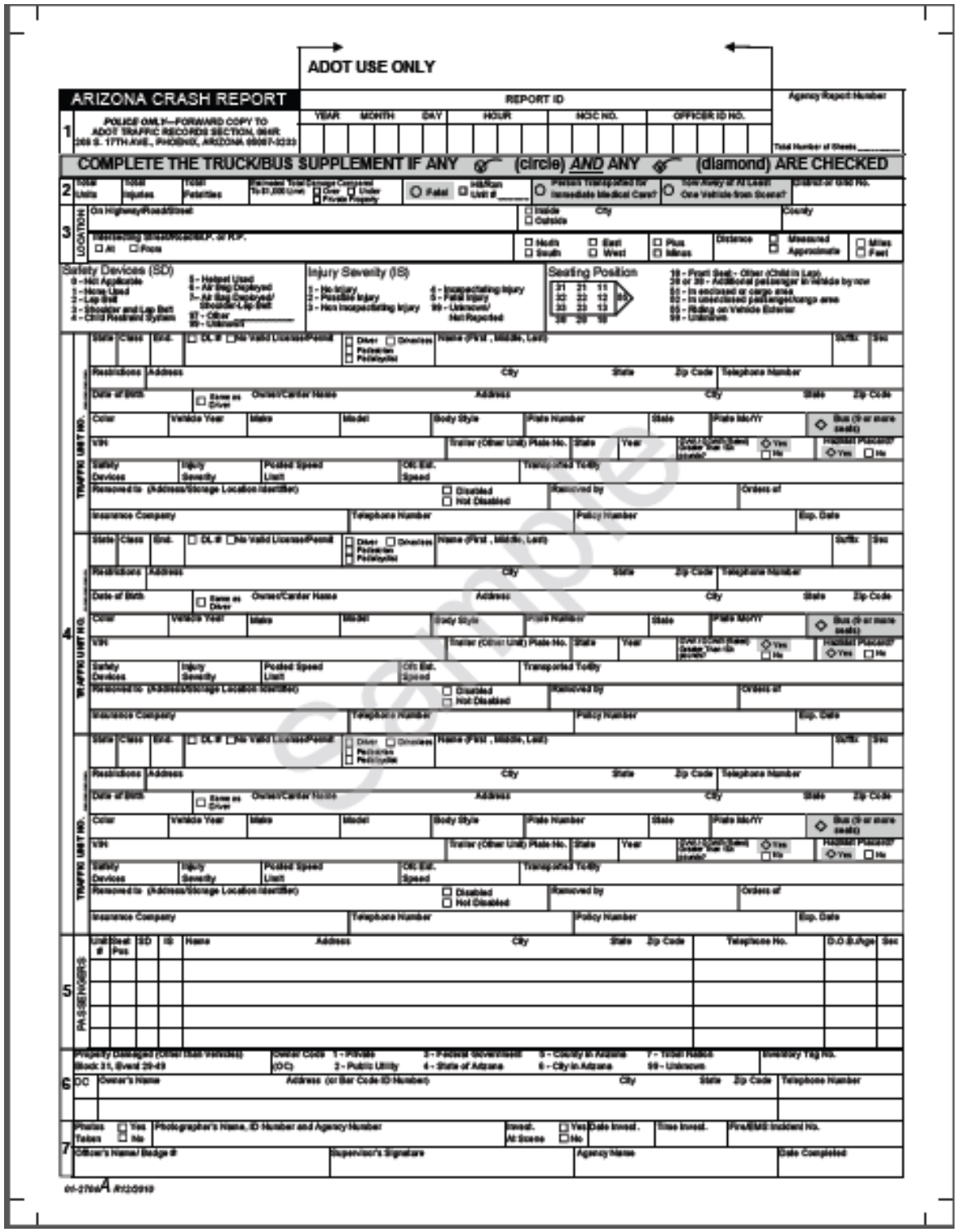
Conclusion

This analysis shows that valuable information can be determined by establishing a link between ALISS and ASTR. By establishing a reliable method for linking specified data elements of the databases, ADOT and ADHS, their constituent resources, system stakeholders and the public will be able to study motor vehicle crashes where injuries or loss of life have occurred, and identify new relationships between the data that will prompt better or new preventative measures and treatment modalities that reduce mortality and morbidity. Although neither model meets the threshold for a valuable predictive curve, this evaluation shows that there are significant differences in the outcome based upon collision type, use of a restraint systems, estimated speed, gender, extrication and age. Motorcycles, as a subset of this data, have significant differences in outcome based on collision type, helmet use, estimated speed and age. Finally, the linkage will provide measureable data to support new or revisions of state and/or local highway safety laws.

The employees of the BEMSTS section at ADHS were very helpful and provided valuable feedback throughout the project as bumps in the road were discovered. The original plan for this data, after executing a probabilistic linkage was to analyze the data for correlations between vehicle speed and injury severity. Additionally the nature of the crash was going to be explored at in order to determine if injury severity differs among the types of crashes. This was changed to an analysis of a predictive model for predicting trauma injury based on the characteristics of the crash. The change was made due to the type of reliable data that was available. The large amount of missing data was not anticipated. The data analysis was modified to minimize the number of records that

would be deleted. This issue will hopefully be rectified if the two departments continue to work together to make an effort to get complete information about the incident. The most important goal of this internship was the linkage of the database. While it is a tentative linkage because of the lack of unique identifiers, it does show that a linkage between the ASTR and the ALISS can be set up and utilized for further data analysis.

Appendix I- Arizona Crash Report



ADOT USE ONLY

ARIZONA CRASH REPORT

REPORT ID

Agency Report Number

1 POLICE ONLY—FORWARD COPY TO ADOT TRAFFIC RECORDS SECTION, 06HR 206 S. 17TH AVE., PHOENIX, ARIZONA 85007-3233

YEAR MONTH DAY HOUR NCIC NO. OFFICER ID NO.

Total Number of Sheets

COMPLETE THE TRUCK/BUS SUPPLEMENT IF ANY (circle) AND ANY (diamond) ARE CHECKED

2 Total Units | Total Injuries | Total Fatalities | Estimated Total Damage Compared To $1,000 Limit: Over / Under / Private Property | Fatal | Hit/Run Unit # | Person Transported for Immediate Medical Care? | Tow Away of At Least One Vehicle from Scene? | District or Grid No.

3 LOCATION: On Highway/Road/Street | Inside / Outside | City | County

Intersecting Street/Road/M.P. or R.P. At / From | North / South / East / West | Plus / Minus | Distance | Measured / Approximate | Miles / Feet

Safety Devices (SD): 0 - Not Applicable; 1 - None Used; 2 - Lap Belt; 3 - Shoulder and Lap Belt; 4 - Child Restraint System; 5 - Helmet Used; 6 - Air Bag Deployed; 7 - Air Bag Deployed/Shoulder-Lap Belt; 97 - Other; 99 - Unknown

Injury Severity (IS): 1 - No Injury; 2 - Possible Injury; 3 - Non Incapacitating Injury; 4 - Incapacitating Injury; 5 - Fatal Injury; 99 - Unknown/Not Reported

Seating Position: 31 21 11; 32 22 12; 33 23 13; 38 28 18; 19 - Front Seat - Other (Child in Lap); 29 or 39 - Additional passenger in vehicle by row; 51 - In enclosed or cargo area; 52 - In unenclosed passenger/cargo area; 55 - Riding on Vehicle Exterior; 99 - Unknown

TRAFFIC UNIT NO.

State | Class | End. | DL # | No Valid License/Permit | Driver / Pedestrian / Pedalcyclist | Drivers | Name (First, Middle, Last) | Suffix | Sex

Restrictions | Address | City | State | Zip Code | Telephone Number

Date of Birth | Same as Driver | Owner/Carrier Name | Address | City | State | Zip Code

Color | Vehicle Year | Make | Model | Body Style | Plate Number | State | Plate Mo/Yr | Bus (9 or more seats)

VIN | Trailer (Other Unit) Plate No. | State | Year | Gross Vehicle Weight Greater Than 10k pounds? Yes / No | Hazmat Placard? Yes / No

Safety Devices | Injury Severity | Posted Speed Limit | Off. Est. Speed | Transported To/By

Removed to (Address/Storage Location Identifier) | Disabled / Not Disabled | Removed by | Orders of

Insurance Company | Telephone Number | Policy Number | Exp. Date

4 TRAFFIC UNIT NO.

State | Class | End. | DL # | No Valid License/Permit | Driver / Pedestrian / Pedalcyclist | Drivers | Name (First, Middle, Last) | Suffix | Sex

Restrictions | Address | City | State | Zip Code | Telephone Number

Date of Birth | Same as Driver | Owner/Carrier Name | Address | City | State | Zip Code

Color | Vehicle Year | Make | Model | Body Style | Plate Number | State | Plate Mo/Yr | Bus (9 or more seats)

VIN | Trailer (Other Unit) Plate No. | State | Year | Gross Vehicle Weight Greater Than 10k pounds? Yes / No | Hazmat Placard? Yes / No

Safety Devices | Injury Severity | Posted Speed Limit | Off. Est. Speed | Transported To/By

Removed to (Address/Storage Location Identifier) | Disabled / Not Disabled | Removed by | Orders of

Insurance Company | Telephone Number | Policy Number | Exp. Date

TRAFFIC UNIT NO.

State | Class | End. | DL # | No Valid License/Permit | Driver / Pedestrian / Pedalcyclist | Drivers | Name (First, Middle, Last) | Suffix | Sex

Restrictions | Address | City | State | Zip Code | Telephone Number

Date of Birth | Same as Driver | Owner/Carrier Name | Address | City | State | Zip Code

Color | Vehicle Year | Make | Model | Body Style | Plate Number | State | Plate Mo/Yr | Bus (9 or more seats)

VIN | Trailer (Other Unit) Plate No. | State | Year | Gross Vehicle Weight Greater Than 10k pounds? Yes / No | Hazmat Placard? Yes / No

Safety Devices | Injury Severity | Posted Speed Limit | Off. Est. Speed | Transported To/By

Removed to (Address/Storage Location Identifier) | Disabled / Not Disabled | Removed by | Orders of

Insurance Company | Telephone Number | Policy Number | Exp. Date

5 PASSENGERS

Unit #	Seat Pos	SD	IS	Name	Address	City	State	Zip Code	Telephone No.	D.O.B./Age	Sex

6 Property Damaged (Other than Vehicles) Block 31, Event 29-49 | Owner Code (OC): 1 - Private; 2 - Public Utility; 3 - Federal Government; 4 - State of Arizona; 5 - County in Arizona; 6 - City in Arizona; 7 - Tribal Nation; 99 - Unknown | Inventory Tag No.

OC | Owner's Name | Address (or Bar Code ID Number) | City | State | Zip Code | Telephone Number

7 Photos Taken Yes / No | Photographer's Name, ID Number and Agency Number | Invest. At Scene Yes / No | Date Invest. | Time Invest. | Fire/EMS Incident No.

Officer's Name/ Badge # | Supervisor's Signature | Agency Name | Date Completed

01-2764A R12/2010

Sample

8 WITNESSES	Name	Address	City	State	Zip Code	Telephone Number	D.O.B./Age

9 CITATION/CHARGES	UNIT #	A.R.S. NO. OR CITY CODE	UNIT #	A.R.S. NO. OR CITY CODE

BLOCKS 10 - 24: CHECK ONLY ONE OR ONE BLOCK PER UNIT UNLESS NOTED

10 — LIGHT CONDITION

- ☐ 1 DAYLIGHT
- ☐ 2 DAWN
- ☐ 3 DUSK
- ☐ 4 DARK—LIGHTED
- ☐ 5 DARK—NOT LIGHTED
- ☐ 6 DARK—UNKNOWN LIGHTING

11 — WEATHER CONDITIONS

- ☐ 1 CLEAR
- ☐ 2 CLOUDY
- ☐ 3 SLEET/HAIL (freezing rain/drizzle)
- ☐ 4 RAIN
- ☐ 5 SNOW
- ☐ 6 SEVERE CROSSWINDS
- ☐ 7 BLOWING SAND, SOIL, DIRT
- ☐ 8 FOG, SMOG, SMOKE
- ☐ 9 BLOWING SNOW
- ☐ 97 OTHER _____
- ☐ 99 UNKNOWN

12 — ROAD SURFACE CONDITION

UNIT #

- ☐ ☐ ☐ 1 DRY
- ☐ ☐ ☐ 2 WET
- ☐ ☐ ☐ 3 SNOW
- ☐ ☐ ☐ 4 SLUSH
- ☐ ☐ ☐ 5 ICE/FROST
- ☐ ☐ ☐ 6 WATER (standing, moving)
- ☐ ☐ ☐ 7 SAND
- ☐ ☐ ☐ 8 MUD, DIRT, GRAVEL
- ☐ ☐ ☐ 9 OIL
- ☐ ☐ ☐ 97 OTHER _____
- ☐ ☐ ☐ 99 UNKNOWN

13 — ROAD GRADE

UNIT #

- ☐ ☐ ☐ 1 LEVEL
- ☐ ☐ ☐ 2 DOWNHILL
- ☐ ☐ ☐ 3 UPHILL
- ☐ ☐ ☐ 4 HILLCREST
- ☐ ☐ ☐ 5 SAG/DIP/BOTTOM
- ☐ ☐ ☐ 99 UNKNOWN

14 — RELATION TO JUNCTION

- ☐ 0 NOT JUNCTION RELATED

NON-CONTROLLED ACCESS AREA

- ☐ 1 INTERSECTION (within)
- ☐ 2 INTERSECTION-RELATED
- ☐ 3 ENTRANCE/EXIT RAMP (not areas)
- ☐ 4 RAILWAY GRADE CROSSING
- ☐ 5 MEDIAN CROSSOVER-RELATED
- ☐ 6 FRONTAGE ROAD
- ☐ 7 DRIVEWAY
- ☐ 8 ALLEY-ACCESS-RELATED
- ☐ 9 UNKNOWN NON-INTERCHANGE

CONTROLLED ACCESS AREA

- ☐ 10 THRU ROADWAY
- ☐ 11 INTERSECTION (within)
- ☐ 12 INTERSECTION-RELATED
- ☐ 13 ENTRANCE/EXIT RAMP
- ☐ 14 FRONTAGE ROAD
- ☐ 15 OTHER PART OF INTERCHANGE
- ☐ 99 UNKNOWN

15 — TYPE OF INTERSECTION

- ☐ 1 FOUR- WAY INTERSECTION
- ☐ 2 T - INTERSECTION
- ☐ 3 Y - INTERSECTION
- ☐ 4 INTER. AS PART OF INTERCHANGE
- ☐ 5 TRAFFIC CIRCLE
- ☐ 6 ROUNDABOUT
- ☐ 7 FIVE POINT, OR MORE
- ☐ 99 UNKNOWN

16 — TRAFFIC WAY DESCRIPTION

- ☐ 1 ONE WAY TRAFFICWAY
- ☐ 2 TWO-WAY, NOT DIVIDED (no median present)
- ☐ 3 TWO-WAY, (NOT DIVIDED) WITH A CONTINUOUS LEFT TURN LANE
- ☐ 4 TWO-WAY, DIVIDED, UNPROTECTED (PAINTED< 4 FEET) MEDIAN
- ☐ 5 TWO-WAY, DIVIDED, POSITIVE MEDIAN BARRIER
- ☐ 99 UNKNOWN

17 — MANNER OF CRASH IMPACT

- ☐ 1 SINGLE VEHICLE
- ☐ 2 ANGLE (front to side) (other than left turn)
- ☐ 3 LEFT TURN
- ☐ 4 REAR END (front-to-rear)
- ☐ 5 HEAD-ON (front-to-front) (other than left turn)
- ☐ 6 SIDESWIPE, SAME DIRECTION
- ☐ 7 SIDESWIPE, OPPOSITE DIRECTION
- ☐ 8 REAR-TO-SIDE
- ☐ 9 REAR-TO-REAR
- ☐ 97 OTHER _____
- ☐ 99 UNKNOWN

18 — DIRECTION OF UNIT TRAVEL (Compass)

BEFORE 1ST CRASH EVENT

UNIT #

- ☐ ☐ ☐ 1 NORTH
- ☐ ☐ ☐ 2 SOUTH
- ☐ ☐ ☐ 3 EAST
- ☐ ☐ ☐ 4 WEST
- ☐ ☐ ☐ 5 NORTHWEST
- ☐ ☐ ☐ 6 NORTHEAST
- ☐ ☐ ☐ 7 SOUTHWEST
- ☐ ☐ ☐ 8 SOUTHEAST
- ☐ ☐ ☐ 99 UNKNOWN

19 — CONTRIBUTING CIRCUMSTANCES

UP TO TWO CHOICES PER UNIT

UNIT #

- ☐ ☐ ☐ 0 NO CONTRIBUTING CIRCUMSTANCE

ENVIRONMENTAL

- 1 GLARE
  - ☐ ☐ ☐ A. SUNLIGHT
  - ☐ ☐ ☐ B. HEADLIGHTS
- 2 PHYSICAL OBSTRUCTION(S)
  - ☐ ☐ ☐ A. STOPPED/PARKED VEHICLE
  - ☐ ☐ ☐ B. MOVING VEHICLE
  - ☐ ☐ ☐ C. LOAD ON VEHICLE
  - ☐ ☐ ☐ D. TREES/SHRUBS/BRUSH

ROAD

- ☐ ☐ ☐ 3 ROAD SURFACE CONDITION
- ☐ ☐ ☐ 4 DEBRIS
- 5 WORK ZONE
  - ☐ ☐ ☐ A. LANE CLOSURE
  - ☐ ☐ ☐ B. LANE SHIFT/CLOSURE
  - ☐ ☐ ☐ C. WORK ON SHOULDER OR MEDIAN
  - ☐ ☐ ☐ D. INTERMITTENT OR MOVING WORK
  - ☐ ☐ ☐ E. OTHER _____
  - ☐ ☐ ☐ F. WORKERS PRESENT
- ☐ ☐ ☐ 6 OBSTRUCTION IN ROADWAY
- ☐ ☐ ☐ 7 CHANGING ROAD WIDTH
- ☐ ☐ ☐ 8 NON-HIGHWAY WORK

MOTOR VEHICLE

- ☐ ☐ ☐ 9 BRAKES
- ☐ ☐ ☐ 10 STEERING
- ☐ ☐ ☐ 11 SUSPENSION
- ☐ ☐ ☐ 12 TIRES
- ☐ ☐ ☐ 13 WHEELS
- ☐ ☐ ☐ 14 LIGHTS (head, signal, tail)
- ☐ ☐ ☐ 15 WINDOWS/WINDSHIELD
- ☐ ☐ ☐ 16 MIRRORS
- ☐ ☐ ☐ 17 WIPERS
- ☐ ☐ ☐ 18 TRUCK COUPLING/TRAILER/HITCH/SAFETY CHAINS
- ☐ ☐ ☐ 97 OTHER _____
- ☐ ☐ ☐ 99 UNKNOWN

20 — TRAFFIC CONTROL DEVICE

UNIT #

- ☐ ☐ ☐ 0 NO CONTROLS
- ☐ ☐ ☐ 1 SIGNAL
- ☐ ☐ ☐ 2 STOP SIGN
- ☐ ☐ ☐ 3 YIELD SIGN
- ☐ ☐ ☐ 4 WARNING SIGN
- ☐ ☐ ☐ 5 RAILROAD CROSSING DEVICE
- ☐ ☐ ☐ 6 FLASHING TRAFFIC SIGNAL
- ☐ ☐ ☐ 7 PERSON (law enforcement, crossing guard, flagger, etc.)
- ☐ ☐ ☐ 97 OTHER _____
- ☐ ☐ ☐ 99 UNKNOWN

21 — CONDITION INFLUENCING Driver/Pedi/Cyclist

UP TO TWO CHOICES PER UNIT

UNIT #

- ☐ ☐ ☐ 0 NO APPARENT INFLUENCE
- ☐ ☐ ☐ 1 ILLNESS
- ☐ ☐ ☐ 2 PHYSICAL IMPAIRMENT
- ☐ ☐ ☐ 3 FELL ASLEEP/FATIGUED
- ☐ ☐ ☐ 4 ALCOHOL
- ☐ ☐ ☐ 5 DRUGS
- ☐ ☐ ☐ 6 MEDICATIONS

CHECK ONE IF BLOCKS 4, 5, OR 6 CHECKED

- ☐ ☐ ☐ A. NO TEST GIVEN
- ☐ ☐ ☐ B. TEST GIVEN
- ☐ ☐ ☐ C. TEST REFUSED
- ☐ ☐ ☐ D. TESTING UNKNOWN
- ☐ ☐ ☐ 97 OTHER _____
- ☐ ☐ ☐ 99 UNKNOWN CONDITION

22 — VIOLATIONS/BEHAVIOR

UP TO TWO CHOICES PER UNIT

UNIT #

- ☐ ☐ ☐ 1 NO IMPROPER ACTION
- ☐ ☐ ☐ 2 SPEED TOO FAST FOR CONDITIONS
- ☐ ☐ ☐ 3 EXCEEDED LAWFUL SPEED
- ☐ ☐ ☐ 4 FOLLOWED TOO CLOSELY
- ☐ ☐ ☐ 5 RAN STOP SIGN
- ☐ ☐ ☐ 6 DISREGARDED TRAFFIC SIGNAL
- ☐ ☐ ☐ 7 MADE IMPROPER TURN
- ☐ ☐ ☐ 8 DROVE/RODE IN OPPOSING TRAFFIC LANE
- ☐ ☐ ☐ 9 KNOWINGLY OPERATED WITH FAULTY/MISSING EQUIPMENT
- ☐ ☐ ☐ 10 REQUIRED MOTORCYCLE SAFETY EQUIPMENT NOT USED
- ☐ ☐ ☐ 11 PASSED IN NO PASSING ZONE
- ☐ ☐ ☐ 12 UNSAFE LANE CHANGE
- ☐ ☐ ☐ 13 FAILED TO KEEP IN PROPER LANE
- ☐ ☐ ☐ 14 DISREGARDED PAVEMENT MARKINGS
- ☐ ☐ ☐ 15 OTHER UNSAFE PASSING
- ☐ ☐ ☐ 16 INATTENTION/DISTRACTION
- ☐ ☐ ☐ 17 DID NOT USE CROSSWALK
- ☐ ☐ ☐ 18 WALKED ON WRONG SIDE OF ROAD
- ☐ ☐ ☐ 19 ELECTRONIC COMMUNICATIONS DEVICE
- ☐ ☐ ☐ 20 FAILED TO YIELD RIGHT-OF-WAY
- ☐ ☐ ☐ 97 OTHER _____
- ☐ ☐ ☐ 99 UNKNOWN

23 — TRAFFIC UNIT MANEUVER/ACTION

UNIT #

- ☐ ☐ ☐ 1 GOING STRAIGHT AHEAD
- ☐ ☐ ☐ 2 SLOWING IN TRAFFICWAY
- ☐ ☐ ☐ 3 STOPPED IN TRAFFIC WAY
- ☐ ☐ ☐ 4 MAKING LEFT TURN
- ☐ ☐ ☐ 5 MAKING RIGHT TURN
- ☐ ☐ ☐ 6 MAKING U-TURN
- ☐ ☐ ☐ 7 OVERTAKING/PASSING
- ☐ ☐ ☐ 8 CHANGING LANES
- ☐ ☐ ☐ 9 NEGOTIATING A CURVE
- ☐ ☐ ☐ 10 BACKING
- ☐ ☐ ☐ 11 AVOIDING VEHICLE/OBJECT/PEDIC/YCLIST
- ☐ ☐ ☐ 12 ENTERING PARKING POSITION
- ☐ ☐ ☐ 13 LEAVING PARKING POSITION
- ☐ ☐ ☐ 14 PROPERLY PARKED
- ☐ ☐ ☐ 15 IMPROPERLY PARKED
- ☐ ☐ ☐ 16 DRIVERLESS MOVING VEHICLE
- ☐ ☐ ☐ 17 CROSSING ROAD
- ☐ ☐ ☐ 18 WALKING WITH TRAFFIC
- ☐ ☐ ☐ 19 WALKING AGAINST TRAFFIC
- ☐ ☐ ☐ 20 STANDING
- ☐ ☐ ☐ 21 LYING
- ☐ ☐ ☐ 22 GETTING ON/OFF VEHICLE
- ☐ ☐ ☐ 23 WORKING ON/PUSHING VEHICLE
- ☐ ☐ ☐ 24 WORKING ON ROAD
- ☐ ☐ ☐ 97 OTHER _____
- ☐ ☐ ☐ 99 UNKNOWN

24 — LOCATION OF PEDESTRIAN/CYCLIST

UNIT #

- ☐ ☐ ☐ 1 MARKED CROSSWALK at INTERSECTION
- ☐ ☐ ☐ 2 AT INTERSECTION BUT NO CROSSWALK
- ☐ ☐ ☐ 3 NON-INTERSECTION CROSSWALK
- ☐ ☐ ☐ 4 DRIVEWAY ACCESS CROSSWALK
- ☐ ☐ ☐ 5 SCHOOL CROSSWALK
- ☐ ☐ ☐ 6 IN ROADWAY (not in crosswalk/intersection)
- ☐ ☐ ☐ 7 MEDIAN (not not on shoulder)
- ☐ ☐ ☐ 8 ISLAND
- ☐ ☐ ☐ 9 SHOULDER
- ☐ ☐ ☐ 10 SIDEWALK
- ☐ ☐ ☐ 11 ROADSIDE
- ☐ ☐ ☐ 12 OUTSIDE OF TRAFFICWAY
- ☐ ☐ ☐ 13 DEDICATED BIKE LANE
- ☐ ☐ ☐ 14 SHARED-USE PATH
- ☐ ☐ ☐ 15 INSIDE BUILDING
- ☐ ☐ ☐ 97 OTHER _____
- ☐ ☐ ☐ 99 UNKNOWN

04-27NW B R12/09/99

sample

ARIZONA CRASH REPORT

1 *CONTINUED*
POLICE ONLY—FORWARD COPY TO
ADOT TRAFFIC RECORDS SECTION, 044R
206 S. 17TH AVE., PHOENIX, ARIZONA 85007-3233

REPORT ID						Agency Report Number
YEAR	MONTH	DAY	HOUR	HCIC NO.	OFFICER ID NO.	

25 VEHICLE DAMAGED AREA(S) - (CIRCLE UP TO THREE AREAS PER UNIT)

Unit # — 1 2 3 4 5 6 7 8 — 0—NONE, 10—UNDERCARRIAGE, 97—OTHER, 99—UNKNOWN

Unit # — 1 2 3 4 5 6 7 8 — 0—NONE, 10—UNDERCARRIAGE, 97—OTHER, 99—UNKNOWN

Unit # — 1 2 3 4 5 6 7 8 — 0—NONE, 10—UNDERCARRIAGE, 97—OTHER, 99—UNKNOWN

26 GLOBAL POSITION — Latitude: Longitude:

27 —ROADWAY ALIGNMENT
UNIT #
☐ ☐ ☐ 1 - STRAIGHT
☐ ☐ ☐ 2 - CURVE LEFT
☐ ☐ ☐ 3 - CURVE RIGHT
☐ ☐ ☐ 99 - UNKNOWN

28 —LANE

Please enter unit's number and lane of travel before first crash event

UNIT ___	UNIT ___	UNIT ___

0 TWO-WAY CONTINUOUS LEFT TURN
1-9 1=FIRST LANE NEXT TO A MEDIAN THRU 9
10 CROSSWALK
L1 THRU L1 - LEFT TURN ONLY LANES (L1= 1ST LEFT TURN AFTER MEDIAN/ CENTERLINE)
R1 THRU R1 - RIGHT TURN LANES (R1=1ST RIGHT TURN AFTER THROUGH LANES)
BL DEDICATED BIKE LANE
HOV HIGH OCCUPANCY VEHICLE
97 NON-ROADWAY
99 UNKNOWN

29 —EJECTION

0 NOT APPLICABLE
1 NOT EJECTED
2 EJECTED, PARTIALLY
3 EJECTED, TOTALLY
4 UNKNOWN DEGREE
99 UNKNOWN

30 —EXTRICATION

0 NOT APPLICABLE
1 EXTRICATED
99 UNKNOWN

Unit # and Seat Position from front page.
Driver seat position = 11

Unit #	Seat Pos	Ejection	Extrication

31 —SEQUENCE OF EVENTS

SEE EXAMPLE BELOW

UP TO FOUR CRASH EVENTS FOR EACH UNIT IN THE ORDER OF OCCURRENCE

NON-COLLISION

1 OVERTURN/ROLLOVER
2 FIRE/EXPLOSION
3 IMMERSION
4 JACKKNIFE
5 CARGO/EQUIPMENT LOSS/SHIFT
6 FELL/JUMPED FROM VEHICLE
7 THROWN OR FALLING OBJECT
8 OTHER NON-COLLISION _____
9 EQUIPMENT FAILURE (tires, brakes)
10 SEPARATION OF UNITS
11 RAN OFF ROAD RIGHT
12 RAN OFF ROAD LEFT
13 CROSS MEDIAN
14 CROSS CENTERLINE
15 DOWNHILL RUNAWAY

COLLISION WITH PERSON, MOTOR VEHICLE, OR NON-FIXED OBJECT

16 MOTOR VEHICLE IN TRANSPORT
17 PEDESTRIAN
18 PEDALCYCLE
19 RAILWAY VEHICLE (TRAIN, ENGINE)
20 LIGHT RAIL/WAY/RAILCAR VEHICLE
21 ANIMAL, WILD—NON GAME _____
22 ANIMAL, WILD—GAME _____
23 ANIMAL—PET _____
24 ANIMAL—LIVESTOCK _____
25 PARKED MOTOR VEHICLE
26 WORK ZONE/MAINT. EQUIP.
27 STRUCK BY FALLING, SHIFTING CARGO OR ANYTHING SET IN MOTION BY ANOTHER VEHICLE
28 OTHER NON-FIXED OBJ. _____

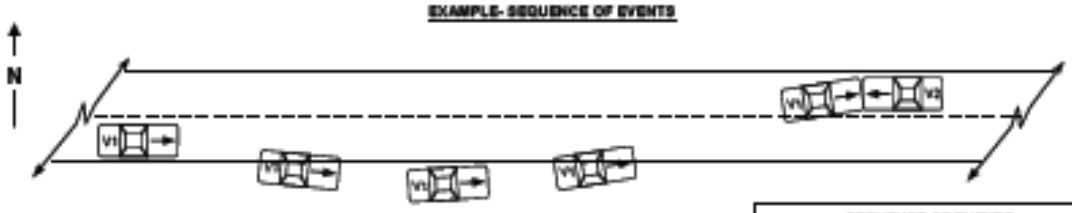
COLLISION WITH FIXED OBJECT

29 IMPACT ATTENUATOR/CRASH CUSHION
30 BRIDGE/OVERHEAD STRUCTURE
31 BRIDGE RAIL
32 CULVERT
33 CURB
34 DITCH
35 EMBANKMENT
36 GUARDRAIL FACE
37 GUARDRAIL END
38 CONCRETE TRAFFIC BARRIER
39 CABLE TRAFFIC BARRIER
40 OTHER TRAFFIC BARRIER
41 TREE, BUSH, STUMP (standing)
42 TRAFFIC SIGN SUPPORT
43 TRAFFIC SIGNAL SUPPORT
44 UTILITY POLE/LIGHT SUPPORT
45 OTHER POST, POLE, OR SUPPORT
46 FENCE
47 MAILBOX
48 BUILDING
49 OTHER FIXED OBJ. _____

99 UNKNOWN

SEQUENCE OF EVENTS			
UNIT ___	UNIT ___	UNIT ___	
			FIRST EVENT
			SECOND EVENT
			THIRD EVENT
			FIRST HARMFUL (based on the crash)

EXAMPLE- SEQUENCE OF EVENTS



VEHICLE 1—SEQUENCE OF EVENTS

11— RAN OFF ROAD RIGHT
14— CROSS CENTERLINE
16— MOTOR VEHICLE IN TRANSPORT

VEHICLE 2—SEQUENCE OF EVENTS

16— MOTOR VEHICLE IN TRANSPORT

SEQUENCE OF EVENTS			
UNIT 1	UNIT 2	UNIT _	
11	16		FIRST EVENT
14			SECOND EVENT
16			THIRD EVENT
Note: Fill FIRST HARMFUL based on the crash			FIRST HARMFUL

04-276B *B* R12/2010

Sample

ARIZONA CRASH REPORT	REPORT ID						Agency Report Number
1 CONTINUED POLICE ONLY—FORWARD COPY TO ADOT TRAFFIC RECORDS SECTION, 064R 206 S. 17TH AVE., PHOENIX, ARIZONA 85007-3233	YEAR	MONTH	DAY	HOUR	NCIC NO.	OFFICER ID NO.	

32 CRASH DIAGRAM

☐ MEASUREMENTS ARE APPROXIMATE AND NOT TO SCALE
☐ MEASUREMENTS ARE SCALED (SCALE = _____)

33 INDICATE NORTH

Sample

01-2764C R08/2010

ARIZONA CRASH REPORT	REPORT ID						Agency Report Number
1 CONTINUED POLICE ONLY—FORWARD COPY TO ADOT TRAFFIC RECORDS SECTION, 064R 206 S. 17TH AVE., PHOENIX, ARIZONA 85007-3233	YEAR	MONTH	DAY	HOUR	NCIC NO.	OFFICER ID NO.	

34 NARRATIVE Describe what happened

Sample

ADDITIONAL PASSENGERS

Unit #	Seat Pos	SD	IS	Name	Address	City	State	Zip Code	Telephone No.	D.O.B./Age	Sex

ADDITIONAL WITNESSES

Name	Address	City	State	Zip Code	Telephone Number	D.O.B./Age

01-2704C R06/2010

Appendix II- Data Completeness for Meeting with GOHS

Preliminary Completeness Analysis: Selected Variables in the ADOT Crash Database (Erin Campbell, MPH Intern; September 30, 2012)

For calendar years 2010 and 2011 more than 580,000 people in 406,789 vehicles were involved in 210,382 crashes reported to the Arizona Department of Transportation (ADOT) Crash Database comprised of more than 170 elements. Of the 580,000 individuals involved in a crash, only 129,761 (22.4%) were documented as having at least a possible injury. These 129,761 individuals occupied 107,328 vehicles involved in 85,623 crashes.

Accurate and consistent linkage of the Crash Database with the Arizona State Trauma Registry (ASTR) will require completeness and consistency of the variable capture in the ADOT Crash Database. Currently, crash records are filled out by responding law enforcement officers and then input into the Crash Database by ADOT/Traffic Records Section staff. This report analyzes the Crash Database's completeness recorded data in selected (targeted) field variables.

Table 1A shows the count and percent missing for variables reported for each individual documented as having at least a possible injury:

- The percentage of records missing Airbag usage is significantly higher than the other variables in Table 1A. This may be for one of two reasons: 1) uncertainty of how officers record airbag deployment for passengers not seated in the first row other than 0 Not Applicable, 1 Not Deployed; and 2) Airbag use is recorded in the "Safety Device" variable (6 Air Bag Deployed, and 7 Air Bag Deployed/Shoulder-Lap Belt).
- Zip Code is another field with a high percentage of a missing variable. In this instance, Zip Code is referring to the individual's Zip Code of residence, which may be difficult to obtain at a busy crash scene, and populated later by obtaining the individual's driver's license or state identification.
- The "Transported To/By" variable is not consistently populated (e.g., variations in hospital name, hospital address, blank, etc.). Ideally, the alpha-numeric ADHS facility codes (e.g., Banner Good Samaritan Medical Center = MED0219) should populate this variable. If one looks only at the "Transported To/By" variable for persons transported, about one-third of the observations include the alpha-numeric ADHS facility code. Consistent entry of the ADHS facility codes in the "Transported To/By" variable would greatly facilitate matching ASTR records with Crash Records.

Table 1A. Completeness of Variables for Persons (n=129,761)

Variable	Count (%)	
Gender	Reported	115,756 (89.2%)
	Unknown or Missing	14,005 (10.8%)
Age	Reported	110,234 (85%)
	Missing	19,527 (15%)
Seat Position	Reported	119,542 (92.1%)
	Missing or Unknown	10,219 (7.9%)
Safety Device	Reported	102,374 (78.9%)
	Missing or Unknown	27, 387 (21.1%)
Airbag	Reported	31,824 (24.5%)
	Missing or Unknown	97,937 (75.5%)
Injury Status	Reported	101,615 (78.3%)
	Missing or Unknown	28,146 (21.7%)
Ejection	Reported	75,453 (58.1%)

	Missing or Unknown	54,308 (41.9%)
Extraction	Reported	75,741 (58.4%)
	Missing or Unknown	54,020 (41.6%)
Zip code	Reported	18,236 (14.1%)
	Missing	111,525 (85.9%)
Transport Source	Reported	33,521 (25.8%)
	Missing or Unknown	96240 (74.2%)
Transported To	Reported***	118,953 (92%)
	Missing or Unknown	10,808 (8%)
Transported To/By (Transport is EMS) N= 13,019	Correctly Reported	4,500 (35%)
	Report Some Destination	4,240 (32.6%)
	Missing	4,279 (32.4%)
***Report some value. Not analyzed for validity		

There were 11 total variables reported in Table 1A. Table 1B shows the number of variables that are missing for each person record. Currently only 3.69% of records contain all 11 elements looked at in Table 1B.

Table 1B. Completeness of Records by Person

Number of Variables Missing	Frequency	Percent
0	4,785	3.69
1	11,336	8.74
2	18,987	14.63
3	27,712	21.36
4	12,515	9.64
5	27,379	21.10
6	7,692	5.93
7	7,167	5.52
8	3,542	2.73
9	6,727	5.18
10	1,905	1.47
11	14	0.01

Unlike the incomplete records per person, the location of the incident is more accurately reported. Over a quarter of incidents are missing the street crossing, but given the small numbers that are missing the latitude and longitude, it makes the location of the accident a very accurate and complete variable. There are only 819 incidents that are missing both latitude/longitude and a crossing street, which represents just 1% of the incidents.

Table 2. Completeness of Variable for Incident (n=85,623)

Variable	Count (%)	
Latitude/Longitude	Reported	84,574 (98.8%)
	Missing	1,049 (1.2%)
Crossing	Reported	62,895 (73.5%)
	Missing	22,728 (26.5%)
On Road	Reported	85,623 (100%)
	Missing	0 (0%)
Hour	Reported	85,623(100%)
	Missing	0 (0%)

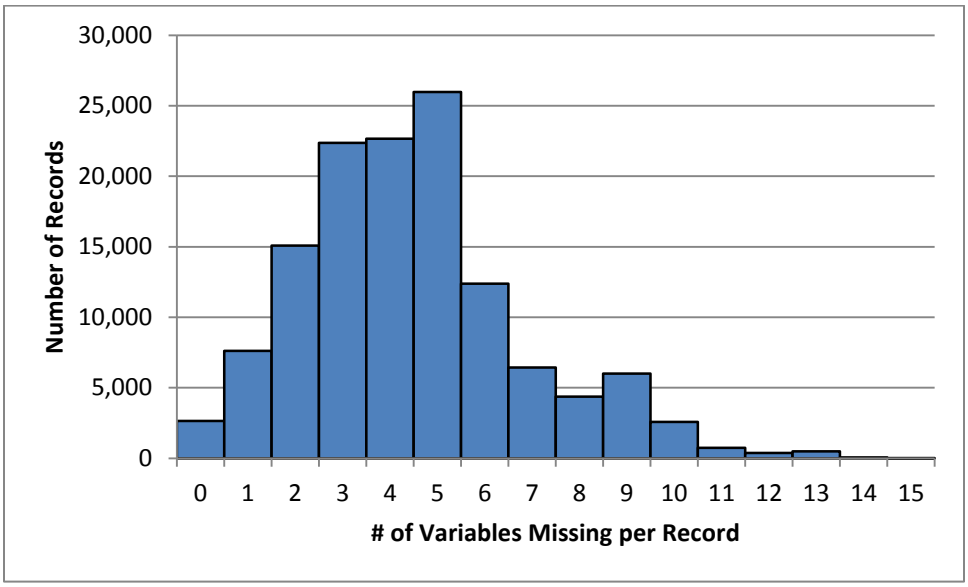
Table 3 looks at the estimated speed of the vehicle at time of impact, the posted speed of the location, and the difference between the two values. Speed has been shown to be highly predictive of the severity of injuries resulting from a motor vehicle accident. In regards to posted speed, it was necessary to correct a small number of records that were reported incorrectly. This small error in addition to the missing posted speeds can be corrected by utilizing the location of the accident. There are also a large amount of records (10.9%) with a posted speed of zero miles per hour. This may reflect the lack of posted speed limits on some roads. The actual speed of the vehicle is more predictive of the severity of the accident but the posted speed is necessary to determine if a vehicle was being operated at a speed in violation of law. Of the 107,328 vehicles involved in a potential injuring crash, just over 1% (1,272 vehicles) of those records is missing all three elements of speed.

Table 3. Completeness of Variable for Unit (n=107,328)		
Variable	**Count (%)**	
Posted Speed	Reported	105,711 (81.5%)
	Missing	1,617 (18.5%)
Estimated Speed	Reported	99,615 (76.8%)
	Missing	7,713 (23.2%)
Speed Over	Reported	99,270 (76.5%)
	Missing	8,058 (23.5%)

Table 4 shows the number of missing variables when the three files (Incident, Person, and Unit) are merged. There were 15 variables in all analyzed for missing variables.

Table 4. Completeness of Records (Person, Incident, Unit) N=129,761		
Number of Variables Missing	**Frequency**	**Percent**
0	2,644	2.04
1	7,610	5.86
2	15,087	11.63
3	22,355	17.23
4	22,662	17.46
5	**25,982**	**20.02**
6	12,373	9.54
7	6,425	4.95
8	4,380	3.38
9	6,004	4.63
10	2,575	1.98
11	728	0.56
12	381	0.29
13	498	0.38
14	50	0.04
15	7	0.01

Figure 1. Graphic Representation of Table 4



CURRENTLY WORKING ON:

Completeness of Variables in ALISS records linked to ASTR

Currently, the two databases have been linked utilizing date of injury, age, gender and the hour of the incident. There were 23,787 ASTR records that indicated that the mechanism of injury was motor vehicle related. Of these 23,787 records, 17,573 or 73.9% have been potentially linked to records in the crash database. Of the more than 17,000 potential matches, more than 10,000 of them (61.8%) match on all four fields. The remaining values match on date and gender and are potentially linked based on missing ages or time in one or both databases. A cushion of +/- 2 hours was used to account for possible time discrepancies.

Appendix III- IRB Approval

Mel & Enid Zuckerman College of Public Health

Non-Applicable Form for a Student Field-based Project

*(***Note: This form is used for projects which do **NOT** involve human subjects, (i.e. course/ program/policy development, or evaluation, etc., **OR** uses data which is **de-identified or** is already **public domain** ***)*

PRINCIPAL INVESTIGATORS: Name(s) of both Advisor and Student:
(By signing this form (below) the student and advisor assure that all information provided on this form is accurate and complete.)
(**Please type this form out OR please write legibly.)

Student Name: Erin Campbell (MEZCOPH) **Advisor Name:** Dr. Denise Roe

Student Email: erin06@email.arizona.edu **Advisor Email:** droe@email.arizona.edu

Title of Project: Biostatistical Analysis of AZ Prehospital Trauma Incident Data

Project Period: 07/2012 to 04/2013 **Project Location and Sponsoring Agency (if pertinent):** ADHS

Is this an **MPH** ■ or **MD-MPH** ☐ (select one) internship project? Yes■ No☐ **If no**, for what course or program or purpose is this project being performed?
(**Note that this form does not apply if you are an MS or PhD student, since these programs require thesis publication. **)

Human Subjects Certification Completed? Yes ■ No ☐ Date of Completion: 06/03/2012

Please indicate which **CITI** certification test completed: **SBR** ■ **BIOMED** ☐ (attach completion report) ☐

Do you intend to present (poster or presentation) this project at any other academic or public forum **other than** the **MEZCOPH Internship Conference?** Yes ☐ No ■

Does the student (or advisor) intend to have the results of the project published in any way? Yes ■ No ☐

(1) **Purpose and Objectives of Project:** (Describe the details (briefly) of **what you will be doing** in this project.)

*Please type this out on a separate Word document & only provide a **brief** paragraph summary.*

******DO NOT cut and paste your MEZCOPH Internship Plan/Contract or paste an Abstract*****

(2) **Check the appropriate category(s) for why this project should be considered as Non-applicable to Official IRB review and explain "why":** ☐ No Human Subjects Involved ☐ Using publicly available Data ☐ Creating a course/program ☐ Evaluating a program ■Using De-identified data (must provide an auth letter)

(2B) *Provide an **explanation**, with your Objectives summary above, and only provide a **brief** paragraph.*
******DO NOT cut & paste your MEZCOPH Internship Plan or attach research definitions******

Student Signature: Erin Campbell (MEZCOPH) **Advisor Signature** Denise Roe

Student Contact Phone: 480-215-1919 Advisor Contact Phone: 520-626-3975

Date: 7/06/12 Date: 7/18/12

(By signing above, the Student and Advisor verify that the proposed student project falls entirely within the scope of the MEZCOPH Non-Applicable requirements, including that any data used are based on publicly available or previously collected de-identified data as defined on page 2 of this form.)

If questions, please contact Dr. Duane Sherrill at (520) 626-7513, dls3@email.arizona.edu
OR Suzanna Trejo at 626-8315, suzanna@email.arizona.edu
When completed, turn form in to:
MEZCOPH Research Office (RmA317C or A317DD)
PO Box 245163, Tucson, AZ 85724 (Fax: 520-626-6093)

The goal of this project is to perform a probabilistic match between two databeses, the Arizona State Trauma Registry and the Arizona Department of Transportation Traffic Records. Both databases provide valuable information as to the nature of injuries sustained, treatment received and overall outcome. By matching up the data between the two databases, the data can be analyzed to show trends in injury severity based on accident type or identify areas of high fatality concentration. The results of this data analysis can be used to better prepare emergency responders and help identify areas where new crash mitigation strategies may be helpful. All Data that is used in this study has been de-identified, as confirmed by the following letter.

Arizona Department of Health Services

Division of Public Health Services
Office of the Assistant Director
Public Health Preparedness Services
150 N. 18th Avenue, Suite 540
Phoenix, Arizona 85007
(602) 364-3150
(602) 364-3567 Fax

JANICE K. BREWER, GOVERNOR
WILL HUMBLE, DIRECTOR

TO: **Erin Renee Campbell, MPH Candidate, University of Arizona**
FROM: **Dr. David James Harden, JD, Data & Quality Assurance Section Chief**
RE: **Secondary Analysis of De-identified Data**
ATTN: **MEZCOPH Research Office**
DATE: **June 13, 2012**

I have reviewed your data request for your research project entitled, *Biostatistical Analysis of AZ Prehospital Trauma Incident Data and Crash Data*, analyzing epidemiologic trends using de-identified data from the Arizona State Trauma Registry and Department of Transportation Traffic Records for years 2008-2011. I support your project and approve your data request. Please note, all data provided for your study are de-identified, and exclude the following identifiers:

1. names
2. geographic subdivisions smaller that a state, including street address, city, county, precinct, zip code and equivalent geo-codes, except for the initial five digits of a zip code to 000
3. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death, and all ages over 89
4. telephone numbers
5. fax numbers
6. electronic mail addresses
7. social security numbers
8. medical record numbers
9. health plan beneficiary numbers
10. account numbers
11. certificate/license numbers
12. vehicle identifiers and serial numbers including license plate numbers
13. device identifiers and serial numbers
14. Web Universal Resource Locator (URL)
15. biometric identifiers, including finger or voice prints
16. full face photographic images and any comparable images
17. internet protocol address numbers
18. any other unique identifying number characteristic or code

This authorization covers the time period of July 2012 through April 2013. We look forward to working with you.

Sincerely,

Dr. David J. Harden, JD, NREMT
Data and Quality Assurance Section Chief
Bureau of EMS and Trauma System

Health and Wellness for all Arizonans

CITI Collaborative Institutional Training Initiative

Human Research Curriculum Completion Report
Printed on 7/6/2012

Learner: Erin Campbell (username: erinrcamphell)
Institution: University of Arizona
Contact Information Department: Public Health
Email: erin06@email.arizona.edu
Social & Behavioral Research Investigators:

Stage 1. Basic Course Passed on 06/03/12 (Ref # 8037820)

Required Modules	Date Completed	Score
UA - Native American Module	06/03/12	no quiz
History and Ethical Principles - SBR	06/03/12	4/4 (100%)
Defining Research with Human Subjects - SBR	06/03/12	5/5 (100%)
Basic Institutional Review Board (IRB) Regulations and Review Process	06/03/12	5/5 (100%)
Informed Consent - SBR	06/03/12	4/5 (80%)
Privacy and Confidentiality - SBR	06/03/12	3/5 (60%)
Conflicts of Interest in Research Involving Human Subjects	06/03/12	4/5 (80%)
Biomedical 200 Refresher Course - Research Involving Vulnerable Subjects	06/03/12	1/1 (100%)

For this Completion Report to be valid, the learner listed above must be affiliated with a CITI participating institution. Falsified information and unauthorized use of the CITI course site is unethical, and may be considered scientific misconduct by your institution.

Paul Braunschweiger Ph.D.
Professor, University of Miami
Director Office of Research Education
CITI Course Coordinator

Return

References

1. Naumann RB, Dellinger AM, Zaloshnja E, Lawrence BA, Miller TR. Incidence and total lifetime costs of motor vehicle-related fatal and nonfatal. *Traffic Inj Prev.* Aug 2010;11(4):353-360.

2. *Arizona Motor Vehicle Crash Facts 2010.* Phoenix, AZ2010.

3. *Arizona Motor Vehicle Crash Facts 2011.* Phoenix, AZ2011.

4. Administration UDoTNHTS. *Problems, Solutions and Recommendations for Implementing CODES.* 2001.

5. Officials NAoSEMS. Incorporation of EMS Patient Care Data. *In State Data Linkage Programs*. Virginia2012.

6. van Schalkwyk I. *The Economic And Injury Burden of Motor Vehicle Crashes In Arizona For 2005.* Arizona: Arizona State University; July, 2007 2007.

7. Cook LJ, Olson LM, Dean JM. Probabilistic record linkage: relationships between file sizes, identifiers and match weights. *Methods Inf Med.* Vol 40. Germany2001:196-203.

8. Brenner H, Schmidtmann I, Stegmaier C. Effects of record linkage errors on registry-based follow-up studies. *Stat Med.* Vol 16. England1997:2633-2643.

9. Fleming M, Kirby B, Penny KI. Record linkage in Scotland and its applications to health research. *J Clin Nurs.* Oct 2012;21(19-20):2711-2721.

10. Newcombe HB, Kennedy JM, Axford SJ, James AP. Automatic linkage of vital records. *Science.* Oct 16 1959;130(3381):954-959.

11. Mason CA, Tu S. Data linkage using probabilistic decision rules: a primer. *Birth Defects Res A Clin Mol Teratol.* Nov 2008;82(11):812-821.

12. Administraion NHTS. Seat belt use in 2011: use rates in the states and territories. Traffic Safety Facts, (D)T HS 811 561. Washington (DC): US Department of Transportation; 2012.

13. Thygerson SM, Merrill RM, Cook LJ, Thomas AM, Wu AC. Epidemiology of motor vehicle crashes in Utah. *Traffic Inj Prev.* Feb 2011;12(1):39-47.

14. Harrell FE. *Regression Modeling Strategies.* New York: Springer-Verlag New York, Inc; 2001.

15. Cummins JS, Koval KJ, Cantu RV, Spratt KF. Risk of injury associated with the use of seat belts and air bags in motor. *Bull NYU Hosp Jt Dis.* 2008;66(4):290-296.

16. McGovern MK, Murphy RX, Jr., Okunski WJ, Wasser TE. The influence of air bags and restraining devices on extremity injuries in motor. *Ann Plast Surg.* May 2000;44(5):481-485.

17. Burgess AR, Dischinger PC, O'Quinn TD, Schmidhauser CB. Lower extremity injuries in drivers of airbag-equipped automobiles: clinical and. *J Trauma.* Apr 1995;38(4):509-516.

18. Orsay E, Holden JA, Williams J, Lumpkin JR. Motorcycle trauma in the state of Illinois: analysis of the Illinois Department. *Ann Emerg Med.* Oct 1995;26(4):455-460.

19. Houston DJ, Richardson LE, Jr., Neeley GW. Mandatory seat belt laws in the states. A study of fatal and severe occupant. *Eval Rev.* Apr 1996;20(2):146-159.

20. *Seat Belt Use in 2012- Overall Results.* National Highway Traffic Safety Administration;2012.

21. Aarts L, van Schagen I. Driving speed and the risk of road crashes: a review. *Accid Anal Prev.* Vol 38. England2006:215-224.

22. Kent R, Viano DC, Crandall J. The field performance of frontal air bags: a review of the literature. *Traffic Inj Prev.* Mar 2005;6(1):1-23.