Driver Factors in Fatal and "A" Injury Crashes in North Carolina (2001-2005)

Final Report

Prepared for
North Carolina Department of Transportation
Research and Analysis Group
Raleigh, North Carolina

by

Libby Thomas
Scott Masten
Carol Martell

THE UNIVERSITY OF NORTH CAROLINA
HIGHWAY SAFETY RESEARCH CENTER

May 2008

THIS PAGE IS INTENTIONALLY BLANK

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/NC/2007-14A		

4. Title and Subtitle

Driver Factors in Fatal and "A" Injury Crashes in North Carolina (2001-2005)

5. Report Date
May 15, 2008

6. Performing Organization Code

7. Author(s)
Libby Thomas, Scott Masten, Carol Martell

8. Performing Organization Report No.

9. Performing Organization Name and Address
University of North Carolina
Highway Safety Research Center
730 Martin Luther King Jr Blvd., Suite 300
Chapel Hill, North Carolina 27514

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
North Carolina Department of Transportation
Research and Analysis Group
1 South Wilmington Street
Raleigh, North Carolina 27601

13. Type of Report and Period Covered
Final Report

September 1, 2006 – August 31, 2007

14. Sponsoring Agency Code
2007-14

Supplementary Notes:

16. Abstract
This project used data from the North Carolina crash database system and driver license history files to determine driver risk factors that predict fault in fatal and serious injury crashes involving passenger vehicles. After analyzing univariate trends, logistic regression models were developed using driver, vehicle, and environmental factors from the crash data only, from the driver history data only, and using both (combined models) to predict fault in multi-vehicle crashes and single-vehicle crash involvement. The comparison group for both types of analyses was comprised of the non-culpable (no contributing circumstances cited) drivers involved in multi-vehicle collisions.

Results were similar to earlier studies indicating that driver alcohol/drug impairments, impairments due to sleep/fatigue, or other causes, lack of safety belt use, young driver age at the time of the study crash; and having prior at-fault crashes, driving on a suspended license or unlicensed, and having a graduated driver license restriction in the prior five years were reliable predictors of fault in both multi-vehicle and single-vehicle crashes. Older drivers were associated with fault in multi-vehicle, but not single-vehicle collisions. Few other driver history factors were reliable predictors. Driver risk factors were more strongly associated with involvement in single-vehicle collisions than fault in multi-vehicle crashes and model prediction efficiency reflected this result. Additional environmental factors were also associated with single-vehicle crashes including night-time occurrence, rural, and higher-speed locations. These latter factors may also suggest in part driver behavioral patterns. Discussion of potential countermeasures and further research are also provided.

17. Key Words	18. Distribution Statement
High risk drivers, impaired drivers, teenage drivers, single-vehicle accidents, multiple vehicle accidents, accident exposure, traffic violations, driving without a license	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	123	

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

Disclaimer

The contents of this report reflect the views of the author(s) and not necessarily the views of the University. The author(s) are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of either the North Carolina Department of Transportation or the Federal Highway Administration at the time of publication. This report does not constitute a standard, specification, or regulation.

Acknowledgments

The authors would like to thank North Carolina Department of Transportation (NCDOT) for the opportunity to perform this research, and the NCDOT project advisory panel. David Harkey, Director, and William Hunter, Senior Research Scientist of the University of North Carolina Highway Safety Research Center reviewed portions of this manuscript.

Executive Summary

Since 2001, North Carolina (NC) has experienced more than 1500 traffic fatalities and on average, another 4200 serious (A-type) injuries each year. While the rate of fatalities per 100 million vehicle miles traveled has declined about 12%, from 1.7 to 1.5 between 2002 and 2005, the absolute number of crashes has remained essentially the same since 2001 and risen since 1995. Both rising State population and increasing vehicle miles of travel per person contribute to increasing congestion and challenges to further reduce fatalities and injuries on the State's roadways. The margin for driver error may be growing ever smaller, and it is increasingly important that drivers safely share the road by avoiding aggressive speeding, alcohol use, and other behaviors that may lead to serious crashes.

The North Carolina Department of Transportation (NCDOT) recently adopted a new mission statement and goals that specifically include making the transportation network safer [http://www.ncdot.org/programs/dashboard/]. The NC Executive Committee for Highway Safety has also established goals to prioritize, implement, and evaluate multi-disciplinary policies and programs to reduce the toll of crash fatalities and injuries on the State's roadways. Efforts to improve traffic safety must clearly target driver factors as well as roadway elements, as driver error is involved in nearly all collisions. This study contributes to several state and national safety goals by identifying and describing key driver-related risk factors that will aid efforts to prioritize and develop effective strategies to reduce crashes and injuries on the State's roadways.

State departments of transportation or licensing offices have long been interested in ways of identifying risky drivers so that licensing control or other actions may target those most likely to contribute disproportionately to future crashes. There is significant prior evidence that, particularly at-fault drivers in serious crashes differ from other drivers in identifiable ways, but as far as we are aware, an in-depth analysis of multiple, driver crash factors including driver history data elements has not been carried out using NC data.

This project compiled and merged data from fatal and "A-type" injury crash records for the years 2001 to 2005 with driver license history data and crash data for five years prior to the study crash for an in-depth examination of driver characteristics based on the information available in those databases. The aims of the study were to:

- Identify driver factors associated with greater risk of 'causing' serious injury and fatal crashes.
- Assess whether data and analyses resulting from the addition of license histories adds significantly to our understanding of at-fault drivers in serious injury crashes beyond the information available from crash reports alone, and improves ability to correctly 'predict' actual fault status of drivers.
- Complement engineering and other current highway safety efforts by providing details of driver and other crash risk factors that may be targeted by changes in policies and procedures, enforcement and adjudication efforts, and innovative programs. The information contained in this report may contribute to efforts to

address some of the more intractable problems related to reducing the toll of highway fatalities and injuries in North Carolina, and start a serious discussion about ways to create a traffic safety culture in the State.

Distributional tables of driver factors and limited environmental factors were created for the entire study sample of drivers involved in fatal and A-injury crashes that included only passenger vehicles on North Carolina Roadways during the years 2001-2005. Tables of factors by fault group (not at fault, multi-vehicle at fault, and single-vehicle) and cross-tabulations of certain driver factors were examined to determine general involvement trends and likely predictors of fault in the multiple logistic regression equations. Fault was determined by whether drivers were indicated to have any contributing circumstances in the crash, with those having no contributing circumstances in multi-vehicle collisions serving as the not-at-fault comparison group for both multi-vehicle at-fault and single-vehicle groups. Representation of contributing circumstances differed between the two groups of at-fault drivers as did a number of driver characteristics. Speeding, reckless driving, and alcohol-related contributing circumstances were most highly represented in single-vehicle crashes, whereas failure to yield right-of-way, crossing centerline, and inattention/distraction were most highly represented among those deemed at fault in multi-vehicle collisions. These characteristics supported the decision to conduct separate models for multi-vehicle involvement and single-vehicle crashes.

Eight different multivariate models were developed to analyze the relationship between driver factors and fault in NC fatal and serious injury crashes. In particular, separate models were developed to predict fault in ***multi-vehicle crashes*** (models 1 to 4) and ***single vehicle crashes*** (models 5 to 8) using:

- All crash involved drivers (including out-of-state licensees or residents) – using crash data only.
- NC drivers using crash data only.
- NC drivers (licensed or those that could otherwise be matched with a driver history) using license history data only.
- NC drivers using crash and license history data (combined models).

The models supported the supposition that driver characteristics and behaviors may predict fault in a crash and identified a number of factors associated with fault in multi-vehicle and single-vehicle collisions. Driver factors consistently predictive of fault in **both multi-vehicle and single-vehicle** collisions included:

- Young driver age (especially 16 to 19 years; also 20 to 29 years, compared with middle-aged drivers)
- Driver unbelted in study crash
- Driver impaired by alcohol, drugs, or medication
- Driver sleepy/fatigued

- Driver impaired by other condition (medical, physical)
- Alcohol/drug use suspected/detected
- Driving a passenger car (versus other passenger vehicles – pickups, SUVs, vans/minivans; SUVs were also associated with fault in single-vehicle collisions)
- Driving on suspended license
- Driving unlicensed
- Having prior at-fault crashes in five years preceding the study crash
- Having a graduated driver license restriction in the five preceding years

Other factors associated with fault in **multi-vehicle** collisions included:

- Being an older driver (beginning with slight increase for 60 to 69 years and increasing with age)
- Having a commercial driver license
- Reported to be Hispanic or Asian/other in study crash
- Having no passengers compared with having one or more passengers.
- Driving an older model vehicle
- Having no prior not-at fault crashes

Other factors consistently associated with fault in **single-vehicle** collisions included:

- Having a daylight driving only restriction
- Having two or more "Other" serious convictions (such as reckless driving)
- Having two or more passengers was associated with fault in single-vehicle crashes.

Other variables were predictive of either multi-vehicle or single-vehicle collisions, or predictive in the license-only, or crash-data only, or combined models only. These are discussed more fully in the report body.

While the oldest drivers, 80 and up, had the highest odds of being at fault of any age group, even higher than teens, their overall crash involvement was only 2% of the entire study sample. Teen drivers (only 16 to 19) accounted for 14% of those involved and the 20 to 29 year group for nearly 28% of all drivers. The relationship of the 20 to 29 year age group to fault was not nearly as pronounced as that for teen drivers, suggesting that part of their involvement is due to greater exposure.

Being male was predictive of fault only in single-vehicle collisions when only license data were used, suggesting again that exposure and risk factors captured in the crash itself may account for the apparent over-representation of males in at-fault crashes, (71% of single-vehicle and 62% of multi-vehicle fault group compared with 58% for not-at-fault drivers). However, both males and the 20 to 29 year age group were highly over-represented among those indicated to have alcohol impairments or reported not using a seat belt in the study crash, both of which were identified risk factors

Certain environmental factors also characterized involvement in single-vehicle collisions compared with multi-vehicle (not at fault comparison group) including driving on rural roads and interstates, on higher speed limit roadways, and at nighttime.

The models to predict single-vehicle fault were more efficient at classifying drivers correctly than those developed for multi-vehicle involvement. These results are perhaps not unexpected, since the group of at-fault drivers in multi-vehicle crashes were more similar to the not-at-fault drivers involved in the same collisions. The strongest driver predictors of fault including impairments were also even more strongly predictive (based on odds ratios) of fault in single-vehicle collisions. Although not incorporated as model factors since they were used to determine fault in the crash, representation of contributing circumstances were also different between the two groups of drivers. Speeding, reckless driving, and alcohol-related contributing circumstances were most highly represented in single-vehicle crashes, whereas failure to yield right-of-way, crossing centerline, and inattention/distraction were most highly represented among those deemed at fault in multi-vehicle collisions.

Models based on driver history data only were less efficient at discriminating accurately among not-at-fault and at-fault drivers than those based on crash data only. And the addition of data currently available from the drivers' histories in combined models added only marginally (1.1%, single-vehicle, 4%, multi-vehicle) to the ability of crash-based models to correctly classify risky (at-fault) drivers. Possible reasons for these findings include: 1) The data from the crash record itself is obviously strongly associated with fault in that crash (including such factors as alcohol and drug-related impairments, sleep/fatigue, etc. 2) The two sets of data (crash-based, and driver prior records) may be measuring similar sets of underlying attributes and risk factors so there will be significant overlap in predictive power. And, 3) Drivers' prior histories do not adequately capture drivers' risk for a number of reasons – drivers must first of all be detected in violations and accurately identified in records for both violations and prior crashes, and they must be convicted for the violations actually committed. If convictions do not reflect guilty violations, then the driver history will not capture risk associated with these prior behaviors. There is considerable anecdotal evidence, including from investigative news reports in The News and Observer on what happens to those caught speeding in NC that convictions often do not reflect driver violations. Unfortunately, for the present study, we were unable to incorporate the use of violation data from the Administrative Office of the Courts to determine if these data capture additional predictive power.

The driver characteristics associated with fault in serious injury crashes in NC are supported by findings from a number of previous studies in other jurisdictions. A key follow-up objective should be to review, identify, and otherwise develop effective methods to target the problems that are associated with a high proportion of serious injury crashes, particularly those with a known causal association. Discussion of the findings, potential countermeasures and recommendations for further research are provided.

THIS PAGE IS INTENTIONALLY BLANK

Table of Contents

List of Tables

List of Figures

THIS PAGE IS INTENTIONALLY BLANK

Introduction

Since 2001, North Carolina (NC) has experienced more than 1500 traffic fatalities and on average, another 4200 serious (A-type) injuries each year. While the NC rate of fatalities per 100 million vehicle miles traveled has declined about 12%, from 1.7 to 1.5 between 2002 and 2005, the absolute number of fatalities has remained essentially the same since 2001. NC fatalities were nearly 6% higher than in 1995 (Figure 1). Both rising State population and increasing vehicle miles of travel per person contribute to increasing congestion and challenges to further reduce fatalities and injuries on the State's roadways. The margin for driver error may be growing ever smaller, and it is increasingly important that drivers safely share the road by avoiding aggressive speeding, alcohol use, and other behaviors that may lead to serious crashes.

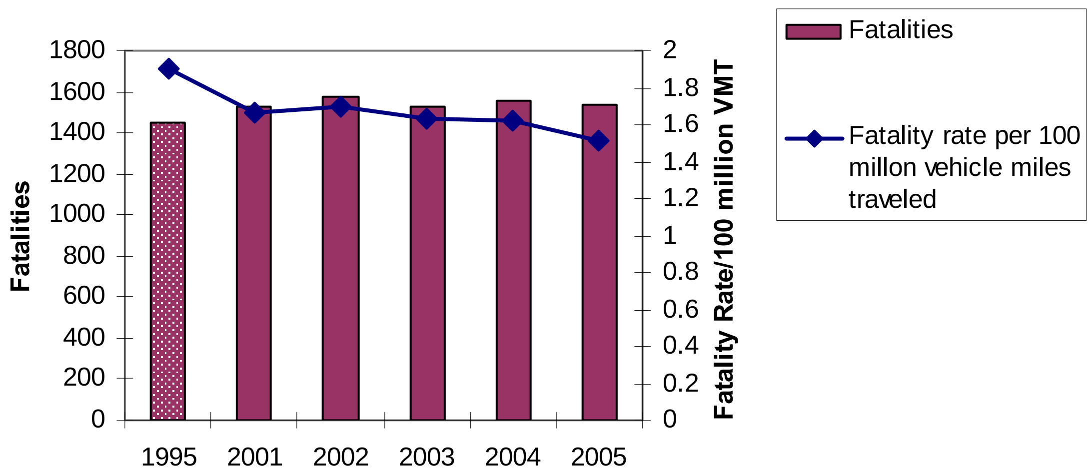


Figure 1. NC Traffic Fatalities and Fatality Rates. (Data from NHTSA's National Center for Statistics and Analysis, Traffic Safety Facts: State Traffic Data, 1995, 2001, 2002, 2003, 2004, and 2005)

The social and economic impacts of these crashes are tremendous and require urgency in prioritizing and dealing with this serious public health issue. Traditional enforcement and roadway engineering efforts to maximize safety cannot prevent nor reduce the severity of all crashes, given the role of driver error in crashes. The human element must be addressed in efforts to meet North Carolina's policy goal of reducing the fatal crash rate to 1.0 per 100 million vehicle miles of travel by 2008, a goal that seems likely to be missed. The North Carolina Department of Transportation (NCDOT) has also recently adopted a new mission statement and goals that clearly identify safety as a priority of the department [http://www.ncdot.org/programs/dashboard/]. Other national and state efforts have focused attention on driver behavioral safety problems including AASHTO's (American Association of State Highway and Transportation Officials) Strategic Highway Safety Plan, the National Forum on Speeding in 2005, and continued efforts to curb impaired driving, reduce young driver crashes, and increase safety belt use.

State departments of transportation have long been interested in ways of identifying risky drivers so that licensing control or other actions may target those most likely to contribute disproportionately to future crashes. Penalty point programs aim to deter risky driving as evidenced through prior crash and conviction records, providing a mode to sanction and limit drivers through license restrictions or suspensions. Certain driver demographic characteristics and behavioral indicators, including young age, older ages, lack of seat belt use, and psycho-physical conditions (alcohol, drug, medical conditions, sleepiness), have also frequently been shown to be correlated with higher crash propensity.

In the present study, we analyze the associations of driver factors including demographic characteristics and behavioral indicators from the crash itself, as well as conviction and crash data from the drivers' prior histories, with each driver's culpable association with crashes resulting in fatal or serious injury. As far as we are aware, an in-depth analysis of crash factors that include driver history information for this subset of drivers (at-fault drivers in serious crashes), has not been carried out using North Carolina (NC) data.

Objectives of the Present Study

This project selected fatal and "A-type" crash records for crashes involving only passenger vehicles for the years 2001 to 2005 as the crashes of interest, or study crashes. Information specific to each crash was combined with five years of prior crash data and driver history data for each driver involved in the crash to profile drivers' behavior and other characteristics in the five years preceding the study crash. This compilation enabled an in-depth examination of driver (and limited vehicle) characteristics based on the data available in those databases. The aims of the study were to:

- Identify driver factors associated with greater risk of 'causing' serious injury and fatal crashes.
- Assess whether data and analyses resulting from the addition of license histories adds significantly to our understanding of at-fault drivers in serious injury crashes beyond the information available from crash reports alone, and improves ability to correctly 'predict' actual fault status of drivers.
- Complement engineering and other current highway safety efforts by providing details of driver and other crash risk factors that may be targeted by changes in policies and procedures, enforcement and adjudication efforts, and innovative programs.

The information contained in this report may be used to help begin efforts to address some of the more intractable problems related to traffic safety in North Carolina, and building a traffic safety culture among the public, the media, and involved agencies and institutions State-wide.

The following sections describe results of a review of literature related to identifying driver crash risk factors and risky drivers, the methods and results of the present study, a summary of key findings and discussion of the findings, potential countermeasures, and recommendations for further research.

Background

A number of key earlier studies examining driver risk factors were reviewed in the preparation for developing the methodology for the present study. Efforts to profile drivers for targeted interventions have been around for decades. Some of the earliest research explored identifying risky drivers from their prior records in order to target interventions such as post-licensing control actions. Other approaches have used crash and license history data to gain a greater understanding of driver risk factors in general, which corresponds more closely to the objective of the present study. We were, however, also interested in the possible future application of models that might be used to identify and target risky drivers for post-licensing actions, or in some cases actions targeting unlicensed and suspended-licensed drivers. The primary aim of this review was to evaluate study objectives, methods, and shortcomings of prior research with the goal of developing a sound methodological approach in the present study. A secondary objective was to identify factors that have proven to be associated with crash risk in prior studies. The diversity of objectives, study approaches and methods makes direct comparison of outcomes among studies infeasible, other than identifying in a general way, factors that have proven to be correlated with crash involvement in a number of earlier studies.

Driver "profiles"

An early study conducted by the Highway Safety Research Center examined the relationships between prior crashes and convictions with crashes in a subsequent period (Stewart and Campbell, 1972). Part of the intent of this study was to illustrate that although subsets of drivers with prior convictions and/or crashes do account for more than their share of future crashes, the proportion of all crashes that can be "predicted" by prior crash involvement or convictions is relatively small (Stewart and Campbell, 1972). Stewart and Campbell pointed out that focusing only on drivers with prior convictions or crashes would target a small proportion of the overall crash problem. While drivers with 1 or more convictions in a 2-year period (16% of all drivers) were involved in 29% of all crashes in a subsequent 2-year period, 71% of collisions involved those with no convictions in the prior 2-year period. Similarly drivers with one or more crashes on their records (10.7% of all drivers) were involved in about 19% of all crashes, but that leaves 81% of crashes involving drivers with no prior crashes.

A similar statistical tenet also underlies a difficulty in efforts including other measures of risk to identify individual risky drivers for targeted interventions that will reduce crashes significantly. Any particular 'group' of drivers is apt to account for a relatively small percentage of crashes. Despite this, efforts have therefore continued over the years to refine prediction methods to more accurately identify risky drivers and reduce the proportion of false positive identifications. In addition, depending on the particular 'group' characteristic, some groups may account for a significant share of crashes. As the rate of decline of crash rates has slowed, it may be time to think seriously about some of the particular driver crash problems. And some have argued that there is less risk in falsely identifying a driver who turns out to be "innocent" than in missing drivers in need of intervention (Chandraratna 2006). In particular, for non-punitive interventions such as warning letters, safe driver programs, etc., the risk of falsely identifying ‘innocent” drivers is low. In any case, current driver demerit points systems and insurance rates tied to these penalties, are considered to reflect driver risk, perceived or otherwise determined. Establishing increased risk using sound scientific methods should be sufficient grounds for taking action.

Researchers in the California DMV have, for example, carried out extensive research, beginning in the 1970’s, aimed at using existing crash and driver data to characterize and identify high crash-risk drivers (Gebers, 2003, 1999; Gebers and Peck 2003, 2000, 1994; and others). The California work has also focused on identifying risky drivers from among the general driving population (all licensed drivers) with an objective of targeting post-license control or educational actions. A 2003 report updated an inventory of California Driver Accident Risk Factors (Gebers, 2003). Prior citations (total, countable, and moving), and prior crashes correlations with future crashes (and relative risk factors compared with the no priors cases) were reported. Some general results were that crash risk increases as a function of number of crashes and citations on the prior record. A combination of prior citations and prior crashes produced the best prediction, although “total number of citations” was the strongest single predictor. Total number of crashes was a better indicator of subsequent crash risk than culpable crashes. Current period crashes and convictions were also more strongly correlated than those for non-concurrent time periods. Men had higher risk than women (largely due to exposure per the authors) and younger drivers had the worst records, with older drivers’ (above age 69) risk increasing somewhat.

Stewart and Campbell and some of the California reports focused on single-variable associations with subsequent crashes. Current analysis methods utilize multivariate methods, most commonly multiple logistic regression, which allows the dependent variable to have non-linear relationships with the predictor variables. Furthermore, effects of variables can be determined while controlling for co-variation of other variables in the model. Efforts to develop predictive equations based on prior convictions, prior crashes, and both, have generally been consistent with the early results in that predictive power is better than chance, but accuracy of prediction on an individual basis is still low, resulting in either high false positives or high false negatives (Gebers and Peck, 2000, 2003). In models incorporating demographic factors, Gebers and Peck identified prior 3-year citations, prior 3-year crash involvements, possession of a commercial license, age, gender, being Black, being Hispanic, median household income, presence of a physical or mental condition on record, and presence of a driver license restriction on record as being correlated (in regression models and canonical correlation functions) with total future crashes and total future convictions. A correct classification rate of future crash-involvement of 23% (from a data set of all licensed driers) was obtained with the models. In addition to the continuing trade-offs between identifying a larger set of risky drivers and netting a higher proportion of ‘innocent’ (in the future time) drivers, these authors expressed concerned about the inclusion of demographic factors such as age, gender, ethnic group, and socioeconomic indicators that may be problematic to use in identifying drivers for remedial actions (Gebers and Peck 2000).

Hauer et al. (1991) showed that multivariate models using a number of factors including age, gender, convictions, and crashes (from a large sample of all licensed drivers) from one two-year period were better predictors of subsequent crash involvement than an existing Toronto, Canada driver points/demerit system, particularly when prior crashes were included in the models. Little additional predictive power was gained by distinguishing between types of convictions, or, in these models, whether the driver was at fault in the prior crashes. This result is not surprising however, given that they did not predict fault in the outcome crash, but any crash involvement. These authors concluded that reasonable estimation accuracy could be achieved by simply attaching a weight of 1 to a conviction and 1.88 to any prior crash. Unfortunately, these authors felt compelled to exclude drivers whose licenses were suspended during the study period due to the confounding of reduced exposure. Since drivers with suspended licenses are likely to be among the most risky drivers, if they in fact continue to drive, efforts to remedy this problem in analyses should prove worthwhile.

Another Canadian study focused on the risk of elderly drivers and performed separate regression analyses examining the correlations between prior convictions with crashes and prior crashes with crashes among those 65 years and up (Daigneault, et al. 2002). Regression analyses showed that the correlation coefficient for crash involvement was higher with prior crashes and increased with increasing age whereas the correlation with convictions declined with increasing age.

None of the work discussed thus far accounted for fault in the outcome (predicted) crashes. Thus the lack of improved prediction by distinguishing fault in the earlier crashes as reported by Gebers and Hauer et al is not unexpected, since the outcome measure included both at-fault collisions, and collisions in which the participants may have just been involved by chance. Nor was crash severity considered in the prior models. Another factor that may reduce the strength of the correlation between predictor and outcome variables is that periods are arbitrarily divided into before and after periods. Thus, the most recent prior crashes within the 'after' period cannot serve as predictors for the later crashes that happen to fall in that same period.

Causal crash involvement

The research discussed thus far has focused on all types of crash-involvement and severity of crashes, and without assessing the effects of blame or driver contribution to causing the crash. Different types of crashes may have different causes and therefore should not be predicted as a group according to Wahlberg (2003). Wahlberg also indicates that researchers have tended to find stronger relationships between predictor variables and culpable crashes than with all crashes. Walhberg goes on to provide an extensive analysis of prior research on traffic crash predictors as to whether 1) test-retest reliability of variables has been addressed; 2) the study time period allows for stable crash frequencies and sufficient variance, and is sufficient to overcome the Poisson distribution of crashes per individual, while not unduly reducing the expected correlation between the predictors and the outcome (e.g. relationship of age with behaviors); and 3) whether culpability has been accounted for. He argues that without culpability, the only correlation of crash involvement to be expected is with exposure. His definition of culpability, however, cannot practically be met; that is the line should be drawn between culpable and not culpable if the non-culpable driver could not have influenced the occurrence of the crash (other than by not being there).

Some past research related to identifying risky drivers has accounted for fault status of drivers. Chen et al (1995) estimated "future" at-fault crash involvement based on prior convictions (aggregated into 15 groups based on similarity of offense and comparable points attached) and prior at-fault crashes (not explained). This study also divided the study period into before and after periods. Using this method, a prior crash may end up in the after period with a subsequent crash and thus, will not be able to act as a 'predictor,' as pointed out by Chandraratna et al. (2006). Personal characteristics such as age, gender, territory of residence, etc. were not used in the models. Prior at-fault crash involvement was a stronger predictor than prior convictions. The models successfully identified about 49% of the top 1000 crash involved drivers and 35% of the top 20,000. Using only convictions, about 40% of the top 1000 drivers were correctly identified as risky (as determined by subsequent crash-involvement), while 32% of the top 20,000 were so.

Cooper (1995 and 1997) also reported on driver factors, including types of convictions, and prior crash involvements to predict subsequent culpable crashes. There was some evidence from the 1997 work that associations with types of convictions may not be very reliable due to low frequencies for some types, as well as potentially how "types" are grouped. Earlier work by Cooper et al. (1995) employed logistic regression to identify drivers most likely to have one or more at-fault crash involvements in a subsequent period. Both increasing numbers of prior convictions and prior at-fault crashes were related to increasing post-period crashes. "Prior crash involvements" was more predictive than prior convictions. The number of convictions by category (15 groups), and the number of prior (3-year) crash involvements, were used as the independent variables. The occurrence of one or more (yes/no) at-fault crash involvements in the final 2 years was the dependent variable.

Induced exposure methods

The induced-exposure or quasi-induced exposure methods may provide a solution to dealing with unequal exposures as well as causal crash involvement. These methods assign relative blame for a crash (based on citations or driver contributing circumstances) and use the "innocent" party as a comparison group to represent the general driving population. These methods assume that the 'innocent' party in a crash serves as an effective exposure control. Stamatiadis and Deacon (1997) reviewed the history of induced-exposure methodology tracing back to Thorpe (1967). The method of quasi-induced exposure, described by Carr (1970) introduced the concept of identifying the (most) culpable (or at-fault) driver in a multi-vehicle crash and using the non-responsible group to control for relative exposure.

A 1991 study used Fatal Accident Reporting System (national, fatal collision) data in what essentially was a quasi-induced exposure study with the added component of pairing 'cases' (driver responsible for 'initiating' the collision) with controls (passively involved) from the same collision (Perneger and Smith, 1991). This method, according to the authors, avoids confounding by environmental factors, exposure to traffic, and differences in severity. The strongest independent (driver factor) predictors of causing a crash were alcohol, ages over 80, and young age (teens). Other driver factors included not wearing a seat belt, driving without a valid license, and having had a crash within the past year (but earlier crashes were not significant predictors). These authors concluded that driver errors do not occur randomly, "but are associated with specific driver characteristics." These authors also limited their dataset to studies involving passenger vehicles or light pickups, excluded single vehicle collisions and collisions involving more than two vehicles, and included only ‘clean crashes’ - ones in which one and only one driver was assigned ‘driver-related factors.’ Perneger and Smith did not test the success of their model at predicting crash involved drivers.

Stamatiadis and Deacon (1997) also included only drivers involved in “clean crashes” or those in which only one driver in multi-vehicle crashes was assigned ‘fault,’ in this case determined by the recording of a violation or citation. There was a slight under-representation of younger drivers in the “clean crash” data set and a slight over-representation of older drivers, but given the large sample of clean crashes, this bias was not considered to have a significant effect. Additionally, this is an acknowledged attribute of younger drivers – that is they are less likely to be able to ‘avoid’ a crash; even if they are not the principal at-fault driver.

Stamatiadis and Deacon (1997) tested some of the assumptions of the quasi-induced exposure method and found that the use of non-responsible drivers in multi-vehicle collisions provided an acceptable surrogate for exposure (to multi-vehicle crashes) based on analyses of vehicle type and driver age characteristics. They also found that there were differences in population and time/location characteristics of at-fault drivers in multi-vehicle crashes and drivers in single-vehicle collisions. Thus, it is important to model the two types of outcomes – 1) causing a multi-vehicle crash, and 2) single-vehicle crash involvement – separately since different risk factors may be at work.

Lardelli-Claret et al. carried out induced-exposure studies in Spain to examine risk factors for causing collisions involving two-wheeled motor vehicles (2005) and to compare two quasi-induced exposure methods for examining driver and vehicle-factors on the risk of causing a crash involving at least one four-wheeled motor vehicle (2006). Again, both of these studies included “clean crashes” where only one driver was at fault. In each study Lardelli-Claret et al. determined adjusted odds ratio estimates for each of the driver-related factors related to the outcome measure (causing a ‘clean crash). These analyses included factors from the crash itself (hours driving, medical/physical condition, helmet use, and crash-related infractions recorded at the time of the collision) as well as age, sex, nationality, and years in possession of a valid license at the time of the crash. Inappropriate speed, excessive speed, and driving under the influence of alcohol, were the strongest predictors, followed by younger and older driver age, being a foreign driver, and driving without a valid license.

The later study compared two induced exposure methods, one of which ‘paired’ the case and the control within a crash in analyses to account for potential environmental confounders, and the other used classic quasi-induced exposure (not paired within crash), but allowed environmental factors to be covariates. Both types of analyses (classic and paired) yielded similar estimates of driver risk factors, with psycho/physical condition (alcohol and sleepiness) being the strongest predictors of causing a crash. The odds ratios (OR) for most risk factors also differed significantly between those causing two-vehicle crashes and those involved in single-vehicle collisions, with sleepiness/fatigue being much more correlated with single-vehicle collisions. Thus, the populations of drivers causing single-vehicle collisions do not seem to be the same as those causing multiple vehicle collisions. These models were not used in an attempt to identify risky drivers per se, but to develop an understanding of risk factors for causing a crash.

In the 2006 study, Chandraratna et al. accounted for fault and used only the latest crash involvement so that all of the earlier crash and citation history (eight years) could be used as predictors (without arbitrarily dividing the data into before and after time periods) in a quasi-induced exposure type study. This type of analysis therefore eliminated drivers who did not have at least one earlier crash involvement from the database since their intent was to examine ability to predict recurrent crashes. As in many of the earlier studies, they seem to have used all crashes (any severity) in which only one driver was assigned responsibility, but whether or not this included single-vehicle collisions isn't clear. Several problems may be associated with using all crashes. Reporting thresholds may change over time for property-damage crashes due to increasing costs. And minor fender benders may simply represent greater exposure to congested driving conditions and temporary lapses in attention as opposed to more 'intentional' or inherent risk-taking tendencies.

Through the quasi-induced exposure methods, we seem to be coming closer to developing models that account for exposure, that account for driver 'causality,' or fault in the crash, and that may further our understanding of underlying behavioral/psychological/ demographic driver risk factors. The present study therefore used a quasi-induced exposure method to analyze driver factors associated with fault in serious injury crashes in NC. It is also possible that by combining information from crash files, and driver history files, multivariate analyses of driver crash propensity may be able to better predict who will 'cause' a crash in future.

The following sections describe the methods and formation of the study population and groups, results of univariate and multivariate analyses, a summary of key findings, discussion of results and potential countermeasures as well as recommendations for further research and suggestions for implementation of study findings.

Methods and Study Population

As noted, it was determined to use quasi-induced exposure methods in logistic regression analyses to identify driver characteristics associated with fault in the outcome crash (latest involvement). We determined to select "clean crash" cases in which only one driver was at fault, but did not exclude crashes with more than two vehicles involved. Therefore, the additional "not-at-fault" drivers became part of the control group. Variables from the crash records as well as from five years of prior driver history were compiled and used in the analyses. The determination to use five years of prior history was a compromise to obtain sufficient prior history data without being so long that the strength of the relationship between predictor variables and outcome variable (latest crash involvement) is weakened. Evidence from some of the prior studies showed that more recent crashes, for example are more highly predictive of subsequent crashes, and the same is likely to be true for conviction record. Driver age and experience are also factors that depend more on recent history than older history. Wahlberg noted the importance of having a sufficiently long record to establish stability, while not unduly reducing the expected association between predictors and outcome. In addition, when prior crashes were determined to have occurred from a driver's history (license) file, that crash was looked up in the crash database in order to determine fault status for prior crashes as well.

Crashes resulting in fatal or A-type injuries and that involved only passenger vehicles were selected for the study sample in order to avoid confounding by severity and by very different characteristics of sub-populations of drivers such as commercial truck operators and motorcycle drivers. (Other NC studies were underway about both of the aforementioned groups.) The subsequent sections describe the formation and characteristics of the study population revealed through single variable and cross-tabulations.

Driver Selection

All crash-involved drivers for the years 2001 to 2005 meeting the following criteria were obtained from the North Carolina Crash Database System (NCCDS):

1. The crash involved a severe or fatal injury (KABCO levels K&A) for at least one involved person;
2. Only passenger vehicles, pickup trucks, SUVS, and vans/minivans were involved in the crash;
3. One and only one driver was deemed to be at-fault for the collision based on contributing circumstances coded by law enforcement (contributing circumstances for only one driver);
4. The crash occurred on a publicly accessible roadway (not private property);
5. The driver was at least 16 years old.

The total number of crash-involved drivers during the 5-year period meeting all of these criteria was 21,786. Study characteristics of these drivers are shown in the next section, Table 1.

Variables Considered for Inclusion

Based on a thorough review of the literature focused on developing models to profile or identify crash-involved drivers, or drivers who were at-fault in crashes, a wide range of variables was identified for possible inclusion as predictors of being at-fault in a serious or fatal crash in North Carolina. Both the NCCDS and North Carolina DMV licensing history database were used to create an initial list of potential variables, with an explicit purpose of including or creating best approximations of those used in past studies that were found to be strong predictors. Crash-based data are collected for all persons and vehicles involved in an injury crash in North Carolina, but driver history data are only reliably collected for persons who are residents of the state (though both licensed and unlicensed drivers). Therefore, out-of-state drivers and those with no licensing activity on record are only included in the subset of analyses involving solely crash database predictors. The next two sections describe the initial sets of predictor variables identified and coded from the NCCDS and North Carolina licensing history databases for further analysis.

Crash database predictors

Driver or vehicle-related variables that could be used directly or otherwise created from the study crash records were eligible predictors in the models. Driver factors such as seatbelt or alcohol use in the study crash may seem to lack the proper temporal nature to "predict" fault in that crash in the crash since they come from same crash that was used to identify the driver for the study and in some cases were likely a factor used in determining the fault of the driver. For example, a variable that represented whether or not the driver was exceeding the speed limit prior to the crash would likely be a strong predictor of fault, because such information is considered by law enforcement officers when they are coding the contributing circumstances we used to establish fault for purposes of this study. For these reasons, we did not use the contributing circumstances cited to determine fault in the crash as predictor variables since having contributing circumstances would be wholly confounded with the fault groups. But we did incorporate factors such as seatbelt indicators or alcohol indicators outside of the contributing circumstances variables. These variables are likely, to capture underlying driver risk that may reflect patterns of behavior, and thus, since such measures are scarce, they were used in these models. Other variables from the study crash records such as sex and age of the drivers indicate pre-crash status or risk characteristics of the drivers and were also used as predictors.

Univariate analyses were used primarily to characterize the study population – to determine the proportion of drivers having various levels of study factors. Cross-tabulations of study factors by fault group were examined to determine which factors would most likely be predictive of fault in the study crashes. Other cross-tabulations were performed to assess how certain factors varied across some other factor – for example, impairments by age group.

The set of 24 crash-based predictor variables is shown in Table 1, along with frequencies and percentages of involvement for the various levels of the factors across all 21,786 crash-involved drivers. The initial crash database predictor variables were grouped into the following categories in the table to ease their presentation: (a) License/Registration-Related, (b) Driver-Related, (c) Vehicle-Related, and (d) Environment-Related.

The environment-related variables such as hour of the crash, roadway type and speed limit, and crash locality would not be expected to be predictive of fault when comparing the two groups of multiple-vehicle crash-involved drivers because they were by definition closely matched on these factors given they were involved in the same set of crashes. The multiple-vehicle not-at-fault and single-vehicle at-fault drivers are not matched in any way on these environmental factors, however, so it is necessary to try to statistically equalize these groups based on the available information so that the other crash and licensing characteristics can be compared in a less-biased fashion. Hence, beyond their use for descriptive purposes, these environmental variables were also coded for use as covariates when comparing the single-vehicle at-fault group to the multiple-vehicle not-at-fault drivers as was done by Lardelli-Claret et al. (2005) since these groups are not inherently matched on environmental factors.

Table 1. North Carolina Crash Database Predictor Variables with Descriptions, Coded Categories, Frequencies, and Percentages for All Study Drivers Combined, 2001-2005

Variable description	Categories	*n*	%
License/registration related			
DRIVER RESIDENCE STATE/COUNTRY IN STUDY CRASH	**0** = NORTH CAROLINA	20,511	94.1
	1 = NEIGHBORING STATE	743	3.4
	2 = OTHER	532	2.4
	9 = MISSING	0	0.0
DRIVER LICENSE ISSUANCE STATE/COUNTRY IN STUDY CRASH	**0** = NORTH CAROLINA	20,545	94.3
	1 = NEIGHBORING STATE	608	2.8
	2 = OTHER	631	2.9
	9 = MISSING	2	0.0
VEHICLE REGISTRATION STATE/COUNTRY IN STUDY CRASH	**0** = NORTH CAROLINA	19,416	89.1
	1 = NEIGHBORING STATE	782	3.6
	2 = OTHER	687	3.1
	9 = MISSING	901	4.1
DRIVER LICENSE & REGISTRATION STATE/COUNTRY COMBINED IN STUDY CRASH	**0** = NC LICENSE, NC REGISTRATION	19,148	87.9
	1 = NC LICENSE, OUT-OF-STATE REGISTRATION	597	2.7
	2 = OUT-OF-STATE LICENSE AND REGISTRATION	872	4.0
	3 = OUT-OF-STATE LICENSE, NC REGISTRATION	266	1.2
	9 = MISSING	903	4.1
DRIVER LICENSE & RESIDENCE STATE/COUNTRY COMBINED IN STUDY CRASH	**0** = LICENSED IN RESIDENCE STATE/ COUNTRY	21,057	96.6
	1 = NOT LICENSED IN RESIDENCE STATE/COUNTRY	729	3.3
	9 = MISSING	0	0.0
DRIVER COMMERCIAL DRIVER LICENSE (CDL) STATUS IN STUDY CRASH	**0** = NO CDL OR UNKOWN	21,758	99.9
	1 = CDL LICENSE	28	0.1
	9 = MISSING	0	0.0
Driver-related			
DRIVER AGE GROUP IN STUDY CRASH	**0** = 16-19	3,105	14.2
	1 = 20-29	6,042	27.7
	2 = 30-39	4,279	19.6
	3 = 40-49	3,415	15.7
	4 = 50-59	2,293	10.5
	5 = 60-69	1,342	6.2
	6 = 70-79	876	4.0
	7 = 80+	434	2.0
	9 = MISSING	0	0.0
DRIVER SEX	**0** = FEMALE	7,942	36.4
	1 = MALE	13,840	63.5
	9 = MISSING	4	0.1
DRIVER RACE	**0** = WHITE	15,401	70.7
	1 = BLACK	4,010	18.4

Variable description	Categories	*n*	%
	2 = NATIVE AMERICAN	261	1.2
	3 = HISPANIC	1,735	8.0
	4 = ASIAN/OTHER	325	1.5
	9 = MISSING	54	0.2
DRIVER BELT USE IN STUDY CRASH	**0** = BELTED, ANY TYPE	15,774	72.4
	1 = NO BELT	4,606	21.1
	9 = MISSING	1,406	6.4
DRIVER PHYSICAL CONDITION IN STUDY CRASH	**0** = NORMAL	15,676	71.9
	1 = IMPAIRED (ALCOHOL, DRUGS , MEDICATION)	3,188	14.6
	2 = SLEEPY/FATIGUED	574	2.6
	3 = OTHER IMPAIRMENT	311	1.4
	9 = MISSING	2,037	9.3
DRIVER ALCOHOL/ DRUG USE IN STUDY CRASH	**0** = NO ALCOHOL OR DRUGS	16,931	77.7
	1 = ALCOHOL ONLY SUSPECTED/DETECTED	3,332	15.3
	2 = DRUGS ONLY SUSPECTED/DETECTED	166	0.8
	3 = BOTH ALCOHOL/DRUGS SUSPECTED/DETECTED	295	1.3
	9 = MISSING	1,062	4.9
Vehicle related			
MODEL YEAR OF VEHICLE IN STUDY CRASH	**0** = 1995 OR NEWER MODEL YEAR	11,921	54.7
	1 = OLDER THAN 1995 MODEL YEAR	9,680	44.4
	9 = MISSING	185	0.8
VEHICLE TYPE IN STUDY CRASH	**0** = CAR	12,695	58.3
	1 = PICKUP	4,479	20.6
	2 = SPORT UTILITY VEHICLE	3,011	13.8
	3 = VAN/MINI VAN	1,601	7.3
	9 = MISSING	0	0.0
VEHICLE DEFECTS PRESENT IN STUDY CRASH	**0** = NONE	14,224	65.3
	1 = ONE OR MORE DEFECTS	611	2.8
	9 = MISSING	6,951	31.9
VEHICLE INSURED IN STUDY CRASH	**0** = INSURED VEHICLE	21,308	97.8
	1 = NOT INSURED	475	2.2
	9 = MISSING	3	0.0
NUMBER OF PASSENGERS IN STUDY CRASH	**0** = NONE	12,941	59.4
	1 = ONE	5,364	24.6
	2 = TWO+	3,481	16.0
	9 = MISSING	0	0.0
Crash Environment related			
SEASON OF YEAR OF STUDY CRASH	**0** = WINTER	5,135	23.6
	1 = SPRING	5,683	26.1
	2 = SUMMER	5,488	25.2
	3 = FALL	5,480	25.1
	9 = MISSING	0	0.0
DAY OF WEEK OF STUDY CRASH	**0** = WEEK DAY (MON-THU)	11,363	52.2
	1 = WEEKEND (FRI-SUN)	10,423	47.8
	9 = MISSING	0	0.0
HOUR (TIME) OF STUDY	**0** = 5:00-8:59 AM	3,067	14.1

Variable description	Categories	*n*	%
CRASH	**1** = 9:00 AM-12:59AM	3,696	17.0
	2 = 1:00-4:59 PM	5,450	25.0
	3 = 5:00-8:59 PM	4,792	22.0
	4 = 9:00 PM-12:59PM	3,009	13.8
	5 = 1:00-4:59 AM	1,772	8.1
	9 = MISSING	0	0.0
ROAD CLASS OF STUDY CRASH	**0** = LOCAL ROAD	4,767	21.9
	1 = NC/US/STATE ROAD	8,060	37.0
	2 = INTERSTATE	1,195	5.5
	3 = RURAL ROAD	7,764	35.6
	9 = MISSING	0	0.0
LOCALITY TYPE OF STUDY CRASH LOCATION	**0** = URBAN	5,734	26.3
	1 = MIXED	4,017	18.4
	2 = RURAL	12,035	55.2
	9 = MISSING	0	0.0
DEVELOPMENT TYPE IN LOCATION OF STUDY CRASH	**0** = RESIDENTIAL	5,206	23.9
	1 = COMMERCIAL/ INDUSTRIAL	5,214	23.9
	2 = FARM/UNDEVELOPED	11,366	52.2
	9 = MISSING	0	0.0
SPEED LIMIT OF STUDY CRASH ROAD	**0** = < 55 MPH	9,707	44.6
	1 = 55 MPH+	11,823	54.3
	9 = MISSING	256	1.2

The numbers of observations with missing values for particular crash database predictors was low (<5%) for most variables. The exceptions were driver belt use (missing for 6%), driver physical condition (missing for 9%), and vehicle defects (missing for 32%). Given that this study used a complete case analytic strategy (not imputing missing values), these variables were less desirable as predictors of culpability. From a practical standpoint, their use would exclude all observations with missing values, which would be almost 1/3 of the observations if the vehicle defects variable were used.

Some characteristics of drivers involved in fatal and A-level injury crashes include:

- Slightly more than 3% of drivers were reported to reside in an adjacent (abutting) state to NC; somewhat less than 3% of drivers were licensed in an adjacent state.
- Slightly more than 2% of drivers were reported to reside in a more distant state (non-adjacent) or country; nearly 3% of drivers were licensed in a non-adjacent state or country.
- When driver license state and vehicle registration state were compared, about 4% were different (licensed in NC, but registered out-of-state or vice-versa).
- When residence state and license state were compared, about 3% were different.
- The youngest age group, spanning only 4 years, accounted for 14% of drivers.

- Drivers aged 20 – 29 represented the largest percentage involvement by age group (28%).
- 30 – 39 year olds accounted for (20%) of drivers.
- Males accounted for 64%; females 36%.

Among obvious risk factors,

- 21% of drivers were indicated to be unbelted.
- Impairments due to alcohol, drugs, or medication were indicated for 15%,
- Sleepy/fatigued, 3%, and
- Other impairments (medical/physical) less than 2%.
- Alcohol/drugs were similarly suspected/detected in 15% of cases.
- Twenty-two percent of drivers were involved in fatal and A-injury crashes during the hours from 9 pm to 5 am.

Other driver, vehicle and environmental characteristics of the crashes can be obtained from Table 1.

Driver physical condition (for non-missing values) in the study crash was distributed by age group as shown in Table 2. The 20 to 29 year age group was most over-represented in the impaired physical condition category (24% of drivers in this age group compared with 16% over all ages). The 70 to 79 year age group was most over-represented in the sleepy/fatigued category, followed by 16 to 19 year olds. Other impairments (medical, etc.) were increasingly associated with age.

Table 2. Driver physical condition at time of study crash by age group of the driver (NC drivers only)

Driver age	DRIVER PHYSICAL CONDITION IN STUDY CRASH: NORMAL	IMPAIRED (A,D,M)	SLEEPY/ FATIGUED	OTHER IMPAIRMENT	Total
16-19	2,383 84.2%[1] 15.2%[2]	323 11.4% 10.1%	109 3.9% 19.0%	15 0.5% 4.8%	2,830 100.0% 14.3%
20-29	3,942 72.3% 25.1%	1,299 23.8% 40.7%	157 2.9% 27.4%	54 1.0% 17.4%	5,452 100.0% 27.6%
30-39	3,011 77.3% 19.2%	757 19.4% 23.7%	86 2.2% 15.0%	43 1.1% 13.8%	3,897 100.0% 19.7%
40-49	2,425 78.1% 15.5%	546 17.6% 17.1%	80 2.6% 13.9%	55 1.8% 17.7%	3,106 100.0% 15.7%
50-59	1,801 85.4% 11.5%	204 9.7% 6.4%	60 2.8% 10.5%	43 2.0% 13.8%	2,108 100.0% 10.7%
60-69	1,090 89.9% 7.0%	42 3.5% 1.3%	40 3.3% 7.0%	40 3.3% 12.9%	1,212 100.0% 6.1%
70-79	695 89.0% 4.4%	16 2.0% 0.5%	33 4.2% 5.7%	37 4.7% 11.9%	781 100.0% 4.0%
80+	329 90.6% 2.1%	1 0.3% 0.0%	9 2.5% 1.6%	24 6.6% 7.7%	363 100.0% 1.8%
All Ages	15,676 79.4% 100.0%	3,188 16.1% 100.0%	574 2.9% 100.0%	311 1.6% 100.0%	19,749[3] 100.0% 100.0%

[1]Percent of age group
[2]Percent of impairment group
[3]Missing cases excluded

Younger drivers up to the 30 to 39 year age group were over-represented among the unbelted drivers, compared to all ages (Table 3).

Table 3. Driver seat belt use by age group of driver

Age group	DRIVER BELT USE IN STUDY CRASH		
	BELTED, ANY TYPE	NO BELT	Total
16-19	2,226 76.4%[1] 14.1%[2]	687 23.6% 14.9%	2,913 100.0% 14.3%
20-29	4,030 72.4% 25.5%	1,539 27.6% 33.4%	5,569 100.0% 27.3%
30-39	3,010 75.7% 19.1%	966 24.3% 21.0%	3,976 100.0% 19.5%
40-49	2,497 77.8% 15.8%	714 22.2% 15.5%	3,211 100.0% 15.8%
50-59	1,794 82.5% 11.4%	380 17.5% 8.3%	2,174 100.0% 10.7%
60-69	1,120 86.4% 7.1%	176 13.6% 3.8%	1,296 100.0% 6.4%
70-79	743 89.2% 4.7%	90 10.8% 2.0%	833 100.0% 4.1%
80+	354 86.8% 2.2%	54 13.2% 1.2%	408 100.0% 2.0%
All Ages	15,774 77.4% 100.0%	4,606 22.6% 100.0%	20,380[3] 100.0% 100.0%

[1]Percent of age group
[2]Percent of belt use group
[3]Missing cases excluded

Males were more than three times as likely to be impaired by alcohol, drugs, or medication (22%) than females (7%) (data not shown). Males were also twice as likely to be unbelted (28%) during the study crash as were females (14%).

Licensing database predictors

The next set of potential predictor variables all came from the licensing history database of the North Carolina DMV (NCDMV). The driver license numbers on the crash record were matched to the official database and up to five years of pre-study crash licensing history was obtained for each driver. Although it was possible to track the ‘licensing’ and conviction/action history of unlicensed North Carolina residents, there is no mechanism for accurately doing so for out-of-state drivers or (obviously) those that did not match a person in the licensing database. Therefore, licensing history data were not available for out-of-state drivers and those for whom licensing status could not be determined. This latter group may have included miscoded driver license numbers, instances where drivers provided false identification, or where there was missing driver license information on the crash record.

The approach to identifying and coding the licensing history variables was to include factors found to be predictive in past studies (e.g., past crash and conviction histories), as well as include more specific information about related factors that might add additional predictive power when available (e.g., types of prior license suspensions rather than just the overall number of past licensing suspensions). The set of 42 licensing database predictor variables is shown in Table 4 with the overall distributions of various driver factors.

Table 4. North Carolina Driver History Database Predictor Variables with Descriptions, Coded Categories, Frequencies, and Percentages for All Study Drivers Combined, 2001-2005

Variable description	Categories	*n*	%
LICENSE STATUS AT STUDY CRASH	**0** = FULLY NC LICENSED	16,180	74.3
	1 = SUSPENDED/LIMITED PRIVILEGE NC LICENSE	2,252	10.3
	2 = NC UNLICENSED	1,823	8.4
	3 = OUT-OF-STATE LICENSE	1,236	5.7
	4 = NOT DETERMINABLE	295	1.3
Restriction history			
OVERALL NUMBER OF LICENSE RESTRICTIONS IN PRIOR 5 YEARS	**0** = NONE	13,325	61.6
	1 = ONE	6,451	29.6
	2 = TWO+	479	2.2
	9 = MISSING	1,531	7.0
CORRECTIVE LENSES RESTRICTION IN PRIOR 5 YEARS	**0** = NO	14,740	67.7
	1 = YES	5,515	25.3
	9 = MISSING	1,531	7.0
45MPH/NO INTERSTATE RESTRICTION IN PRIOR 5 YEARS	**0** = NO	20,195	92.7
	1 = YES	60	0.3
	9 = MISSING	1,531	7.0
DAYLIGHT DRIVING ONLY RESTRICTION IN PRIOR 5 YEARS	**0** = NO	20,197	92.7
	1 = YES	58	0.3
	9 = MISSING	1,531	7.0
INTRASTATE DRIVING ONLY RESTRICTION IN PRIOR 5 YEARS	**0** = NO	20,192	92.7
	1 = YES	63	0.3
	9 = MISSING	1,531	7.0
OTHER LICENSE RESTRICTION IN PRIOR 5 YEARS	**0** = NO	19,428	89.2
	1 = YES	827	3.8
	9 = MISSING	1,531	7.0
ONLY DRIVE WHILE SUPERVISED RESTRICTION IN PRIOR 5 YEARS	**0** = NO	20,111	92.3
	1 = YES	144	0.7
	9 = MISSING	1,531	7.0
AUTOMATIC TRANSMISSION ONLY RESTRICTION IN	**0** = NO	20,241	92.9
	1 = YES	14	0.1

Variable description	Categories	*n*	%
PRIOR 5 YEARS	**9** = MISSING	1,531	7.0
DRIVE CLASS B AND C ONLY RESTRICTION IN PRIOR 5 YEARS	**0** = NO	20,228	92.8
	1 = YES	27	0.1
	9 = MISSING	1,531	7.0
GRADUATED DRIVER LICENSING RESTRICTION IN PRIOR 5 YEARS	**0** = NO	19,658	90.2
	1 = YES	597	2.7
	9 = MISSING	1,531	7.0
BLOOD ALCOHOL CONTENT < .04 RESTRICTION IN PRIOR 5 YEARS	**0** = NO	20,133	92.4
	1 = YES	122	0.6
	9 = MISSING	1,531	7.0
BLOOD ALCOHOL CONTENT = .00 RESTRICTION IN PRIOR 5 YEARS	**0** = NO	20,241	92.9
	1 = YES	14	0.1
	9 = MISSING	1,531	7.0
Crash history			
*TOTAL PRIOR CRASHES IN PRIOR 5 YEARS	**0** = NONE	13,140	60.3
	1 = ONE	4,828	22.2
	2 = TWO	1,551	7.1
	3 = THREE	467	2.1
	4 = FOUR+	269	1.2
	9 = MISSING	1,531	7.0
NUMBER OF PRIOR NOT-AT-FAULT CRASHES IN PRIOR 5 YEARS	**0** = NONE	19,265	88.4
	1 = ONE	2243	10.3
	2 = TWO	241	1.1
	3 = THREE+	37	0.2
NUMBER OF PRIOR AT-FAULT CRASHES IN PRIOR 5 YEARS	**0** = NONE	16,758	76.9
	1 = ONE	2,865	13.1
	2 = TWO	495	2.3
	3 = THREE+	137	0.6
	9 = MISSING	1,531	7.0
AT-FAULT IN MOST RECENT PRIOR CRASH INVOLVEMENT	**0** = NO/NO CRASH	17,317	79.5
	1 = YES	2,938	13.5
	9 = MISSING	1,531	7.0
AT-FAULT 2ND MOST RECENT PRIOR CRASH INVOLVEMENT	**0** = NO/NO CRASH	19,345	88.8
	1 = YES	910	4.2
	9 = MISSING	1,531	7.0
AT-FAULT 3RD MOST RECENT CRASH INVOLVEMENT WITHIN 5 YEARS PRIOR	**0** = NO/NO CRASH	19,951	91.6
	1 = YES	304	1.4
	9 = MISSING	1,531	7.0
AGE DIFFERENCE AT MOST RECENT PRIOR CRASH INVOLVEMENT	**0** = SAME AGE	1,256	5.8
	1 = ONE YEAR YOUNGER	2,029	9.3
	2 = TWO YEARS YOUNGER	1,441	6.6
	3 = THREE YEARS YOUNGER	1,146	5.3
	4 = FOUR YEARS YOUNGER	864	4.0
	5 = FIVE YEARS YOUNGER	377	1.7
	6 = NO CRASH PRIOR 5 YEARS	13,140	60.3
	9 = MISSING	1,533	7.0

Variable description	Categories	*n*	%
Conviction history			
OVERALL NUMBER OF CONVICTIONS IN PRIOR 5 YEARS	**0** = NONE	12,313	56.5
	1 = ONE	3,559	16.3
	2 = TWO	1,851	8.5
	3 = THREE	1,051	4.8
	4 = FOUR	583	2.7
	5 = FIVE+	896	4.1
	9 = MISSING	1,533	7.0
NUMBER OF LICENSE DEMERIT POINTS IN PRIOR 5 YEARS	**0** = NONE	14,104	64.7
	1 = 1-2 POINTS	1,499	6.9
	2 = 3-4 POINTS	2,595	11.9
	3 = 5-6 POINTS	1,019	4.7
	4 = 7-8 POINTS	436	2.0
	5 = 9+ POINTS	600	2.7
	9 = MISSING	1,533	7.0
NUMBER OF NON-MOVING CONVICTIONS IN PRIOR 5 YEARS	**0** = NO CONVICTIONS	19,628	90.1
	1 = ONE	527	2.4
	2 = TWO+	98	0.4
	9 = MISSING	1,533	7.0
NUMBER OF LICENSE-RELATED CONVICTIONS IN PRIOR 5 YEARS	**0** = NO CONVICTIONS	18,345	84.2
	1 = ONE	1,245	5.7
	2 = TWO+	663	3.0
	9 = MISSING	1,533	7.0
NUMBER OF SPEED-RELATED CONVICTIONS IN PRIOR 5 YEARS	**0** = NO CONVICTIONS	15,009	68.9
	1 = ONE	3,365	15.4
	2 = TWO+	1,879	8.6
	9 = MISSING	1,533	7.0
NUMBER OF ALCOHOL-RELATED CONVICTIONS IN PRIOR 5 YEARS	**0** = NO CONVICTIONS	18,581	85.3
	1 = ONE	441	2.0
	2 = TWO+	1,231	5.6
	9 = MISSING	1,533	7.0
NUMBER OF OTHER MOVING CONVICTIONS IN PRIOR 5 YEARS	**0** = NO CONVICTIONS	18,559	85.2
	1 = ONE	1,460	6.7
	2 = TWO+	234	1.1
	9 = MISSING	1,533	7.0
NUMBER OF OTHER SERIOUS CONVICTIONS IN PRIOR 5 YEARS	**0** = NO CONVICTIONS	19,694	90.4
	1 = ONE	482	2.2
	2 = TWO+	77	0.3
	9 = MISSING	1,533	7.0
Suspensions			
OVERALL NUMBER OF LICENSE SUSPENSIONS IN PRIOR 5 YEARS	**0** = NONE	15,543	71.3
	1 = ONE	2,962	13.6
	2 = TWO	1,147	5.3
	3 = THREE	528	2.4
	4 = FOUR	75	0.3
	9 = MISSING	1,531	7.0
ANY LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	15,543	71.3
	1 = YES	4,712	21.6
	9 = MISSING	1,531	7.0
ADMINISTRATIVE LICENSE SUSPENSION IN PRIOR 5	**0** = NO	20,222	92.8
	1 = YES	33	0.1

Variable description	Categories	*n*	%
YEARS	**9** = MISSING	1,531	7.0
ALCOHOL-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	18,005	82.6
	1 = YES	2,250	10.3
	9 = MISSING	1,531	7.0
DRIVING WHILE SUSPENDED LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	19,145	87.9
	1 = YES	1,110	5.1
	9 = MISSING	1,531	7.0
FATAL INVOLVEMENT LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	20,250	92.9
	1 = YES	5	0.0
	9 = MISSING	1,531	7.0
FAILURE TO APPEAR/ FAILURE TO PAY LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	17,460	80.1
	1 = YES	2,795	12.8
	9 = MISSING	1,531	7.0
NOT TRAFFIC RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	19,974	91.7
	1 = YES	281	1.3
	9 = MISSING	1,531	7.0
OTHER LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	20,181	92.6
	1 = YES	74	0.3
	9 = MISSING	1,531	7.0
PHYSICAL OR MENTAL CONDITION LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	20,218	92.8
	1 = YES	37	0.2
	9 = MISSING	1,531	7.0
POINTS-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	20,127	92.4
	1 = YES	128	0.6
	9 = MISSING	1,531	7.0
GRADUATED DRIVER LICENSING-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	20,145	92.5
	1 = YES	110	0.5
	9 = MISSING	1,531	7.0
RECKLESS DRIVING LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	20,212	92.8
	1 = YES	43	0.2
	9 = MISSING	1,531	7.0
SERIOUS OFFENSE LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	20,209	92.8
	1 = YES	46	0.2
	9 = MISSING	1,531	7.0
SPEED-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS	**0** = NO	20,021	91.9
	1 = YES	234	1.1
	9 = MISSING	1,531	7.0

Note. Included in the missing values are 1,531 out-of-state and unmatched drivers (7% of total).
*Total prior crashes provided for information only and not used as a predictor variable in models since prior at-fault and not-at-fault sum to total crashes for an individual driver.

The first variable shown in the table describes the licensing status classification results from matching the study driver license numbers to those in the licensing history database. Over 93% of the study drivers were successfully matched to the license or driver history database (NCDMV).

- Approximately 19% of the drivers were North Carolina residents were either unlicensed (8%) or driving on a suspended/limited privilege license (10%) at the time of the crash. (Among the suspended license drivers, above 300 were suspended for virtually the entire prior five years while more than 2000 were suspended for a year or more.)

- About 7% of the study drivers had either out-of-state licenses or it was not possible to determine their licensing status because of non-matching or missing driver license information.

The out-of-state and not determinable drivers (n = 1,531) represent the majority of missing values for the licensing database variables shown in the table. Excluding these observations, no licensing-related variable was missing for 5% or more of the drivers, making them all good practical candidates as predictors of culpability for state residents. Although they may represent a risk group, the out-of-state and not determinable license status drivers are by necessity excluded from further analyses involving the licensing database variables, though they will be included in the analyses involving only crash-based predictors.

In general, sample sizes of drivers with license restrictions were less than one percent. The exceptions were:

- Those with corrective lens restriction (25% of drivers)

- Graduated driver license (GDL) restriction (level 1 or 2, 2.7%)

- Combined “other” restrictions (4%)

Regarding prior crash histories:

- 33% of all drivers had one or more prior crashes within five years prior to their most recent crash. 1.2% or 269 drivers had four or more crashes in the five years preceding their last crash.

- More drivers had one or more prior at-fault crashes (16%) than not-at-fault crashes (12%)

Other license history results for the sample of involved NC drivers included:

- Above 36% of drivers had one or more prior convictions

- 24% had speeding-related convictions

- 8% had alcohol-related convictions

- 28% had demerit points

- Nearly 22% had one or more driver license suspensions within the prior five years; 8% had two or more license suspensions in the prior five years.

Numbers of drivers with some types of suspensions were fairly small (1% or less of drivers), including serious offenses, reckless, administrative suspension, points-related suspension, GDL, physical or mental condition, and fatal-involvement suspension. The most common types of license suspensions were failure to appear/failure to pay (13%), alcohol-related (10%), and 5% were suspended for driving while suspended. Non-traffic related suspensions accounted for a little over 1% of drivers, as did speed-related suspensions. (Suspension groupings for the types of suspensions that appeared on this set of drivers' records are included in Appendix 1.)

Drivers with more than one prior conviction (any type) were increasingly over-represented among those who had multiple prior (at-fault) crashes compared to drivers with none or one prior conviction as shown in Table 5.

Table 5. Prior At-fault crashes by prior convictions

OVERALL NUMBER OF CONVICTIONS IN PRIOR 5 YEARS	NUMBER OF AT-FAULT CRASHES IN PRIOR 5 YEARS				
	NONE	ONE	TWO	THREE+	Total
NONE	11,102 66.3%[1]	1,084 37.8%	106 21.4%	21 15.3%	12,313 60.8%[2]
ONE	2,759 16.5%	671 23.4%	102 20.6%	27 19.7%	3,559 17.6%
TWO	1,308 7.8%	422 14.7%	92 18.6%	29 21.2%	1,851 9.1%
THREE	705 4.2%	269 9.4%	60 12.1%	17 12.4%	1,051 5.2%
FOUR	368 2.2%	160 5.6%	43 8.7%	12 8.8%	583 2.9%
FIVE+	514 3.1%	259 9.0%	92 18.6%	31 22.6%	896 4.4%
Total	16,756 82.7%[2]	2,865 14.1%	495 2.4%	137 0.7%	20,253 100.0%

[1] Percent of column total
[2] Percent of table total

In contrast, Table 6 shows the prior not-at-fault crashes by number of convictions. There is not an obvious trend between those having increasing numbers of prior convictions and prior not-at-fault crash involvements (except a possible trend for those with five or more convictions to be over-represented among those with 2 or more "no-fault" crashes in the five years.

Table 6. Prior Not-at-fault crashes by prior convictions.

OVERALL NUMBER OF CONVICTIONS IN PRIOR 5 YEARS	NO FAULT CRASHES in 5-YEARS PRIOR				Total
	0	1	2	3+	
NONE	1,944 42.3%[1]	1,116 49.8%	87 36.1%	15 40.5%	3,162 44.4%[2]
ONE	1,000 21.8%	471 21.0%	60 24.9%	4 10.8%	1,535 21.6%
TWO	632 13.8%	259 11.6%	32 13.3%	9 24.3%	932 13.1%
THREE	381 8.3%	163 7.3%	24 10.0%	2 5.4%	570 8.0%
FOUR	237 5.2%	89 4.0%	11 4.6%	1 2.7%	338 4.8%
FIVE+	401 8.7%	144 6.4%	27 11.2%	6 16.2%	578 8.1%
Total	4,595 64.6%[2]	2,242 31.5%	241 3.4%	37 0.5%	7,115 100.0%

[1] Percent of column total (crashes)
[2] Percent of table total

Group Formation

The fatal and A-injury crash-involved drivers were grouped based on combinations of the number of vehicles involved in the crash (i.e., multiple vs. single) and whether or not the driver was culpable for the crash (i.e., not-at-fault vs. at-fault) to form the following three groups: (a) multiple-vehicle crash, driver not at-fault (n = 7,595); (b) multiple-vehicle crash, driver at-fault (n = 6,571); and (c) single-vehicle crash, driver at-fault (n = 7,620). The first group, those with no contributing circumstances cited (multiple-vehicle, not-at-fault drivers) served as the induced exposure comparison group used to represent the background exposure distribution of drivers on North Carolina roadways. The two at-fault groups of drivers were compared to these drivers to identify driver, vehicle, environment, and licensing history variables that were over-represented in at-fault parties.

For descriptive purposes, Table 7 shows the vehicle maneuvers prior to the crash for each group of drivers. The most common vehicle maneuver prior to the crash for all three groups of drivers was simply driving straight ahead (ranging from 66% for multi-vehicle, at fault to 95% for single-vehicle drivers). The next most common maneuvers for not-at-fault multi-vehicle crash drivers were being stopped in the travel lane (11%), making a left turn (3%), and slowing or stopping (3%). Those for the multi-vehicle at-fault drivers were making a left turn (21%) and starting in the roadway (3%), passing (2%), and changing lanes or merging (2%). The only other maneuver that stands out for the single-vehicle at-fault drivers is passing (1%).

Table 7. Vehicle Maneuver Prior to Study Crash by Driver Fault x Number of Vehicles Involved Cross-Classification, 2001-2005.

	NAF		MVAF		SVAF	
Vehicle maneuver	***n***	**%**	***n***	**%**	***n***	**%**
Stopped in travel lane	862	11.35	36	0.55	14	0.18
Parked out of travel lanes	24	0.32	3	0.05	4	0.05
Parked in travel lanes	6	0.08	3	0.05	2	0.03
Going straight ahead	6,123	80.62	4,371	66.52	7,279	95.52
Changing lanes or merging	14	0.18	124	1.89	63	0.83
Passing	23	0.30	152	2.31	89	1.17
Making right turn	55	0.72	66	1.00	21	0.28
Making left turn	230	3.03	1,387	21.11	42	0.55
Making U turn	4	0.05	44	0.67	4	0.05
Backing	2	0.03	14	0.21	2	0.03
Slowing or stopping	215	2.83	69	1.05	9	0.12
Starting in roadway	17	0.22	205	3.12	3	0.04
Parking	1	0.01	2	0.03	0	0.00
Leaving parked position	2	0.03	6	0.09	3	0.04
Avoiding object in road	4	0.05	12	0.18	29	0.38
Other	13	0.17	77	1.17	56	0.73
Total	7,595	100.00	6,571	100.00	7,620	100.00

Note. NAF = multi-vehicle crash not at-fault drivers; MVAF = multi-vehicle crash at-fault drivers; SVAF = single vehicle at-fault drivers. Percentages represent column percentages within each group of drivers.

The driver contributing circumstances for the two groups of at-fault drivers are summarized in Table 8 for descriptive purposes. These are the contributing circumstances noted by the law enforcement officers that were used to determine that the driver was at fault for the crash in this study. Up to three different contributing circumstances can be coded for each driver; hence the numbers in the table do not sum to the total sample sizes of at-fault drivers, nor do the column percentages add to 100. The table shows frequencies and percentages of drivers that were cited for each contributing circumstance in any of the three fields. For the multi-vehicle at-fault drivers, the most common contributing circumstances were failure to yield the right-of-way (32%), crossing the centerline/driving the wrong way (29%), and inattention (18%). The most common contributing circumstances for the single-vehicle at-fault drivers were exceeding safe speed for conditions (33%), erratic, reckless, careless, negligent, or aggressive driving (33%), exceeding the authorized speed limit (31%), and alcohol use (31%). Overall, 49.3% of the multi-vehicle at-fault and 61.9% of the single-vehicle at-fault drivers had two or more contributing circumstances indicated by the law enforcement officer.

Table 8. Contributing Circumstances in Study Crash for each Group of At-Fault Drivers

	MVAF		SVAF	
Driver Contributing Circumstance	*n*	%	*n*	%
Disregarded yield sign	17	0.26	3	0.04
Disregarded stop sign	567	8.63	141	1.85
Disregarded other traffic signs	32	0.49	10	0.13
Disregarded traffic signals	456	6.94	8	0.10
Disregarded road markings	51	0.78	27	0.35
Exceeded authorized speed limit	535	8.14	2,405	31.56
Exceeded safe speed for conditions	559	8.51	2,547	33.43
Failure to reduce speed	899	13.68	348	4.57
Improper turn	245	3.73	6	0.08
Right turn on red	3	0.05	1	0.01
Crossed centerline/going wrong way	1,903	28.96	1,509	19.80
Improper lane change	83	1.26	11	0.14
Use of improper lane	41	0.62	18	0.24
Overcorrected/oversteered	178	2.71	1,126	14.78
Passed stopped school bus	1	0.02	2	0.03
Passed on hill	10	0.15	1	0.01
Passed on curve	18	0.27	15	0.20
Other improper passing	89	1.35	26	0.34
Failed to yield right of way	2,095	31.88	3	0.04
Inattention	1,212	18.44	782	10.26
Improper backing	11	0.17	5	0.07
Improper parking	3	0.05	1	0.01
Driver distracted	50	0.76	66	0.87
Improper or no signal	3	0.05	0	0.00
Followed to closely	73	1.11	7	0.09
Erratic, reckless, careless, negligent, aggressive	760	11.57	2,498	32.78
Wind, slippery surface, vehicle object, non-motorist	37	0.56	124	1.63
Visibility obstructed	52	0.79	12	0.16
Operated defective equipment	87	1.32	113	1.48
Alcohol use	704	10.71	2,348	30.81
Drug use	77	1.17	134	1.76
Other	147	2.24	354	4.65
Unable to determine	7	0.11	11	0.14
Unknown	10	0.15	17	0.22

Note. MVAF = multi-vehicle crash at-fault drivers; SVAF = single vehicle at-fault drivers. Percentages represent the percentages within each group of at-fault drivers with each contributing circumstance. Each driver could have multiple (up to three) different contributing circumstances, so the summed column percentages exceed 100%.

Univariate Relationships of Predictor Variables to Multi- and Single-Vehicle Culpability

The next sections and tables show the distribution of driver factors from the crash and driver history databases across the three fault groups.

Because the out-of-state and unmatched drivers are excluded from the analyses involving the driver history database predictors, the univariate analyses are presented separately for these two sets of predictors. The crash database variable frequencies include all the fatal and serious injury crash-involved drivers, whereas the driver history database frequencies include only in-state residents that could be matched to the driver history database.

Crash database predictors

The set of 24 potential crash database predictor variables was chosen using a series of cross-tabulations in which the available crash record variable distributions were compared across the three groups of crash-involved drivers and those which differed in a univariate sense across the groups were retained as predictors. Chi square tests of independence were used for the purposes of testing their potential for discriminating between multiple-vehicle not-at-fault drivers and the two groups of at-fault drivers (multiple-vehicle and single-vehicle drivers). These analyses are useful for seeing which variables differ among the groups, without adjusting for the effects of other variables. Continuous variables, such as model year of vehicle and driver ages, were re-coded into categories for easier interpretation and to improve the likelihood of significant findings (larger samples per group). Variables were, in some cases, recoded into categories that were deemed to best capture the differences among the driver groups based on the results of the chi square tests and using adjusted standardized residuals for the individual cells. The percentages of the three groups of drivers in each category of each predictor variable, as well as the results of the omnibus chi square tests, are presented in Table 9. The column percentages are based on all observations with non-missing values for each variable, out of all 7,595 multi-vehicle crash not-at-fault (NAF) drivers, 6,571 multi-vehicle crash at-fault drivers (MVAF), and 7,620 single-vehicle crash at-fault drivers (SVAF). Both standardized residuals for the cells and additional chi square follow-up tests were used to determine where the differences existed when a significant ($p < .05$) or trending ($p < .10$) omnibus chi square test was observed.

Table 9. Percentages of Multiple-Vehicle Not-At-Fault, Multiple-Vehicle At-Fault, and Single-Vehicle At-Fault Fatal and Serious-Injury Crash-Involved Drivers in each Level of the Crash Database Predictor Variables and Associated Omnibus Chi Square Results

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
License/Registration-Related Variables				
DRIVER RESIDENCE STATE/COUNTRY IN STUDY CRASH				18.40*
NORTH CAROLINA	94.4	94.2	93.8	
NEIGHBORING STATE	3.6	3.5	3.2	
OTHER	2.0	2.3	3.0	
MISSING	0.0	0.0	0.0	

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
DRIVER LICENSE ISSUANCE STATE/COUNTRY IN STUDY CRASH				15.10*
NORTH CAROLINA	94.8	94.0	94.0	
NEIGHBORING STATE	2.8	2.9	2.7	
OTHER	2.3	3.1	3.3	
MISSING	0.0	0.0	0.0	
VEHICLE REGISTRATION STATE/COUNTRY IN STUDY CRASH				15.91*
NORTH CAROLINA	89.6	89.4	88.4	
NEIGHBORING STATE	3.4	3.8	3.6	
OTHER	2.7	3.1	3.6	
MISSING	4.2	3.7	4.4	
DRIVER LICENSE & REGISTRATION STATE/COUNTRY COMBINED IN STUDY CRASH				20.54*
NC LICENSE, NC REGISTRATION	88.6	88.0	87.0	
NC LICENSE, OUT-OF-STATE REGISTRATION	2.4	2.7	3.1	
OUT-OF-STATE LICENSE, OUT-OF-STATE REGISTRATION	3.8	4.2	4.1	
OUT-OF-STATE LICENSE, NC REGISTRATION	1.0	1.4	1.3	
MISSING	4.2	3.7	4.4	
DRIVER LICENSE & RESIDENCE STATE/COUNTRY COMBINED IN STUDY CRASH (				12.43*
LICENSED IN RESIDENCE STATE/ COUNTRY	97.2	96.4	96.3	
NOT LICENSED IN RESIDENCE STATE/COUNTRY	2.8	3.6	3.7	
MISSING	0.0	0.0	0.0	
DRIVER COMMERCIAL DRIVER LICENSE (CDL) STATUS IN STUDY CRASH				4.79†
NO CDL OR UNKOWN	99.9	99.8	99.9	
CDL LICENSE	0.1	0.2	0.1	
Driver-Related Variables				
DRIVER AGE GROUP IN STUDY CRASH				1557.65*
16-19	7.1	16.4	19.5	
20-29	22.8	27.1	33.2	
30-39	22.3	16.9	19.4	
40-49	19.6	13.2	13.9	
50-59	15.0	8.6	7.8	
60-69	8.3	7.0	3.3	
70-79	3.9	6.5	2.0	
80+	1.2	4.2	0.9	
MISSING	0.0	0.0	0.0	
DRIVER SEX				306.93*
FEMALE	42.5	38.1	29.1	
MALE	57.5	61.9	70.9	
MISSING	0.1	0.0	0.1	
DRIVER RACE				121.91*
WHITE	72.4	69.7	69.9	
BLACK	19.1	17.7	18.3	
NATIVE AMERICAN	1.0	1.0	1.6	
HISPANIC	5.7	9.4	8.9	
ASIAN/OTHER	1.4	2.0	1.1	
MISSING	0.4	0.2	0.2	
DRIVER BELT USE IN STUDY CRASH				3673.88*
BELTED, ANY TYPE	91.7	77.4	48.9	

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
NO BELT	5.7	16.0	41.0	
MISSING	2.6	6.6	10.1	
DRIVER PHYSICAL CONDITION IN STUDY CRASH				5272.58*
NORMAL	96.8	74.4	45.1	
IMPAIRED (ALCOHOL, DRUGS , MEDICATION)	0.4	11.3	31.7	
SLEEPY/FATIGUED	0.1	2.4	5.4	
OTHER IMPAIRMENT	0.1	1.8	2.4	
MISSING	2.7	10.1	15.4	
DRIVER ALCOHOL/DRUG USE IN STUDY CRASH				4273.47*
NO ALCOHOL OR DRUGS	97.7	81.5	54.5	
ALCOHOL ONLY SUSPECTED/DETECTED	1.1	11.2	32.9	
DRUGS ONLY SUSPECTED/DETECTED	0.1	1.2	1.1	
BOTH ALCOHOL/DRUGS SUSPECTED/DETECTED	0.1	1.4	2.6	
MISSING	1.1	4.6	8.9	
Vehicle-Related Variables				
MODEL YEAR OF VEHICLE IN STUDY CRASH				173.19*
1995 OR NEWER MODEL YEAR	60.7	51.6	51.4	
OLDER THAN 1995 MODEL YEAR	38.4	47.6	47.7	
MISSING	0.9	0.8	0.9	
VEHICLE TYPE IN STUDY CRASH				357.64*
CAR	52.5	64.6	58.6	
PICKUP	21.8	18.2	21.3	
SPORT UTILITY VEHICLE	15.2	10.6	15.2	
VAN/MINI VAN	10.5	6.6	4.8	
MISSING	0.0	0.0	0.0	
VEHICLE DEFECTS PRESENT IN STUDY CRASH				276.90*
NONE	67.0	63.8	64.9	
ONE OR MORE DEFECTS	0.5	3.0	4.9	
MISSING	32.5	33.2	30.2	
VEHICLE INSURED IN STUDY CRASH				145.42*
INSURED VEHICLE	98.9	98.4	96.2	
NOT INSURED	1.1	1.6	3.8	
MISSING	0.0	0.1	0.1	
NUMBER OF PASSENGERS IN STUDY CRASH				73.93*
NONE	56.3	62.8	59.5	
ONE	26.4	23.5	23.8	
TWO+	17.3	13.7	16.6	
MISSING	0.0	0.0	0.0	
Environment-Related Variables				
SEASON OF YEAR OF STUDY CRASH				6.02
WINTER	23.1	23.1	24.4	
SPRING	26.3	26.5	25.5	
SUMMER	25.2	25.1	25.3	
FALL	25.4	25.2	24.8	
MISSING	0.0	0.0	0.0	
DAY OF WEEK OF STUDY CRASH				158.47*
WEEK DAY (MON-THU)	55.4	55.1	46.3	
WEEKEND (FRI-SUN)	44.6	44.9	53.6	
MISSING	0.0	0.0	0.0	

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
HOUR (TIME) OF STUDY CRASH				2813.82*
5:00-8:59 AM	14.2	14.5	13.6	
9:00 AM-12:59AM	20.1	20.0	11.2	
1:00-4:59 PM	30.1	29.1	16.4	
5:00-8:59 PM	24.1	24.0	18.1	
9:00 PM-12:59PM	9.1	9.6	22.1	
1:00-4:59 AM	2.3	2.7	18.6	
MISSING	0.0	0.0	0.0	
ROAD CLASS OF STUDY CRASH				1941.66*
LOCAL ROAD	27.1	26.9	12.3	
NC/US/STATE ROAD	45.1	38.9	27.3	
INTERSTATE	4.7	4.1	7.4	
RURAL ROAD	23.1	30.1	52.9	
MISSING	0.0	0.0	0.0	
LOCALITY TYPE OF STUDY CRASH LOCATION				1359.73*
URBAN	33.7	32.6	13.5	
MIXED	20.2	20.3	15.0	
RURAL	46.1	47.1	71.4	
MISSING	0.0	0.0	0.0	
DEVELOPMENT TYPE IN LOCATION OF STUDY CRASH				1765.71*
RESIDENTIAL	23.0	24.2	24.4	
COMMERCIAL/ INDUSTRIAL	33.4	31.4	8.1	
FARM/UNDEVELOPED	43.5	44.4	67.5	
MISSING	0.0	0.0	0.0	
SPEED LIMIT OF STUDY CRASH ROAD				663.07*
< 55MPH	50.8	51.0	32.8	
55MPH+	48.2	47.8	65.9	
MISSING	1.0	1.2	1.3	

Note. NAF = multi-vehicle crash not at-fault drivers; MVAF = multi-vehicle crash at-fault drivers; SVAF = single vehicle at-fault drivers. Percentages represent column percentages within each group of drivers.
†$p < .10$. *$p < .05$.

In terms of licensing/registration variables from the crash database, drivers who resided or had their vehicle registered in a jurisdiction that was not North Carolina or a neighboring state were over-represented in single-vehicle at-fault crashes relative to multi-vehicle not-at-fault drivers. Furthermore, both single- and multi-vehicle at-fault drivers were somewhat less likely to be licensed in North Carolina and more likely to be licensed in some non-neighboring jurisdiction than were not-at-fault drivers. When the data for the drivers' licensing and registration jurisdictions were combined into a single variable, the single-vehicle at-fault drivers were found to be less likely to have both a NC license and NC vehicle registration, and more likely to have a NC license and out-of-state registration than were not-at-fault drivers. Furthermore, both groups of at-fault drivers were somewhat more likely to have an out-of-state license with a vehicle registered in North Carolina and were also more likely to be licensed in a place other than their reported *residence* jurisdiction than were not-at-fault drivers.

Regarding other risk factors, drivers residing in neighboring states and beyond were more likely to be indicated as sleepy or fatigued (6.3%, adjacent; 11.2% other) under driver physical condition than were North Carolina drivers (2.6%) and less likely to be indicated as impaired due to alcohol, drugs, or medications (15.8%, adjacent; 9.7% other versus 16.3% NC residents) (data not shown). The pattern was similar, though not as pronounced for those licensed out of state.

Though the data were somewhat sparse, multi-vehicle at-fault drivers appeared somewhat more likely to have a commercial driver license (CDL) than did not-at-fault drivers.

For the driver-related variables, it was clear that teens (16-19-years-old) and 20-29-year-olds were over-represented in both types of at-fault crashes relative to not-at-fault drivers. Drivers ages 30-39-, 40-49-, 50-59-, and 60-69-years-old were under-represented in both types of at-fault crashes. The pattern was mixed between the fault groups for older-aged drivers. While drivers aged 70-79 and 80+-years-old were under-represented in single vehicle at-fault crashes, they were over-represented in multi-vehicle at-fault crashes. Not surprisingly, male drivers were over-represented in both multi- and single-vehicle at-fault crashes. With regard to the race/ethnicity of the drivers, lower percentages of Whites were observed for both types of at-fault crashes relative to their representation in the not-at-fault population. On the contrary, Hispanic drivers were over-represented in both types of at-fault crashes, and Native American drivers were over-represented in single-vehicle at-fault crashes. Finally, Black drivers were under-represented and Asian/Other ethnicity drivers were over-represented in multi-vehicle at-fault crashes.

Both groups of at-fault drivers were less likely to have worn a safety belt in the study crash than were not-at-fault drivers, particularly the single-vehicle at-fault drivers. Safety belt information was also more likely to be missing for both groups of at-fault drivers. Both groups of at-fault drivers had overrepresentations of impaired (by alcohol, drugs, or medication), sleepy/fatigued, ‘other impairment’, and missing/unknown driver physical conditions in the study crash. With regard to suspected or detected alcohol and drug use in the study crash as reported by law enforcement, both groups of at-fault drivers were more likely to have evidence/be suspected of alcohol use, drug use, both alcohol and drug use (combined), or missing information regarding alcohol/drug use than were not at-fault drivers.

For the vehicle-related variables, older model year vehicles (i.e., older than model year 1995) were found to be more prevalent among both groups of at-fault drivers in the study crash than was the case for not at-fault drivers. The types of vehicles involved in the study crash also differed as a function of driver fault. For multi-vehicle at-fault drivers, cars were over-represented, and pickup trucks, SUVs, and vans/minivans were under-represented. For single vehicle at-fault drivers, cars were again over-represented, but only vans/minivans were found to be under-represented. Both groups of at-fault drivers were more likely to have one or more vehicle defects noted during the study crash, although there were a large number of missing values (30% or more) across all three groups of drivers. At-fault drivers of both types were also more likely to have not been insured during the study crash, particularly single-vehicle at-fault drivers, and were less likely to have one passenger in the vehicle than not-at-fault drivers. Furthermore, multi-vehicle at-fault drivers were less likely to have two or more passengers in their vehicle during the study crash.

Recall that the environment-related variables would not be expected to differ much between the at-fault and not-at-fault multi-vehicle crash-involved drivers because they are inherently matched by crash. Any differences observed would be due to the fact that there could only be one at-fault driver involved in a multi-vehicle crash in this study, but there could potentially be more than one not at-fault driver. This assumption was mostly supported in the univariate comparisons of the environment variables between at-fault and not-at-fault multi-vehicle crash-involved drivers. The only environment variables for which the multi-vehicle at-fault drivers differed from those not at-fault were road class and development type, for which the multi-vehicle at-fault drivers were over-represented on rural roadways and under-represented on North Carolina/U.S./State roadways and in commercial/industrial areas. These differences were statistically accounted for in the multivariate analyses by clustering the multi-vehicle groups within crash.

For the single-vehicle at-fault drivers there were many more environment-related differences compared to the multi-vehicle not-at-fault drivers, as expected. While they did not differ on season, single vehicle at-fault drivers were over-represented on weekends, during the hours of 9:00 pm to midnight and 1:00 am to 4:00 am, on interstates and local roads, in rural localities and farming/undeveloped and residential development areas, and on roadways with 55 MPH or higher speed limits. These differences were adjusted for in the multivariate analyses by including these variables as covariates before considering other potential predictor variables.

Licensing database predictors

The 42 licensing database predictor variables were also used in a series of chi square tests of independence for the purposes of testing their potential discriminating ability between multiple-vehicle not-at-fault drivers and the two groups of at-fault drivers (Table 10). Standardized cell residuals and follow-up chi square tests were again used to establish which levels of the variables differed between at-fault and not-at-fault drivers when a significant omnibus chi square test was observed. Only the results for the "license status at time of crash" variable include all 21,786 study drivers. All other univariate licensing database variable results in the table exclude the out-of-state and unmatched licensing status drivers because they did not have North Carolina licensing records that could be used to create the variables. The out-of-state and not determinable proportions combined were 6.5% of not-at-fault drivers, 7.2% of multi-vehicle at-fault drivers, and 7.4% of single-vehicle at-fault drivers. The column percentages for these other variables are based on 7,101 multi-vehicle crash not-at-fault (NAF) drivers, 6,098 multi-vehicle crash at-fault drivers (MVAF), and 7,054 single-vehicle crash at-fault drivers (SVAF) for an overall sample size of 20,253.

Table 10. Percentages of Multiple-Vehicle Not-At-Fault, Multiple-Vehicle At-Fault, and Single-Vehicle At-Fault Fatal and Serious-Injury Crash-Involved Drivers in each Level of the Licensing Database Predictor Variables and Associated Omnibus Chi Square Results

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
License Status at Time of Crash				
LICENSE STATUS AT STUDY CRASH				1145.76*
FULLY NC LICENSED	84.9	74.1	63.7	
SUSPENDED/LIMITED PRIVILEGE NC LICENSE	3.4	10.1	17.5	
NC UNLICENSED	5.2	8.6	11.3	
OUT-OF-STATE LICENSE	5.2	6.0	5.9	
NOT DETERMINABLE	1.3	1.2	1.5	
MISSING	0.0	0.0	0.0	
License Restriction History Variables				
OVERALL NUMBER OF LICENSE RESTRICTIONS IN PRIOR 5 YEARS				141.14*
NONE	65.3	61.7	69.8	
ONE	33.2	35.5	27.4	
TWO+	1.5	2.8	2.8	
MISSING	0.0	0.0	0.0	
CORRECTIVE LENSES RESTRICTION IN PRIOR 5 YEARS				304.73*
NO	68.2	69.4	80.2	
YES	31.8	30.6	19.8	
MISSING	0.0	0.0	0.0	
45MPH/NO INTERSTATE RESTRICTION IN PRIOR 5 YEARS				22.81*
NO	99.8	99.4	99.8	
YES	0.2	0.6	0.2	
MISSING	0.0	0.0	0.0	
DAYLIGHT DRIVING ONLY RESTRICTION IN PRIOR 5 YEARS				27.31*
NO	99.9	99.4	99.8	
YES	0.1	0.6	0.2	
MISSING				
INTRASTATE DRIVING ONLY RESTRICTION IN PRIOR 5 YEARS				4.21
NO	99.6	99.8	99.7	
YES	0.4	0.2	0.3	
MISSING	0.0	0.0	0.0	
OTHER LICENSE RESTRICTION IN PRIOR 5 YEARS				228.85*
NO	98.5	95.8	93.4	
YES	1.5	4.2	6.5	
MISSING	0.0	0.0	0.0	
ONLY DRIVE WHILE SUPERVISED RESTRICTION IN PRIOR 5 YEARS				12.93*
NO	99.6	99.2	99.1	
YES	0.4	0.8	0.9	
MISSING	0.0	0.0	0.0	
AUTOMATIC TRANSMISSION ONLY RESTRICTION IN PRIOR 5 YEARS				2.73
NO	99.9	99.9	99.9	
YES	0.1	0.1	0.1	

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
MISSING	0.0	0.0	0.0	
DRIVE CLASS B AND C ONLY RESTRICTION IN PRIOR 5 YEARS				18.78*
NO	99.7	99.9	99.9	
YES	0.3	0.1	0.1	
MISSING	0.0	0.0	0.0	
GRADUATED DRIVER LICENSING RESTRICTION IN PRIOR 5 YEARS				154.02*
NO	99.1	96.1	95.9	
YES	0.9	3.9	4.1	
MISSING	0.0	0.0	0.0	
BLOOD ALCOHOL CONTENT < .04 RESTRICTION IN PRIOR 5 YEARS				13.07*
NO	99.6	99.4	99.2	
YES	0.4	0.6	0.8	
MISSING	0.0	0.0	0.0	
BLOOD ALCOHOL CONTENT = .00 RESTRICTION IN PRIOR 5 YEARS				8.28*
NO	99.9	99.9	99.9	
YES	0.1	0.1	0.1	
MISSING	0.0	0.0	0.0	
Crash History Variables				
NUMBER OF PRIOR NOT-AT-FAULT CRASHES IN PRIOR 5 YEARS				18.70*
NONE	74.7	77.0	77.0	
ONE	19.9	17.9	17.6	
TWO	4.3	3.9	4.0	
THREE+	1.2	1.2	1.4	
MISSING	0.0	0.0	0.0	
NUMBER OF PRIOR AT-FAULT CRASHES IN PRIOR 5 YEARS				166.15*
NONE	86.9	81.2	79.9	
ONE	11.4	14.8	16.3	
TWO	1.5	3.1	2.9	
THREE+	0.2	0.9	0.9	
MISSING	0.0	0.0	0.0	
AT-FAULT IN MOST RECENT PRIOR CRASH INVOLVEMENT IN PRIOR 5 YEARS				131.28*
NO/NO CRASH	89.3	84.3	82.7	
YES	10.7	15.7	17.3	
MISSING	0.0	0.0	0.0	
AT-FAULT 2ND MOST RECENT PRIOR CRASH INVOLVEMENT IN PRIOR 5 YEARS				53.87*
NO/NO CRASH	97.0	94.6	94.8	
YES	3.0	5.4	5.2	
MISSING	0.0	0.0	0.0	
AT-FAULT 3RD MOST RECENT PRIOR CRASH INVOLVEMENT IN PRIOR 5 YEARS				28.44*
NO/NO CRASH	99.1	98.1	98.2	
YES	0.9	1.9	1.8	
MISSING	0.0	0.0	0.0	
AGE DIFFERENCE AT MOST RECENT PRIOR CRASH				65.41*

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
INVOLVEMENT IN PRIOR 5 YEARS				
SAME AGE	5.0	6.2	7.4	
ONE YEAR YOUNGER	9.3	10.3	10.5	
TWO YEARS YOUNGER	7.5	6.7	7.1	
THREE YEARS YOUNGER	5.5	5.6	5.8	
FOUR YEARS YOUNGER	4.5	4.7	3.6	
FIVE YEARS YOUNGER	2.1	1.5	1.9	
NO CRASH PRIOR 5 YEARS	66.0	65.0	63.6	
MISSING	0.1	0.0	0.1	

Conviction History Variables

OVERALL NUMBER OF CONVICTIONS IN PRIOR 5 YEARS				644.98*
NONE	68.8	61.4	52.2	
ONE	17.3	17.2	18.1	
TWO	6.9	9.4	11.1	
THREE	3.7	4.9	7.0	
FOUR	1.4	2.8	4.4	
FIVE+	1.8	4.3	7.2	
MISSING	0.0	0.0	0.0	
NUMBER OF LICENSE DEMERIT POINTS IN PRIOR 5 YEARS				226.38*
NONE	74.6	69.9	64.4	
1-2 POINTS	7.4	7.4	7.4	
3-4 POINTS	10.5	12.5	15.4	
5-6 POINTS	4.1	4.9	6.0	
7-8 POINTS	1.5	2.3	2.6	
9+ POINTS	1.8	3.0	4.1	
MISSING	0.0	0.0	0.0	
NUMBER OF NON-MOVING CONVICTIONS IN PRIOR 5 YEARS				52.95*
NO CONVICTIONS	97.9	97.1	95.8	
ONE	1.9	2.4	3.5	
TWO+	0.2	0.5	0.7	
MISSING	0.0	0.0	0.0	
NUMBER OF LICENSE-RELATED CONVICTIONS IN PRIOR 5 YEARS				349.14*
NO CONVICTIONS	95.1	90.5	86.0	
ONE	3.4	6.2	8.8	
TWO+	1.5	3.2	5.1	
MISSING	0.0	0.0	0.0	
NUMBER OF SPEED-RELATED CONVICTIONS IN PRIOR 5 YEARS				62.47*
NO CONVICTIONS	76.3	75.1	71.1	
ONE	15.9	15.6	18.2	
TWO+	7.8	9.3	10.7	
MISSING	0.0	0.0	0.0	
NUMBER OF ALCOHOL-RELATED CONVICTIONS IN PRIOR 5 YEARS				693.27*
NO CONVICTIONS	97.6	92.3	85.4	
ONE	0.8	1.9	3.8	
TWO+	1.6	5.8	10.8	
MISSING	0.0	0.0	0.0	
NUMBER OF OTHER MOVING CONVICTIONS IN PRIOR 5 YEARS				64.63*
NO CONVICTIONS	93.7	90.9	90.2	
ONE	5.6	7.8	8.4	
TWO+	0.7	1.3	1.4	

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
MISSING	0.0	0.0	0.0	
NUMBER OF OTHER SERIOUS CONVICTIONS IN PRIOR 5 YEARS				159.39*
NO CONVICTIONS	98.8	97.4	95.5	
ONE	1.1	2.3	3.7	
TWO+	0.1	0.3	0.8	
MISSING	0.0	0.0	0.0	

Suspension History Variables

OVERALL NUMBER OF LICENSE SUSPENSIONS IN PRIOR 5 YEARS				1113.27*
NONE	87.7	77.8	64.8	
ONE	8.9	14.4	20.5	
TWO	2.4	5.4	9.2	
THREE	0.9	2.1	4.8	
FOUR	0.1	0.3	0.7	
MISSING	0.0	0.0	0.0	
ANY LICENSE SUSPENSION IN PRIOR 5 YEARS (ANYSUSP5YR_S)				1048.44*
NO	87.7	77.8	64.7	
YES	12.3	22.2	35.2	
MISSING	0.0	0.0	0.0	
ADMINISTRATIVE LICENSE SUSPENSION IN PRIOR 5 YEARS				5.87†
NO	99.9	99.9	99.7	
YES	0.1	0.1	0.3	
MISSING	0.0	0.0	0.0	
ALCOHOL-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				948.74*
NO	96.7	89.5	80.5	
YES	3.3	10.5	19.5	
MISSING	0.0	0.0	0.0	
DRIVING WHILE SUSPENDED LICENSE SUSPENSION IN PRIOR 5 YEARS				336.33*
NO	97.6	95.3	90.7	
YES	2.3	4.7	9.3	
MISSING	0.0	0.0	0.0	
FATAL INVOLVEMENT LICENSE SUSPENSION IN PRIOR 5 YEARS				4.85†
NO	100.0	99.9	99.9	
YES	0.0	0.1	0.1	
MISSING	0.0	0.0	0.0	
FAILURE TO APPEAR/ FAILURE TO PAY LICENSE SUSPENSION IN PRIOR 5 YEARS				349.83*
NO	91.3	86.8	80.5	
YES	8.7	13.2	19.5	
MISSING	0.0	0.0	0.0	
NOT TRAFFIC RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				68.67*
NO	99.3	98.8	97.7	
YES	0.7	1.2	2.3	
MISSING	0.0	0.0	0.0	
OTHER LICENSE SUSPENSION IN PRIOR 5 YEARS (OTHER5YR_S)				14.44*

Predictor variable	%NAF	%MVAF	%SVAF	χ^2
NO	99.8	99.7	99.4	
YES	0.2	0.3	0.6	
MISSING	0.0	0.0	0.0	
PHYSICAL OR MENTAL CONDITION LICENSE SUSPENSION IN PRIOR 5 YEARS				11.92*
NO	99.9	99.7	99.7	
YES	0.1	0.2	0.3	
MISSING	0.0	0.0	0.0	
POINTS-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				16.51*
NO	99.6	99.4	99.1	
YES	0.4	0.6	0.9	
MISSING	0.0	0.0	0.0	
GRADUATED DRIVER LICENSING-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				27.36*
NO	99.7	99.6	99.1	
YES	0.3	0.4	0.9	
MISSING	0.0	0.0	0.0	
RECKLESS DRIVING LICENSE SUSPENSION IN PRIOR 5 YEARS				12.48*
NO	99.9	99.9	99.6	
YES	0.1	0.1	0.4	
MISSING	0.0	0.0	0.0	
SERIOUS OFFENSE LICENSE SUSPENSION IN PRIOR 5 YEARS				29.82*
NO	99.9	99.8	99.5	
YES	0.1	0.2	0.5	
MISSING	0.0	0.0	0.0	
SPEED-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				42.15*
NO	99.3	99.1	98.2	
YES	0.7	0.9	1.8	
MISSING	0.0	0.0	0.0	

Note. NAF = multi-vehicle crash not at-fault drivers; MVAF = multi-vehicle crash at-fault drivers; SVAF = single vehicle at-fault drivers. BAC = blood alcohol content.
†$p < .10$. *$p < .05$.

With regard to licensing status at the time of the study crash, both groups of at-fault drivers were less likely to be fully licensed in North Carolina when the crash occurred than were not-at-fault drivers. That is, higher percentages of at-fault drivers involved in both multi-vehicle and single-vehicle crashes were driving on a suspended or limited privilege license, driving without any license, or were licensed in another jurisdiction. The proportions of drivers for whom the licensing status was not determinable did not differ across the three groups of drivers.

Of the various licensing restriction variables, the one broadest in scope summarized the total number of restrictions on record, regardless of the type of restriction. Multi-vehicle at-fault drivers were more likely to have either one restriction or two or more restrictions on record than were not-at-fault drivers. This was not the same pattern as was the case for single-vehicle at-fault drivers. Specifically, the single-vehicle at-fault drivers were less likely to have only one license restriction on record compared to the not-at-fault drivers. However, they were more likely to have two or more restrictions on record.

The remaining restriction variables indicated whether or not a driver had a specific type of restriction on record. The multi-vehicle at-fault drivers were more likely to have the following types of restrictions on record: 45MPH/No interstate driving, daylight driving only, some “other” license restriction, supervised-only driving, a GDL restriction, and a < .04 Blood Alcohol Content (BAC) restriction. Multi-vehicle at-fault drivers were *less* likely to have intrastate driving only and Class B/C-only restrictions on record than were not-at-fault drivers. The single-vehicle at-fault drivers were more likely to have restrictions on record for daylight driving only, some “other” license restriction, supervised driving only, a GDL restriction, a < .04 BAC restriction, and an = .00 BAC restriction. They were less likely to have corrective lenses and Class B/C-only license restrictions on record than were not-at-fault drivers.

Past crashes have been found to be one of the best predictors of both future crash involvement and crash culpability. For the most part the expected pattern of high numbers of prior crashes being associated with being at-fault for the study crash was apparent, though not in the clean stepwise fashion that might be expected. Both multi-vehicle and single-vehicle at-fault drivers were more likely to have no prior not-at-fault crashes compared to multi-vehicle not-at-fault drivers. Multi-vehicle at-fault drivers were directionally *less* likely to have one or two 5-year prior not-at-fault crashes than were not-at-fault drivers, but the results were not precise enough to be conclusive. The single vehicle at-fault drivers were significantly less likely to have only one 5-year prior not-at-fault crash. None of the other differences were precise enough to draw conclusions. For both groups of at-fault drivers the expected pattern of prior crashes was much more apparent when at-fault 5-year prior crashes were considered. That is, both multi- and single-vehicle at-fault drivers were more likely to have one, two, or three or more prior at-fault crashes than were not-at-fault drivers. This suggests a stronger association of prior crash record to study crash culpability when only the prior at-fault crashes are considered.

Three other variables shown in Table 10 were used to represent whether or not drivers involved in prior crashes were at-fault in their three temporally nearest crashes, with fault coded separately for each prior crash. These variables suggested that both single- and multi-vehicle at-fault drivers were more likely than not-at-fault drivers to also be at-fault in their three prior temporally proximal crashes. One final variable reflecting the prior crash record of the drivers was intended to represent the age difference of the crash-involved drivers at the most recent prior crash. However, interpretation of the results of the univariate analyses for this variable does not seem to contribute hugely to understanding differences between at-fault and not-at-fault drivers. While both groups of at-fault drivers were more likely to have been the same age or one year younger at their most recent prior crash, the remaining differences were varied. In some cases the at-fault drivers had their crashes more recently in time, but there was not a very clear relationship.

Prior convictions have also been found in past studies to be strong predictors of future crash involvement and crash fault. Indeed, higher percentages of both groups of at-fault drivers had two, three, four, or five+ 5-year prior convictions on record. The single-vehicle at-fault drivers in particular were much more likely to have a non-zero prior conviction record than were the not-at-fault drivers.

The prior numbers of demerit points of drivers have also been found to be useful predictors of subsequent crash history. The use of points attempts to take into consideration that some types of convictions are deemed more serious by society and therefore are weighted more heavily in the assignment of points. The analyses of 5-year prior points indicated that both groups of at-fault drivers were more likely to have 3 or more demerit points on record from the prior 5 years than were not-at-fault drivers.

The remaining conviction-related variables broadly characterized different types of convictions (e.g., speeding-related, non-moving convictions, etc.). With only one exception, the at-fault drivers had worse prior conviction histories, regardless of the type of conviction. Both the multi- and single-vehicle at-fault drivers were more likely to have higher numbers of non-moving convictions, license-related convictions, alcohol-related convictions, 'other' moving convictions, and serious convictions (e.g., reckless driving). The one partial exception was with regard to speeding convictions. Whereas the single-vehicle at-fault drivers were more likely to have one or two+ convictions on record, multi-vehicle at-fault drivers were only more likely to have two+ speeding-related convictions on record than not-at-fault drivers.

The final set of licensing database variables were all related to the past license suspension history of the crash-involved drivers. The first variable was the broadest and characterized the total number of license suspensions on record in the 5-years prior to the study crash. Both groups of at-fault drivers, but particularly the single-vehicle at-fault drivers, were more likely to have one, two, three, or four+ license suspensions in the prior 5 years than were the not-at-fault drivers. This was also true when past suspensions were dichotomized as yes/no with 22% of the multi-vehicle at-fault and 32% of the single-vehicle at-fault drivers having at least one license suspension in the prior 5 years compared to 12% for the not-at-fault drivers.

The remaining suspension variables were used to characterize whether the drivers had specific types of license restrictions. The multi-vehicle at-fault drivers were more likely than not-at-fault drivers to have a prior suspension for alcohol use (e.g., DWI, refusing chemical test), driving while suspended, a failure to appear in court/pay a fine (FTA/FTP), non-traffic related reasons (e.g., child support, writing bad checks), a physical or mental condition (e.g., failure to submit medical report or reexamination), demerit points, or a serious offense (e.g., felony evading police, felony involving motor vehicle). The single-vehicle at-fault drivers were more likely than those not-at-fault to have a suspension of record in the prior 5 years for administrative reasons (e.g., provided false information, false application), alcohol use, driving on a suspended license, involvement in a fatality (e.g., death by vehicle, manslaughter), FTA/FTP, non-traffic related reasons, some 'other' reason (e.g., court suspension, issue error), a physical or mental condition, accumulated demerit points, graduated driver license accumulated convictions, reckless driving, a serious offense, or speed-related convictions (e.g., excessive speeding over limit, speed competition).

Multivariate Prediction of Driver Culpability

The next analyses involved using multivariate combinations of the crash and licensing database predictor variables in logistic regression models to determine which combinations of variables best discriminated the two groups of at-fault drivers from those not at-fault in the crashes. Multivariate models can consider the effects of multiple factors simultaneously, and therefore control for the effects of other factors before determining whether a factor is significantly correlated with the dependent variable, in this case, fault in the study crash.

Table 11. Table showing parameters of eight driver risk models developed.

***Multi*-vehicle Models (two or more drivers involved)**	***Single*-vehicle Models (only one driver involved)**
#1: Includes **all** in- and out-of-state **drivers** (meeting the study criteria) – Model developed using **crash data** variables from NCCDS only	#5: Includes **all** in- and out-of-state **drivers** (meeting the study criteria) – Model developed using **crash data** variables from NCCDS only
#2: Like #1, except **NC drivers only** (same set of drivers as in models #3 and #4) – using **crash data** only	#6: Like #5, except **NC drivers only** (same set of drivers as in models #7 and #8) – using **crash data** only
#3: Includes **NC drivers** - using **driver history** data from NCDMV database	#7: Includes **NC drivers** - using **driver history** data from NCDMV database
#4: Combined model – **NC drivers** – allowing variables from either **crash data *or* driver history** to go into the model	#8: Combined model – **NC drivers** – allowing variables from either **crash data *or* driver history** to go into the model

Separate models were used to identify predictor variables for the multi-vehicle at-fault drivers (models 1 to 4) and single-vehicle at-fault drivers (models 5 to 8, see Table 11). Models 1 and 5 each included the entire set of involved drivers which allowed us to examine predictive factors from the crash data with unlicensed, suspended license, out-of-state licensed, and drivers whose status could not be determined all included. In models 2 and 6, undeterminable and out-of-state drivers were excluded since a driver history could not be identified for these drivers. We kept in suspended and unlicensed drivers who had identifiable records in the driver history database, even though the records could be incomplete.

In the multi-vehicle at-fault models, logistic regression analyses with generalized estimating equations (GEEs) and robust variances were used to account for the inherent clustering of the at-fault and not-at-fault drivers within crashes. In addition, the environment-related variables were used to model any residual differences on these variables due to the inclusion of more than one not-at-fault driver in some crashes. In the single-vehicle at-fault models, the environment-related predictor variables were used to adjust for differences in crash environment between the single vehicle at-fault drivers and the multi-vehicle not-at-fault drivers used to represent the exposure distribution of drivers in general.

For each at-fault model the same procedure was followed in the model development process. First, the crash database predictors were used to create empirical prediction models for discriminating the at-fault from not-at-fault drivers, including drivers with out-of-state and undetermined license status. Next the empirical crash database models were repeated for just the subset of drivers who had a North Carolina license (suspended or not) or who were known North Carolina residents driving unlicensed. This was done to restrict the data to persons for whom licensing database variables were also available. The next models used the same sample of North Carolina residents, but only licensing database variables were used as predictors. The final stage was to create empirical models for these same drivers from combined crash and licensing database variables. The models were developed by first including all variables in the model and then removing through a manual backwards-elimination process those which did not have an observed Wald test with a significance level less than .10. For each stage the degree to which the models were able to discriminate at-fault and not-at-fault drivers was ascertained using 50% and 75% probability cut-points for classification tables.

Results

A total of 24 crash-based and 42 license/driver history factors were tested in the models. The multivariate modeling results are summarized in the following sections. Detailed tables of results and narrative descriptions of significant predictors of fault, along with model classification outcomes are provided in Appendix 2. The models were used to classify the set of drivers that were used in model development to assess the models accuracy at predicting actual fault status.

Multi-vehicle crash fault

Multivariate results reflect the results of controlling for the influence of other factors in the models.

Crash data models

In general, results of models 1 and 2 of fault in multi-vehicle crashes were very similar, except that when all drivers were included in the model (model 1), two additional factors were associated with fault. All other significant factors were the same between models, and had the same relationship with fault and similar odds ratios. In model 1, being licensed in other states, particularly non-adjacent states, was more highly associated with being at fault in a crash than being licensed in NC (from Appendix 2, Model 1). In addition, having a license from a state other than the indicated residence state also increased the odds of being at fault.[1]

---

[1] Odds ratios are a measure of the likelihood of an event relative to the likelihood of another event, in this case, the odds of one level of the variable being at fault compared to another level. Odds ratios are provided in the detailed tables of results, but due to the difficulty in interpreting them,

Only NC drivers were included in model 2 for comparison and matching to models including driver histories (only available for NC drivers), and neither of these factors remained as significant predictors. Examining the other driver characteristics from the ***crash data*** that were significantly predictive (within 95% confidence limits) of being at fault in a multi-vehicle crash (from model 2, NC drivers only), factors were:

- Possession of a commercial driver license (CDL)
- Driver ages:
  - Younger drivers, 16 to 19 years
  - 20 to 29 years (increased odds compared with middle-aged drivers)
  - Older drivers, 60 to 69 years, 70 to 79 years, and 80+ years (increasing with age)
- Driver ethnic group:
  - Hispanic
  - Asian/Other
- Driver belt use (reported): Unbelted
- Driver physical condition:
  - Alcohol impaired
  - Sleepy/fatigued
  - Other impairment
- Alcohol/drug use suspected/detected:
  - Alcohol only
  - Drugs use only suspected/detected
  - Both alcohol and drugs suspected/detected suspected
- Number of passengers: Carrying no passengers at time of the crash
- Driving a car compared with pickup, sport utility, or van/minivan
- Driving older than 1995 model year vehicle.(trend)

---

they will not be included specifically in the results summarized below. In the case of driver age, the youngest group was used as the reference group, so, in this case the relative odds ratio of young drivers being at fault compared to middle aged drivers are 1 / (OR for aged 40 to 49 = .27) = 3.7

License history model

Turning to the ***driver history*** (including demographic) ***predictors***, factors that were associated with fault in a multi-vehicle crash (model 3, including driver history data only, plus certain demographic variables available in either database) included:

- Possession of a commercial driver license
- Driver age:
  - 16 to 19 years
  - 20 to 29 years
  - 70 to 79 years
  - 80+ years
- Driver race: Native American
  - Hispanic
  - Asian/Other
- License status at time of crash:
  - Unlicensed
  - Suspended/limited privilege
- Overall number of license restrictions in prior 5 years: None was associated with a greater risk of being at fault (corrective lenses accounts for the majority of restrictions)
- Individual restrictions that were over-represented among at-fault drivers included:
  - 45 mph/no interstate driving restriction in prior 5 years
  - Daylight driving only restriction
  - Drive only while supervised restriction
  - Automatic transmission restriction
  - GDL restriction
  - Blood alcohol content < .04
- Number of not-at-fault crashes in prior five years: Having ***no*** prior crashes was more associated with being at fault in the study crash than having one or two prior not-at-fault crashes.
- Number of at-fault crashes in prior 5 years:
  - Two, and

- Three or more was associated with increasing odds of being at fault

- At-fault in 2nd most recent crash involvement in prior five years

- Age difference to most recent prior crash involvement: Being the same age (crash occurred within the same year) or three years younger were associated with fault

- Speed related convictions: Having one (significant), or two or more (non-significant trend)

- Number of other moving convictions in prior five years:

  - Two+ other moving convictions

  - Having only one other moving conviction was associated with *reduced* odds of being at fault

- Overall number of license suspensions in prior five years:

  - Drivers with one prior suspension

  - Drivers with three prior suspensions

  - (Two and four also had higher odds compared with no prior, but non-significant.)

- Failure to Appear/Failure to Pay License Suspension: Reduced odds of being at fault compared with not having this suspension

Combined model

When variables from both the crash databases and the driver history data were allowed into the model (model 4), all of the same crash-based predictors remained in the model as in model 2 using ***crash-based predictors*** only (results shown above). The odds ratios for the crash-based predictors decreased somewhat, but remained substantively the same, as some of the variance was accounted for by the driver history predictors. In particular, characteristics associated with fault in multi-vehicle crashes included the following:

- Possession of a commercial driver license
- Young driver age, particularly teen-age (16 to 19)
- Older driver age, particularly 70 to 79 and 80+ years
- Hispanic and Asian-American/Other ethnic groups
- Driver (reported) unbelted at time of study crash
- Any kind of physical impairment at time of the study crash (alcohol, drugs, and medications; sleepy/fatigued; other)
- Alcohol use, drug use, or both (in study crash)

- ➢ Older model vehicles (older than 1995) (trend)
- ➢ Driving a car, compared to pickups, vans, SUVs
- ➢ Carrying no passengers

Nine ***driver history*** predictors were also selected in this model as being associated with fault in the crash.

- ➢ License status:
  - Unlicensed drivers
  - Suspended/limited privilege license drivers
- Drivers without a Class B or C only license restriction in prior 5 years
- ➢ Drivers with a GDL restriction within prior 5 years
- ➢ Drivers with no prior not-at fault crashes in the prior 5 years were somewhat more likely to be at fault than those with one prior not-at-fault crash
- ➢ Drivers with one, two, or three or more prior at-fault crashes were increasingly likely to be at fault in the study crash.
- Those with 3-4, or 5-6 driver demerit points were somewhat more likely to be at fault than those with no driver demerit points in the prior 5 years. The relationship for those with more driver demerit points was less certain.
- Drivers with two or more non-moving violations convictions in the prior five years
- Drivers with two or more alcohol-related convictions
- Those with a driving while suspended license suspension were *less* likely to be at fault
- Drivers with a GDL-related suspension in the prior 5 years were *less* likely to be at fault

Comparatively, the only license database predictors that added predictive power in both the driver history only (model 3) and crash and driver history (combined, model 4) models (shown with an arrow symbol) were license status, number of prior crashes (both at fault and not at fault), and GDL restriction in the prior five years. The associations with fault of these four factors were in general the same, or in the same direction. The lack of consistent predictive power for many of the driver history factors suggests that many of these variables may lack the stability needed for good predictors. These results may be due in some part to the small sample sizes for many of the variables, as well as to overlap in predictive power with variation that was better captured by risk factors from the crash data variables.

Some environmental covariates also remained in the multi-vehicle models, although analyses clustered drivers within a crash to control for most environmental factors. These results were likely due to having more than one not-at-fault driver for some collisions which would tend to over-emphasize the locations and times where collisions involving more than two drivers occurred.

Single-vehicle Crash Fault

Crash data models

Models 5 to 8 examined fault in single-vehicle collisions using the not-at-fault group in multi-vehicle collisions as the comparison group. The models were completed in similar order as for the multi-vehicle models.

Model 5 examined fault for all drivers including out-of-state drivers involved in single-vehicle collisions, using crash data only. Model 6 was the same as model 5, except that only NC drivers were included. In the case of single-vehicle drivers, there were essentially no differences in the findings of models 5 and 6; no factors associated predominantly with out-of-state drivers helped to discriminate fault in single-vehicle collisions. A very similar set of predictors emerged compared with models of fault in the multi-vehicle collisions. The strength of the relationship was generally stronger, however, in models of single-vehicle fault. The following summary is based on results from model 6, NC drivers only.

Crash-data based predictors associated with fault (from model 6, NC drivers only) included:

- Driver ages: 16 to 19 years
  - 20 to 29 years (slightly increased odds)
  - 60 to 69 year-olds had the lowest odds, but not significantly different than middle-aged.
- Driver ethnic group:
  - Hispanic
  - Asian/Other
  - Native American (reduced odds of being at fault, trend)
- Driver belt use (reported): Unbelted
- Driver physical condition: Alcohol/drugs/medications impaired
  - Sleepy/fatigued
  - Other impairment
- Alcohol/drug use suspected/detected: Alcohol only

- Drugs use only suspected/detected
- Both alcohol and drugs suspected/detected

- Vehicle type: Car- and SUV-drivers had about twice the OR of pickups and vans for at-fault involvement
- Number of passengers: Carrying two or more passengers at time of the crash slightly increased the odds, compared with having none, or especially only 1 passenger (trend)
- Vehicle not insured

Additionally, crashes were more likely to occur at night or during early morning hours, between the hours of 9 pm to 9 am, but especially between the hours of 1 and 5 am. Mixed and rural locality types were over-represented, as were residential, and farm/undeveloped, higher speed limit roads, and rural and interstate roads.

License history model

Driver history or demographic factors that were predictive of single-vehicle crashes (driver history only model, # 7) included the following:

- Driver age: 16 – 19 years
    - 20 – 29 years
- Driver sex: males were at increased odds of being at fault
- Driver ethnic group: Native Americans
    - Blacks, reduced odds of being at fault
- ➢ License status at study crash:
    - Unlicensed drivers
    - Suspended license drivers
- Overall number of license restrictions in prior 5 years: Drivers with one or two plus restrictions were under-represented compared with those with no restrictions
- Restrictions that were over-represented among at-fault drivers:
    - GDL restriction
    - Daylight driving only
    - Driver only while supervised
    - "Other" restrictions
- Restrictions that were under-represented among at-fault drivers:

- Automatic transmission restriction
- Drive class B and C only restriction
- BAC < .04 restriction

- Number of at-fault crashes in prior 5 years, from One prior to Three or more, increasingly associated with fault
- Age difference to most recent prior crash
  - Same age (within same year) – overrepresented compared with two and four years younger (and trend compared with one, three, or five years younger or no prior crashes)
- Overall number of convictions over prior 5 years
  - One, Two, and Four, significantly associated with fault
  - Three and Five or more (trend)
- Number of 'other' serious convictions (such as reckless driving) in prior 5 years, especially two or more
- Overall number of license suspensions in prior 5 years, from 1 to 4+ were increasingly associated with fault:
- Some individual types of license suspension (groups) that were *under-represented* among at-fault drivers:
  - Administrative
  - Driving while suspended, suspension
  - Failure to appear/Failure to Pay
  - Not traffic-related
  - Other suspension
  - Points-related suspension
  - GDL-related suspension
  - Reckless driving-related suspension
  - Speed-related suspension

Combined model

When driver history and crash variables were allowed in the ***combined model*** to determine factors associated with single-vehicle crash involvement, many of the license history predictors were no longer significant, whereas the factors from the crash-only model remained similarly predictive in the combined model. Significant predictors of fault in single-vehicle crashes were as follows:

- The youngest two groups of drivers were still predictive of fault compared with other ages:
  - 16 to 19 year-olds more so than 20 to 29-year-olds
- There were again trends associated with race/ethnicity, with Hispanic and Asian/Other tending to be over-represented and Native American tended to be underrepresented compared with Whites and Blacks (but not significant at .05 level)

Other driver factors from the crash variables had the same associations with fault as in the crash-based only model (model 6) and odds ratios were very similar. These included:

- Driving unbelted (reported)
- Impaired by alcohol, drugs or medications, sleep/fatigue, or other
- Suspected alcohol or drug use or both
- Driving an SUV or car in the crash, compared to pickups and vans
- Driving an uninsured vehicle
- Carrying two or more passengers (trend)

The ***driver history*** variables that remained in model 8 to predict single-vehicle crash involvement included:

- ➢ License status at study crash (although the strength of the predictors was lower than in the license-history only mode):
  - Unlicensed status
  - Suspended license status

Overall restrictions were not predictive of fault in the combined model.

- The only two individual restrictions (and the odds remained very similar as in model 7) associated with fault compared with not having that restriction that remained in the combined model included:
  - ➢ Daylight driving only

- ➢ GDL Restriction

- Having three or more *not*-at-fault crashes in prior five years was associated increased odds of being at fault (This factor was not significant in the license-history only model.)

- ➢ Prior at-fault crashes in prior five years were increasingly associated with fault
    - Two
    - Three or more

- ➢ Two or more other serious convictions in prior five years

- Alcohol-related license suspension in prior five years

- ➢ A reckless driving suspension was associated with reduced odds compared with not having that suspension

As in the models of multi-vehicle fault, the license history predictors were less stable between the license-data only models and the combined models. The factors that were consistently predictive of fault (again shown with the arrow bullet) between the two models were license status (increased risk for suspended and unlicensed drivers), prior at-fault crashes, having a daylight-driving only or GDL restriction, having especially two or more other serious convictions, or a reckless driving suspension (reduced odds). These factors added predictive value that was not entirely accounted for by the crash-based variables.

Again, the association of late night hours; rural and interstate, and higher speed roads; farm/undeveloped or residential development type were associated with single-vehicle collisions as compared with crash involvement for multi-vehicle not-at-fault drivers.

Summary of Model Classification Outcomes

To help interpret the myriad models conducted as part of this study, Table 12 presents the percentage of not-at-fault drivers (specificity), percentage of at-fault drivers (sensitivity), and overall percentage of drivers that were correctly classified (efficiency) based on each of the eight models and for both the 50% and 75% probability cut points. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut point) or .75 (75% probability cut point). The higher cut-point identifies fewer at-fault drivers but has higher specificity with regard to not-at-fault drivers.

Table 12. **Summary of Models 1-8 Classification Outcomes for Predicting At-Fault Drivers.**

	50% cut-point			*75% cut-point*		
Model	**%NAF**	**%AF**	***%e***	**%NAF**	**%AF**	***%e***
Multi-vehicle at-fault models						
Model 1 (crash-only, all drivers)	87	45	69	99	19	64
Model 2 (crash-only, NC drivers)	87	45	69	99	18	64
Model 3 (license-only, NC drivers)	80	44	63	98	9	57
Model 4 (crash & license, NC drivers)	87	47	70	99	20	65
Single-vehicle at-fault models						
Model 5 (crash-only, all drivers)	91	76	84	97	62	82
Model 6 (crash-only, NC drivers)	91	76	84	98	62	82
Model 7 (license-only, NC drivers)	77	60	69	95	30	62
Model 8 (crash & license, NC drivers)	92	77	85	98	64	83

Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off). %NAF = percentage not-at-fault predicted by model. %AF = percentage at-fault predicted by the model. %*e* = overall percentage of not-at-fault and at-fault drivers correctly classified by the model (efficiency).

Several summary comments can be made by looking at the table values. First, the overall prediction and efficiency was higher for single-vehicle at-fault drivers. That is, given the equivalent predictor pool of variables, the single-vehicle at-fault crash models were always more efficient than the multiple-vehicle at-fault crash models, (even in test models with no environmental covariates included in the license-only data models). This outcome is particularly true when it came to classifying those who were at-fault, whereby the single vehicle models had much higher sensitivity. Both the single-vehicle and multi-vehicle models had relatively good specificity (classification of not-at-fault drivers). Next, the crash-based predictors were superior for correctly classifying both sets of at-fault drivers. Finally, when the licensed-based predictors were combined with the crash-based predictors, the overall efficiency did not increase more than 2 percentage points (or 4% improvement) for the multiple-vehicle models and by only 1% (1.3% improvement) for the single-vehicle at-fault driver models. Together the results suggest that there is a lot of overlap in predictive power of the two sets of variables given that they provide little synergistic increase in efficiency when combined in the same models. In other words, they may often be measuring similar underlying attributes.

In these exploratory models, we did not attempt to add interaction effects or test model goodness of fit or improvement in predictive power with each variable added to the model due to the large number of variables of interest that were examined for relationship to fault. Thus, it is possible that models with better fit and potentially better predictive power could be developed with a subset of the more valuable predictors of fault that were identified.

Key Findings

Univariate Trends

Drivers involved in single vehicle crashes and multi-vehicle collisions varied substantially in the types of driver contributing circumstances cited in their crashes. At-fault drivers in multiple-vehicle crashes were more likely to be cited for:

- Failure to yield right-of-way (32%)
- Crossing centerline/going wrong way (29%)
- Inattention/distracted (18%)
- Erratic, reckless, careless, negligent, aggressive driving (12%)
- Alcohol use (11%)

Single-vehicle drivers were more likely to be cited for:

- Exceeding safe speed for conditions (33%)
- Erratic, reckless, careless, negligent, aggressive driving (33%)
- Exceeding authorized speed limit (32%)
- Alcohol use (31%)
- Crossed centerline/going wrong way (20%)

Single vehicle drivers were much more likely to be indicated as speeding and reckless/careless/aggressive driving, and alcohol use than multi-vehicle at-fault drivers.

In addition, nearly two-thirds of single-vehicle drivers were cited for two or more contributing circumstances, while about half of multi-vehicle at-fault drivers were cited for two or more. Since driver contributing circumstances were used to establish fault in the study crash, they were not used as predictor variables in the subsequent models.

The predominant crash times and road classes where these collisions occurred were also different. Single-vehicle crash-involvement occurred in locations characterized as predominantly rural and on interstate and rural roads, with 53% of single-vehicle drivers' involvement occurring on roads classed as rural (compared with 30% for multi-vehicle at-fault) and another 7% on interstates (compared with 4% of multi-vehicle at-fault drivers). Multi-vehicle crash involvement was most highly associated with NC/US/State roads (39% of at-fault drivers).

A number of driver risk factors were over-represented among the two fault groups compared with not-at-fault drivers in multi-vehicle collisions (results shown in Tables 9 and 10). Among these were young driver ages (16 to 19, and 20 to 29), older age (especially 80+, for multi-vehicle fault only), males, drivers reported unbelted in the study crash, drivers with any type of physical impairment in the study crash; and various driver history factors including driving on suspended or unlicensed at study crash, having graduated driver or ‘other’ license restrictions in the prior five years, prior at fault crashes, overall number of convictions, license demerit points, alcohol-related convictions, speed-related convictions and drivers with any or increasing numbers of license suspensions as well as a number of individual suspensions, convictions or restriction types. The factors that varied significantly in a univariate sense were eligible as predictors in the multivariate analyses.

A number of driver factors also varied significantly between multi-vehicle at-fault and single-vehicle fault groups. These results, in addition to the different crash characteristics described above, supported the approach of developing separate models for single vehicle and multi-vehicle at-fault status.

Multivariate Outcomes

A number of driver characteristics and behavioral indicators were found to be predictive of fault in the crashes.

Driver factors associated with fault in the study crash

Of the ***crash-based predictors***, those consistently predictive of being at fault (crash-only and combined models) included:

	Multi-vehicle fault	Single-vehicle
Young driver age (16 – 19 years)	✓	✓
Drivers aged 20 – 29	✓	✓
Older driver age (increasing from 60+ years)	✓	
Driver reported to be Hispanic (crash data)	✓	
Driver reported to be Asian/other (crash data)	✓	
Driver reported unbelted in study crash	✓	✓
Being impaired by alcohol, drugs, or medication	✓	✓
Being impaired due to sleep/fatigue	✓	✓
Being impaired due to other causes (medical, physical)	✓	✓
Alcohol use suspected/detected	✓	✓
Drug use suspected/detected	✓	✓
Alcohol and drugs suspected/detected	✓	✓
Driving a passenger car	✓	✓
Driving an SUV	✓	
Driving with no passengers	✓	

Driving with 2 or more passengers		✓
Driving an older model (pre-1995) vehicle	✓	
Having a commercial driver license	✓	

In addition, about 6% of drivers responsible for multi-vehicle collisions, and slightly more of those responsible for single-vehicle collisions were residents of adjacent or other states/countries, and two factors associated with these drivers were predictive in the multi-vehicle crash model that included them. The model including out-of-state drivers showed that drivers with out-of-state licenses, and those with licenses that were not the same as the indicated residence state, were also more likely to be at fault in these collisions.

The ***license-based predictors*** that were reliably predictive (consistent results between license data only and combined models) included:

	Multi-vehicle fault	**Single-vehicle**
Driving on a suspended license at study crash	✓	✓
Unlicensed at study crash	✓	✓
Increasing prior at-fault crashes (esp. > 1)	✓	✓
Having no prior not-at fault crashes	✓	
Graduated Driver License restriction	✓	✓
Daylight driving only restriction		✓
Two+ "Other" serious convictions		✓

Among all of the significant factors, impairments due to alcohol or drugs, sleep/fatigue, lack of belt use, and being a teen driver (all from the crash data) were consistently strong predictors of fault (based on odds ratios) in both single-vehicle and multi-vehicle crashes. In addition, possession of a graduated license restriction (driver history) was predictive of fault in both single-vehicle and multi-vehicle crashes. Driving on a suspended license was a stronger predictor of fault than driving unlicensed in each model. It should be borne in mind, that some suspended license or unlicensed drivers may have had reduced exposure while others obviously continue to drive. In addition, the driver histories created for these drivers may be incomplete due to the difficulty of tracking them. Since there were obviously a number of improperly licensed drivers involved in crashes, it was deemed important to keep them in the analyses, with the assumption being that the induced exposure method would account for the types and amount of driving (exposure). Having prior at-fault crashes was also predictive in all models that included driver history factors.

Although the relative odds of being at fault were lower than for the youngest aged drivers, the 20 to 29 year age group represented 41% of all drivers indicated as being impaired due to alcohol, drugs, or medications, and 23% of this age group, compared with 16% over all ages, were indicated as having an alcohol, drug, or medical impairment. The other risk factors of impairment and non-belt use are also highly over-represented among both 20 to 29-year-old drivers and male drivers

Being male was not independently predictive of fault in the models, with one exception, the single-vehicle, driver history data only model (#7). Other risk factors such as impairments evidently explain the apparent over-representation of males in fault groups.

Relationships of many license-based predictors beyond those listed above were less consistent and less reliable from license-data only to combined models. Many of these predictors, such as specific types of restrictions or suspensions had small sample sizes, and apparently factors in the crash data provided better prediction of fault status. A few additional license-history based factors were predictive of fault in both single-vehicle and multi-vehicle models using license data only, but were non-significant when crash variables were allowed, including:

- Having *no* prior restrictions (recall that most restrictions are corrective lens restriction)
- Drive only while supervised restriction
- Increasing number of total suspensions (especially strong predictor of single-vehicle)
- Being the same age as at most recent prior crash

Finally, certain ***environmental covariates characterized single-vehicle collisions***. These collisions were more apt to occur:

- At nighttime (after 9 pm) to late at night
- On rural roads and interstates compared with NC/US/State roads
- In rural and mixed urban/rural locations compared with urban areas
- In residential and farming/undeveloped areas compared with commercial/industrial areas
- On roadways with 55 mph or greater speed limits.

These latter environmental factors are consistent with what is expected about single-vehicle crash occurrence. The results may also suggest patterns of driver behavior that might be targeted by countermeasures in addition to targeting the types of locations/roadways where, and times of day when, these crashes occur.

Although we used a “nested within crashes” analyses to control for environmental differences, several environmental co-variates continued to be significant predictors of fault in multi-vehicle collisions. This result is likely because we kept in the study population not-at-fault drivers involved in collisions with more than 2 vehicles, as long as there was only a single at-fault driver implicated. We could not justify a reason to exclude them. Therefore the influence of some environmental covariates continued to be significant since we did not have a completely matched-pairs design. The environmental results in these models are somewhat difficult to interpret, and may simply reflect where collisions involving more than two vehicles occurred (more congested times of day, roads, etc.).

In general, variables from the crash itself were more 'predictive' of fault in the crash than were the license history variables. Even summary measures of convictions, suspensions, or restrictions were not reliably predictive. Very slight model improvement (1% for single-vehicle, 4% for multi-vehicle), based on correct classification of fault status over the crash-based only models was gained by allowing either license-history variables or crash-based variables to enter the model predicting fault in multi-vehicle or single-vehicle crashes.

Models predicting fault in single-vehicle collisions were more efficient at correctly classifying drivers than those predicting multi-vehicle crash-involved drivers. These results are perhaps not unexpected, since the group of at-fault drivers were more similar to the not-at-fault drivers involved in the same collisions. The univariate tables as well as odds ratios of predictor variables in the models suggest however that the available predictors/risk factors drivers are more highly associated with drivers involved in single-vehicle collisions than with the set of multi-vehicle, at-fault drivers.

Discussion and Recommendations

In this study, a wide range of driver, vehicle, and environmental factors available or created from the State crash database and driver license history files were explored in connection with involvement in serious injury crashes in North Carolina. Our results support earlier work that indicates that the characteristics of drivers primarily contributing to single-vehicle crashes and those involved in multi-vehicle crashes are different – at least a significant portion of them have different characteristics, as can be seen in the univariate frequency distributions as well as the model outcomes. In these exploratory models, we have attempted, not only to identify driver factors associated with causing these two types of serious crashes, but to explore which set or blend of variables provided better prediction of driver fault status, crash data variables, or driver history variables. Thus, various measures of license history suspensions, restrictions, convictions, and prior crashes were created and tested and compared with the predictive power of crash-based variables. Ultimately, such models might be used to, not only identify general, population-based characteristics of risky drivers, but to actually identify individual risky drivers for targeted interventions. Such efforts have been used, with some success by the California DMV to target interventions.

Fault in multi-vehicle crashes proved more difficult to discriminate than that in single-vehicle crashes. One explanation of the difficulty of 'predicting' driver fault in multi-vehicle collisions from crash-based driver factors and driver history elements may be that a number of these crashes result from errors or temporary lapses in judgment during more congested times of day and on more crowded roadways. These tendencies may not be captured by pre-disposing risk tendencies that are assumed to be represented by, for example the license-history data as well as some of the risk factors in the crash data. As evidence, 18% of multi-vehicle at-fault drivers were cited for inattention, and 14% for failure to reduce speed. It is likely to also be the case that even if a driver is not cited as contributing to the crash, there may have been some culpability. And the at-fault and not-at-fault drivers in these crashes did not different as substantially in a univariate sense, on many of the risk factors.

Our results generally corroborate findings of earlier research from other jurisdictions, and identified driver age (younger and older drivers in multi-vehicle, younger only in single-vehicle), alcohol and other impairments, sleep/fatigue, lack of safety belt use, driving on suspended license or unlicensed, having a GDL restriction, and prior at-fault crash involvements as consistently being predictive of fault in both types of crashes. Our results also support Gebers' (2003) assertion that increased crash risk for males is largely a function of exposure, and/or other risk factors that are more associated with males. In this study, the over-representation of males with other risk factors such as alcohol use may explain their apparent over-representation in crashes.

In general, the variables available from the crash itself, in particular impairments due to alcohol/drugs/or medication, or sleep/fatigue, are more closely and reliably associated with fault in that crash than the license history variables tested. It is notable that one license factor that was reliably predictive - GDL restriction - which also correlates strongly with the youngest age group, provided additive predictive power. We, unfortunately could not determine age at first licensure, or length of licensure from the data. Other factors that also correlated highly with each other including alcohol use indicators and physical condition (impairments), both from the study crash, also provided separate predictive power. These results suggest that as long as these factors are not completely confounded with each other, they are so highly correlated with fault that they continue to add predictive value to the model equations. As to the importance of variables from the study crash, certainly, if a driver is impaired at the moment of the crash, he/she is very likely to bear some responsibility for that crash. Nevertheless, if that reported impairment or lack of belt use, age, and other factors indicate patterns of risky driving behaviors, then those variables might also help to predict future crash risk. In light of the better association of fault with these factors than those from the driver history, it may be possible to use crash data to identify risky drivers for intervention.

Of the driver history variables, the most consistently predictive of fault were increasing number of prior at-fault crash involvements (which actually came from the crash data), driving with a suspended license or being unlicensed at the time of the study crash, having a GDL restriction (associated with young, beginning drivers) or having a daylight-driving only restriction (for single-vehicle crashes). Even when driver history elements were summed, such as total convictions found predictive in previous studies, they were not universally predictive of fault. "Prior total convictions" *was* associated with fault in single-vehicle crashes when only license history data was used. Individual conviction types, suspension types, or other restriction types were also not consistently predictive of fault, with the exception of "Other serious" convictions which was predictive of fault in single-vehicle crashes. In contrast, "prior at-fault crashes" was predictive of fault in both multi- and single-vehicle crashes, with or without crash data variables also in the model. There are other examples of specific types of convictions, suspensions, etc. that were predictive in driver-history only or combined-data only models, but not both types of models. Whether this bodes ill for the possibility of using driver history elements in future model development aimed toward identifying individual risky drivers for targeted interventions is not entirely clear.

The relatively few reliable predictors of fault that derived from the driver history may result from the driver history not being reflective of drivers' actual risk. Obviously it is difficult for prior histories to capture all the risk that is evident from the behaviors and conditions at the time of the crash. However, if drivers are not detected violating traffic laws, or if detected and not convicted of that infraction, the record will not reflect the drivers' past risky behavior. There are evidently just such problems throughout the adjudication, and record-keeping policies and processes in NC. (Some of these may have begun to be addressed by legislation.) News reports published by The News & Observer in a May 2007 series reported that loopholes in prosecuting and sentencing drivers have resulted in flagrant and repeat speeding violators escaping punishment by receiving dismissals, plea agreements to lesser charges that carry no points, or prayer for judgment continued pleas, which do not show up on driver records at all. The second article in the series states that "During the year that ended June 30, four out of every five speeding drivers had their charges reduced or dismissed or were given a 'prayer for judgment continued' (Stith, Raynor, and Locke, 2007a)." Other reports indicated that individual drivers repeatedly escaped punishment for speeding when courts dismissed or lowered charges. According to the articles, many of these dispositions are a result of plea deals between prosecutors and courts to reduce the burden on overloaded courts, but there reports also indicated that some judges are far more lenient than others (Locke, Stith, and Raynor, 2007).

Furthermore, some drivers may use false names and driver licenses that further compound the problem of using driver records to accurately reflect risk factors to identify individual drivers for treatment. Perhaps most importantly, the intended use of enforcement to deter further risky driving through the belief that punishment will be "certain, swift and severe" is seriously weakened.

As a follow-up to the model results in this study, we examined what proportion of those cited (contributing circumstances) for excessive speed or exceeding speed limits had prior records for convictions or other sanctions related to speeding. Similarly, for those with alcohol-related impairments or detected/suspected alcohol use in the study crash, what proportion had a record suggesting prior enforcement and judicial contact relating to alcohol? Our examination pooled any convictions, suspensions, or restrictions that were obviously speeding-related or alcohol-related by driver, and cross-tabulated these summed past contacts with the crash-based evidence. Of those with speed indicated as a factor in the study crash, a three-fourths majority (73%) had no prior evidence of speed-related enforcement actions in their driver records. The proportion of crash-involved at-fault drivers with prior sanctions was slightly higher (27%) than for those who did not have speeding cited as a contributing circumstance (23%). There was also a minority (37 – 38%) of those with alcohol/impairment indicators in the study crash that had any prior alcohol-related driver record. Unknown is what proportion of those drivers considered to be speeding or impaired by alcohol in the study crash are habitual or problem speeders or problem drinkers. Also unknown is what percentage had been detected at all, versus the percentage detected who escaped punishment related to the offense, but the above comparisons suggest a statistical reason that driver license histories are less predictive of fault than crash data.

In addition to the above, some elements in the driver record are not directly related to driving behavior, but may reflect punishments enacted by the courts for other reasons (child support, failure to attend school and so forth). We attempted to capture these non-driving related penalties in “Other” or “Administrative” categories, but they are not always easy to distinguish. Some prior studies have identified other types of non-traffic infractions, and indeed felony convictions as being associated with risky drivers, but the non-traffic infractions did not clearly emerge as a factor in this study. We did not examine felony convictions, except those that were traffic-related and in the driver records.

The relationship between punishments dispensed and drivers’ response is also not clear, however, based on both prior research and the present study results. While some drivers may respond to convictions or more serious penalties by driving within the law/reducing crash risk, others may not. A recent study in Maryland found that legal consequences for speeding had little impact on future citations (Lawpoolsri & Li, 2007). In fact, those receiving citations had a greater risk of receiving a subsequent citation compared with those not cited in the first time interval. More severe penalties (fines and points) also had no impact on drivers’ subsequent citations, although the authors note that the severity of penalties for speeding was relatively low (Lawpoolsri & Li, 2007). In another study, Williams, Kyrychenko, and Retting (2006) identified characteristics of drivers observed speeding more than 15 miles above the limit on 13 high volume Virginia roadways, compared with those driving no more than 5 miles above the limit on the same roadways at the same times. Speeders were younger, drove newer vehicles, and had more speeding and other traffic convictions on their records as well as 60% more prior crashes. Many other studies have documented increased risk of traffic crashes among those with prior convictions. Convictions, however, have been associated with a reduction in personal risk of fatal crash involvement over a short time interval (the two months following, but not beyond 3-4 months) after a conviction (Redelmeier, Tibshirani, & Evans, 2003). Most evidence suggests that there is at least a population of drivers for whom penalties may not have the desired deterrent effect, at least over a longer term.

The relationship between suspensions and driver behavior is perhaps most perplexing. During this project, considerable thought was given to how to treat suspended drivers, unlicensed drivers, or driver with less than a full five years of prior history. Since unlicensed drivers or those with suspensions should not be driving, their exposure should be reduced. It was obvious from the crash evidence that a significant percentage of drivers involved in crashes had a suspended license, or were unlicensed at the time. Therefore, we made the assumption that the induced exposure comparison would account for the amount and types of driving by those with suspended licenses or no licenses (or recently licensed); this assumption could not, however, be tested. Results suggested that drivers with increasing numbers of prior suspensions indeed had greater odds of being at fault, although some types of individual suspensions did seem to reduce the odds, perhaps inducing more care in driving, if not to abandon the idea altogether. It is clear from the results of this study, results that may be understated due to the difficulties in associating prior histories with drivers who use false identities and addresses, that unlicensed and suspended licensed drivers continue to drive and cause serious crashes.

It is important in terms of developing or identifying effective countermeasures to understand the causal relationship between driver risk factors and crash tendency. Most of the factors identified by this study have been previously established to have predictive, and likely causally-related links to driver crash risk. With alcohol or drug use, the link is quite clear. For other factors, understanding the underlying causal association with crash risk is more difficult. For example, why are those with prior convictions and suspensions at increased risk for crashes? Encounters with law enforcement and the judicial system are intended to deter future violations and risky driving, yet for some population of drivers, even severe sanctions obviously don't have the intended effect. Clearly, more serious assessment of driver remediation strategies for these drivers is warranted. A body of literature is building that examine the various driver psychological, social, and behavioral risk factors that may underlie continued disrespect for traffic laws and punishments. Efforts should continue to understand these serious problems and develop appropriate strategies to address them.

Other factors that have a less clear causal association with fault include possession of a CDL license, age of vehicle, and drivers' license state/residence state situation. Most of these factors could be argued to be measuring some underlying risk, including age of vehicles. Vehicle age could be, in part, related to socio-economic/employment/educational status, found in some prior studies to be associated with greater crash propensity. Additionally Williams et al. (2005) found that young drivers who owned their own vehicles were more likely to own older-aged and smaller vehicles, do more driving and more risky driving, and to have more traffic violations and crashes. As for driver licensing/residence state status, drivers that have different licensing states compared to the indicated residence state might be consciously attempting to evade the law, in part, because they are risky drivers who don't want their records to follow them, or they could simply have failed to complete the requirements following a move. Out-of-state drivers may be susceptible to becoming fatigued or sleepy when driving long distances. Further study of commercial-vehicle licensed drivers might help to elucidate why drivers with commercial licenses seem to also be at greater risk of causing multi-vehicle (but not single-vehicle crashes) when behind the wheel of passenger vehicles, as in this study.

It also isn't clear why those driving passenger cars might be at increased risk of causing multi-vehicle collisions compared with other types of passenger vehicles and both SUVs and cars are more associated with single-vehicle crashes. One possibility is that drivers in passenger cars that are involved in crashes with larger vehicles (SUVs, pickups, and vans) might be more often seriously injured, and thus less able to represent their viewpoint when the crash is investigated. Broyles et al (2003) found that occupants of passenger cars were more than twice as likely as occupants of four-wheel-drive types of vehicles (for example, SUVs) to be injured although they did not explore severity of injuries. There is of course, a possibility that vehicle type does explain some risk, beyond that accounted for by driver age, license status and other factors, and is measuring some unknown, underlying phenomenon (such as increased rollover risk in some model SUVs).

The relationship with number of passengers is context dependent. Passengers seem to help drivers prevent multi-vehicle crashes, but having two or more passengers increased the likelihood of being at fault in a single-vehicle crash. This evidence supports the restriction on passengers for young drivers. Ages (and perhaps even gender) of passengers and drivers, time of day, alcohol influence and other factors probably mediate this relationship as illustrated by recent studies of young drivers (Rice, Peek-Asa, and Kraus, 2003 and others).

Countermeasures

To address serious traffic collisions, the study results suggest driver risk factors including alcohol/drug use, young age, non-belt use (which may not cause the crash but certainly contributes to crash severity and suggests the driver may have a pattern of risky behavior), and patterns of driving including speeding, and driving during night-time hours on rural, higher speed roadways, that should clearly be targeted by enforcement, and other, perhaps innovative, programs.

These behavioral characteristic are also ones that involve substantial proportions of at-fault drivers, particularly those involved in single-vehicle collisions. Although single-vehicle and multi-vehicle crashes have differing characteristics in terms of location types, times of occurrence, driver errors, etc., the key, personal driver risk factors are very much the same although more highly correlated with single vehicle collisions. If the same drivers tend to be over-represented in both types of crashes, then targeting countermeasures toward single-vehicle crash involved drivers could pay multiple dividends.

Young drivers

Younger drivers' inexperience, potentially combined with other risk factors including higher speeds and/or alcohol, may contribute to their risk particularly on rural, more dangerous roadways. The 20 to 29 year age group accounted for the largest proportion of impaired drivers, and was highly over-represented for impaired driving. Being the youngest age group that can legally consume alcohol (excepting 20-year-olds), there seems to be a need for interventions aimed at reducing drinking and driving among young adults.

Studies show that graduated driver licensing (GDL) as implemented in NC has substantially reduced crashes among young, beginning drivers (Foss, Feaganes & Rodgman, 2001). Recent analyses conducted by the Highway Safety Research Center have found that crash rates have declined by 34% for 16-year-old drivers and by 18% for 17-year-old drivers following the enactment of North Carolina's GDL system. Nonetheless, young drivers continue to be over-represented in both single- and multi-vehicle crashes. There are many potential countermeasures available to further reduce young driver crashes. Greater publicity and enforcement of GDL restrictions and other laws pertaining to young drivers, such as zero tolerance and seat belt laws, might encourage greater compliance with these laws. One recent program in North Carolina employing the "high visibility enforcement" model was modestly successful at increasing seat belt use among young drivers and compliance with GDL passenger restrictions (Goodwin et al., 2006). There are also many opportunities for assisting parents in managing their teens' driving. Parents are responsible for supervising their teens' early driving experience, determining the timing of licensure, governing access to (and choice of) vehicles, and imposing restrictions on driving privileges. Programs designed to assist parents with this process could potentially improve young driver safety. Other countermeasures for reducing crashes among teens include improving the content and delivery of driver education/training, eliminating early high school start times (which may contribute to drowsy driving among teens), and reviewing transportation plans for new or expanded high schools. Other modes of transportation than driving personal vehicles might be encouraged and supported through more effective transportation plans.

Additionally, creating a ‘safety culture’ among youth should begin much earlier than high school driver education. Psychology-based research shows that habit, or just being used to speeding has an effect, as do intention and personal norms, on reported speeding behavior (De Pelsmacker and Janssens, 2007). Thus, preventing those habits and norms from becoming entrenched among youth could have significant payoffs. Developing, delivering, and evaluating stronger traffic safety programs to help change these habits and perceptions among young children through teens could have a long-term safety improvement effect.

Traffic safety culture

Attention is in fact, being focused nationally, through efforts of the AAA Foundation for Traffic Safety and others, on ways of fostering a traffic safety culture among the driving public, and among enforcement and legal institutions as well as the popular media. While some groups may contribute disproportionately to serious injury crashes, targeting any one group such as those with suspended licenses, or repeat crashers, or young drivers can affect only a small portion of all future crashes. A number of thoughtful papers have been published by the Foundation that calls attention to many of the complex challenges to improving traffic safety in our society. These challenges include addressing a popular culture and media that glorifies speed and personal freedom, and facing the institutional cultures and inertia that often keep us from moving forward with a traffic safety paradigm shift (AAA Foundation for Traffic Safety, 2007).

For example, it is important for all involved agencies and institutions to accept, support, and promote the legitimacy of enforcement and effective punishment to improve safety, and begin to address the continued prevalence and social acceptability of speeding and other risky driving behaviors (McKenna, 2007). Clearly developing a Traffic Safety Culture needs the support of the courts, law enforcement, and the media. Without strong institutional belief in effective traffic safety enforcement, policy-makers nor the public are likely to be ‘sold’ on the idea, whether it is the need for more funds for regular traffic enforcement, enactment of appropriate penalties for infractions, or support for automated enforcement programs.

Enforcement

The perceived threat of detection is probably the most important factor in the effectiveness of an enforcement program (for a majority of drivers). Programs that maximize the perception that enforcement may be encountered anywhere, at any time will go furthest toward achieving real safety benefits. High visibility speed, alcohol, and seat belt enforcement, accompanied by extensive publicity to increase drivers’ perception of being detected are recommended to help to reduce serious crashes (*Countermeasures That Work*, 2007). Automated enforcement could certainly be used to help achieve these objectives, if done appropriately. Although they may be targeted toward problem locations, enforcement locations and times should not be advertised in advance. Programs aiming to reduce the social acceptability of speeding, aggressive driving (and drink-driving), also similar to early belt-use and anti-alcohol programs, could complement high visibility enforcement campaigns.

Enforcement programs should also be sustained over time, as effects of enforcement, even with convictions, are typically short-lived. A program designed to sustain long-term, wide-spread traffic enforcement (even without supplemental publicity - the Random Road Watch program) was apparently successful at reducing crashes of all severities in Queensland, Australia, with increasing effects up to three years after program introduction (Newstead, Cameron & Leggett, 2001). Indeed the Australians seem to be leading the way in learning from enforcement efforts and improving strategies and programs to maximize effectiveness. Targeting roadways that accounted for a high proportion of crashes, and scaling enforcement to available resources in each jurisdiction by using randomly scheduled, less intense enforcement to achieve more widespread coverage were considered to be keys to success (Delaney et al., 2003). Automated enforcement should also be considered to supplement regular enforcement efforts (Sivak, 2007; *Countermeasures That Work*, 2007) and could help improve sustainable enforcement presence and the perception that there is chance of being detected violating traffic laws.

To aid in enforcement efforts, spatial analyses could identify where concentrations of crashes associated with excessive speed, exceeding limit, careless and reckless driving, and alcohol impairment are occurring. If not already being done, both enforcement and engineering countermeasures could then be targeted toward locations where these behaviors are contributing to crash problems. As an aid toward this effort, Appendix 3 of this report provides maps and tables indicating the population-based rates of single-vehicle and multi-vehicle at-fault crash involvements by county. Some of the counties with higher single-vehicle or multi-vehicle crash involvement rates could be potential sites for conducting more in-depths studies of the problem and developing comprehensive model programs. (*Note:* Part of the variation in rates per population that cannot be ruled out through the databases examined in the present study, may be that agencies or personnel in different localities differ in crash reporting and the ways or extent to which contributing circumstances are cited.)

Judicial and administrative

As mentioned previously, the crash-based behavioral factors were stronger predictors of fault than most driver history elements. Driver history predictors would no doubt be better predictors if they were more reflective of drivers' actual behaviors/violations. Therefore, any efforts that may go into improving the association of judicial outcomes and driver records with drivers' actual behaviors (and repeat behaviors) will also aid efforts to identify, track, and treat problem drivers. Improving one of the cornerstones of deterrence, swift and certain punishment of illegal behavior, could also certainly enhance the effectiveness of enforcement.

If further model development were undertaken to identify individual risky drivers for targeted interventions, serious consideration should be given to the potential payoff of such efforts and the types of interventions that might prove most effective. Masten and Peck (2004) reviewed evaluations of 106 individual driver interventions and found that small but significant improvements in crashes and violations followed the use of warning letters, group meetings, individual hearings, and license suspensions/revocations. License suspension/revocation resulted in the most significant improvement, and the threat of license suspension may also have accounted for much of the effect of lesser measures such as warning letters. Again, such measures would be most effective with changes in supporting administrative, enforcement, and judicial policies in NC.

Case for innovative efforts

Enhance enforcement strategies and stringently applied driver sanctions will not alone solve the problems, however, as evidenced in this study and others. Repeat offenders, and those with impaired judgment due to alcohol, continue to drive unsafely even when they do receive severe sanctions such as suspended or revoked licenses. A 1998 study in Cleveland, Ohio examined interventions with 70 injured drivers admitted to a hospital with BACs of .10 or higher and found that 23 were cited for DUI, 15 (21%) were successfully prosecuted and convicted, only eight drivers (11%) were referred for outpatient alcohol counseling, but none were offered counseling as in-patients (Cydulka, et al., 1998). Five of these drivers received subsequent citations during the slightly more than two-year study period. While this was a single, small study, it perhaps speaks to the issue that drivers are being released back to the streets, in a majority of these cases without even a driving conviction, and in almost all cases, apparently without attempting to treat the underlying alcohol-related problem. No doubt, the individuals with alcohol-related traffic violations often have other alcohol-related problems as well. Serious consideration should be given to bringing traffic safety representatives, judicial representatives, social/behavioral scientists, and those involved in treating alcohol-related dysfunction together to originate meaningful interventions including policies for adjudication, treatment, and tracking and following-up on individual cases.

The success of the models in this study at identifying not-at-fault drivers with a high degree of specificity also suggests the possibility of providing some sort of positive reward system, similarly to some recently advertised insurance programs, as opposed to (only) punishing at-fault drivers. Such a program might provide incentive toward reducing the day-to-day risk-factors that creep into driving such as moderate speeding, failure to yield right-of-way, inattention and distraction, following too closely/failure to reduce speed, and other immediate causes that often contribute to multi-vehicle collisions. As an example, one study from the Netherlands reported on a successful program in which lease cars were equipped with technology that continuously monitored and displayed (providing immediate feedback) whether drivers were allowing a safe following distance and complying with the speed limit. Rewards were given by the lease company for good driving behavior over a 16 week period (Mazureck & van Hattem, 2006). Emerging Intelligent Transportation System (ITS) technologies may also help with a number of these issues, by providing in-vehicle speed limiters, immediate feedback, warnings, and/or rewards (Sivak et al., 2007). The State could support research into the effectiveness and optimal driver interfaces of these emerging technologies, some of which hold more promise than others, but none of which will be able to prevent all the crashes that they target (IIHS, 2008).

Priorities

It is of utmost importance to establish what actions will be or are likely to be effective, which is no small challenge. NCHRP's Report 500 Series, Volume 16: A Guide for Addressing Alcohol-Related Collisions describes strategies that have been tried, keys to success, legislative needs, other issues, costs, and considerations for developing effective countermeasures to the problem of alcohol-related collisions (Goodwin, Foss, Hedlund, Sohn, and Pfefer, 2005). Other volumes resulting from AASHTO's Strategic Highway Safety Plan address strategies to deal with non- or improperly-licensed drivers (Vol. 2), collisions involving drowsy and distracted drivers (Vol. 14), and many of the issues identified in this study. A forthcoming volume will address crashes involving young drivers. NHTSA also publishes a compendium, updated yearly, entitled Countermeasures That Work provides information on countermeasures to alcohol-related, speeding and aggressive driving, young driver, older driver, distracted driver, motorcycle, and pedestrian and bicyclist-related crashes. These Guides are just that, however – guides. The burden of developing effective state programs must be borne by state agencies, with perhaps considerable assistance from other public and private institutions.

Both the size of the problem and the potential for payoff should be considered in prioritizing safety programs. For example, very few of the at-fault crash involved drivers apparently possessed a commercial driver license, if the data are accurate. The over-involvement of drivers in possession of a commercial license in multi-passenger-vehicle collisions would likely therefore be a fairly low priority in targeting countermeasures, unless it were relatively easy to tag onto other programs targeting truck driver safety. Similarly, the numbers reported to be uninsured are small, although the accuracy of these data are suspect. If data accuracy could be improved by verifying insurance status, usefulness of the data could be increased. There may also be ways to better verify drivers' license status and personal information.

Suggestions for Additional Research

- Explore the relationship between violations cited and convictions and other penalties enacted to examine the gap between enforcement and adjudication. The full picture is not available when only conviction data is examined - drivers convicted of a lesser offense will appear to drive in a less risky fashion, and those not convicted will appear to have no risky prior behavior at all. Study the effects on subsequent driver behavior on differing adjudication outcomes.

- Conduct a thorough literature review of driver adjudication and remediation practices from other states/countries to identify effective practices. Determine what actions are presently available and what actions are presently taken in North Carolina against unlicensed/suspended license drivers involved in crashes, alcohol-impaired drivers, speeding drivers, repeat offenders, and drivers repeatedly at-fault in crashes. Suggest improvements in targeting penalties and other remediation to reduce crashes.

- Conduct survey or focus group studies of repeat offenders and drivers who violate restrictions and suspensions to determine why they continue to break the law, and what might be done to improve compliance and deter continued risky driving among hard core offenders.

  In addition to driver attitudes toward complying with laws and punishments and poor judgment when impaired, at least part of the problem may be due to perceived need: to drive to a job, to medical care, social opportunities, etc. Certainly those living in rural areas of North Carolina have few transportation options and isolation is probably not a good option. Further study of this question might be fruitful, particularly if transportation options can be provided to those who should not be behind the wheel.

- Refine models and further develop the ability to identify crash-prone drivers for targeted interventions. Based on the results of these model comparisons, and the association of prior at-fault crashes with subsequent crashes, it would be interesting to develop and test model effectiveness at predicting future crash-involved drivers based only on crash data. As adjudication and data issues are hopefully addressed effectively, the usefulness of driver history data to identify risky drivers would also increase.

  Around one-third of all drivers in the present study had one or more prior crashes (any type) within five years, and about 20% of at-fault drivers had prior at-fault crashes within the previous five years. Prior at-fault crash-involvement was the main driver history element (apart from license status) at present that was consistently predictive of at-fault crash involvement, and this element is actually obtained from crash data files, once the fact of a prior crash is gleaned from the license history.

Implementation and Technology Transfer Plan

The products of this research include summary profiles (histories and characteristics) of drivers cited with contributing to serious injury collisions in North Carolina, and suggested countermeasures that would involve law enforcement agencies, the judiciary, social/behavioral scientists, NCDOT/DMV, the NC Governor's Highway Safety Program, and the NC Executive Committee for Highway Safety (NC ECHS).

Uses for the information on driver risk factors developed from these analyses as well as information on the limitations of that information can be shared with appropriate agencies to help develop priorities for addressing the multi-pronged problems of serious injury traffic crashes in NC. Strategies may include changes in policies, procedures, and changes in institutional norms, implementation of new or improved enforcement and other behavioral programs, and possibly changes in legislation. A multi-agency/multi-disciplinary group, perhaps under the aegis of the NC ECHS, could lead the effort. A presentation of results and issues relating to the study may be made to NC ECHS.

Cited References

AAA Foundation for Traffic Safety (2007, April). Improving Traffic Safety Culture in the United States: The Journey Forward. AAA Foundation for Traffic Safety: Washington, D.C., April, 2007, 388 pp. Retrieved from: http://www.aaafoundation.org/pdf/SafetyCultureReport.pdf

Broyles, R.W., L. Narine, S.R. Clarke, and D.R. Baker (2003). Factors associated with the likelihood of injury resulting from collisions between four-wheel drive vehicles and passenger cars. Accident Analysis and Prevention 35: 677-681.

Chandraratna, S., N Stamatiadis, and A. Stromberg (2006). Crash involvement of drivers with multiple crashes. *Accident Analysis and Prevention* 38: 532-541.

Chen, W., P. Cooper, and M. Pinili (1995). Driver accident risk in relation to the penalty point system in British Columbia. *Journal of Safety Research* 26: 9 – 18.

Cooper, P.J. (1997) The relationship between speeding behaviour (as measured by violation convictions) and crash involvement. *Journal of Safety Researc*h 28: 83-95.

Countermeasures That Work: A Highway Safety Countermeasure Guide for State Highway Safety Offices. U.S. D.O.T., National Highway Traffic Safety Administration [DOT HS 809 980].

Daigneault, G., P. Joly, and J-Y. Frigon (2002). Previous convictions or accidents and the risk of subsequent accidents of older drivers. *Accident Analysis and Prevention* 34: 257-261.

Delaney, A., K. Diamantopoulou, and M. Cameron (2003). *MUARC's Speed Enforcement Research: Principles Learnt and Implications for Practice.* Monash University Accident Research Centre Report 200.

De Pelsmacker, P. and W. Janssens (2007). The effect of norms, attitudes and habits on speeding behavior: Scale development and model building and estimation. *Accident Analysis and Prevention* 39: 6-15.

Evans, L. and P. Wasielewski. (1982). Do accident –involved drivers exhibit riskier everyday driving behavior? *Accident Analysis and Prevention* 14: 57-64.

Foss, R.D., J.R. Feaganes, and E.A. Rodgman (2001). Initial effects of graduated driver licensing on 16-year-old driver crashes in North Carolina. *Journal of the American Medical Association (JAMA)* 286: 1588-1592.

Gebers M A (2003). An Inventory of California Driver Accident Risk Factors. California Department of Motor Vehicles, Research and Development Branch: Sacramento, CA [CAL-DMV-RSS-03-204] 86 pp.

Gebers, M.A.(1999). Strategies for Estimating Driver Accident Risk in Relation to California's Negligent-Operator Point System. California Department of Motor Vehicles, Research and Development Section: Sacramento, CA. [CAL-DMV-RSS-99-183] 51 pp.

Gebers, M.A. and R.C. Peck (2003). Using traffic conviction correlates to identify high accident-risk drivers. *Accident Analysis and Prevention* 35: 903-912.

Gebers M A and R.C. Peck (2000). Using Traffic Conviction Correlates to Identify High Accident-risk Drivers. California Department of Motor Vehicles, Research and Development Section: Sacramento, CA [CAL_DMV-RSS-00-187] 26 pp.

Gebers, M.S. and R.C. Peck (1994). An Inventory of California Driver Accident Risk Factors. California Department of Motor Vehicles, Research and Development Section: Sacramento, CA [CAL-DMV-RSS-94-144]

Goodwin, A., R. Foss, J. Hedlund, J. Sohn, and R. Pfefer etal. (2005). A Guide for Reducing Alcohol-Related Collisions. *NCHRP Report 500, Guidance For Implementation of the AASHTO Strategic Highway Safety Plan*, Volume 16. National Cooperative Highway Research Program, Transportation Research Board: Washington, DC.

Goodwin A.H., J.K. Wells, R.D. Foss, and A.F. Williams (2006). Encouraging compliance with graduated driver licensing restrictions. *Journal of Safety Research* 37: 343-351.

Hauer E., B.N. Persauad, A. Smiley, and D. Duncan (1991). Estimating the accident potential of an Ontario Driver. *Accident Analysis and Prevention* 23: 133-152.

IIHS (2008). Crash Avoidance Features. Insurance Institute for Highway Safety, *Status Report* 43.

Janke, M.K. (1991). Accidents, mileage, and the exaggeration of risk. *Accident Analysis and Prevention* 23: 183-188.

Lardelli-Claret, P., J.J. Jiménez-Moleón, J. de Dios Luna-del-Castillo, M. Garcia-Martin, A. Bueno-Cavanillas, and R. Gálvez-Vargas (2005). Driver dependent factors and the risk of causing a collision for two wheeled motor vehicles. *Injury Prevention* 11: 225-231.

Lardelli-Claret, P., J.J. Jiménez-Moleón, J. de Dios Luna-del-Castillo, M. Garcia-Martin, O. Moreno-Abril, and A. Bueno-Cavanillas (2006). Comparison between two quasi-induced exposure methods for studying risk factors for road crashes. *American J Epidemiology* 163: 188-195.

Locke, M., P. Stith, and D. Raynor (2007, May 20). Judge goes easy on 100-mph drivers. The News and Observer. Retrieved from: http://www.newsobserver.com/1545/story/576137.html

Masten, S.V. and R.C. Peck (2004). Problem driver mediation: A meta-analysis of the driver improvement literature. *Journal of Safety Research* 35: 403-425.

Mazureck, U., and van Hattem, J. (2006). Rewards for Safe Driving Behavior: Influence on Following Distance and Speed. *Transportation Research Record*: 1980: 31-38.

McKenna, F.P. (2007). The perceived legitimacy of intervention: A key feature for road safety. In *Improving Traffic Safety Culture in the United States: The Journey Forward*. AAA Foundation for Traffic Safety: Washington, D.C., April, 2007, 388 pp. Retrieved from: http://www.aaafoundation.org/pdf/SafetyCultureReport.pdf

Neuman, T.R., R. Pfefer, K.L. Slack, P. Waller. 2003. *Guidance for Implementation of the AASHTO Strategic highway Safety Plan, Volume 2: A Guide for Addressing Collisions Involving Unlicensed Drivers and Drivers with Suspended or Revoked Licenses.* NCHRP Report 500, Transportation Research Board: Washington, D.C.

Newstead, S. V., M.H. Cameron, and L.M.W. Leggett (2001). The crash reduction effectiveness of a network-wide traffic police deployment system. *Accident Analysis and Prevention*, *33*, 393-406.

Perneger, T. and G.S. Smith (1997). The driver's role in fatal two-car crashes: a paired "case-control" study. *American Journal of Epidemiolog*y 134: 1138-1145.

Redelmeier, D.A.,R. J. Tibshirani, and L. Evans. (2003) Traffic law enforcement and risk of death from motor-vehicle crashes: case-crossover study. *Lancet* 361: 2177-2182.

Rice, T.M., C. Peek-Asa, J.F. Kraus (2003). Nighttime driving, passenger transport, and injury crash rates of young drivers. *Injury Prevention* 9: 245-250.

Stamatiadis, N. and J.A. Deacon (1997). Quasi-induced exposure: Methodology and insight. *Accident Analysis and Prevention* 29: 37-52.

Stewart, J.R. and B.J. Campbell (1972). The Statistical Association between Past and Future Accidents and Violations. The University of North Carolina, Highway Safety Research Center, Chapel Hill, NC, 40 pp.

Stith, P., M. Locke, D. Raynor and (2007, May 15). Speed kills, courts shrug. The News & Observer. Retrieved from: http://www.newsobserver.com/1545/story/573762.html

Stith, P., D. Raynor and M. Locke (2007a, May 15). Cops write tickets; speeders get deals. The News & Observer. Retrieved from: http://www.newsobserver.com/1545/story/574288.html

Stith, P., D. Raynor and M. Locke (2007b, May 18). Speeders race through loopholes. . The News & Observer. Retrieved from: http://www.newsobserver.com/1545/story/575002.html

Wåhlberg, A.E. af (2003). Some methodological deficiencies in studies on traffic accident predictors. *Accident Analysis and Prevention* 35:473-486.

Williams, A.F., W.A. Leaf, B.G. Simons-Morton and J.L. Hartos (2005). Vehicles driven by teenagers in their first year of licensure, Traffic Injury Prevention 7: 23–30.

Appendix 1

Suspension Group used in Multiple Regression Analyses for each individual Suspension Class

Table 13. Suspension Group for each individual Suspension Class

New group	**group**	**SUS_DS**	
P&M	admin	FAILED TO APPEAR FOR REEXAMINATION	60
P&M	admin	FAILED TO APPEAR FOR REEXAMINATION (PRIOR TO CONVERSION)	7
Fta_ftp	admin	FAILURE TO APPEAR	10492
NTO	admin	FAILURE TO COMPLY WITH FINANCIAL RESP LAW OOS	1
Fta_ftp	admin	FAILURE TO DEPOSIT SECURITY	558
Fta_ftp	admin	FAILURE TO PAY FINE	2213
Fta_ftp	admin	RDLSI FAIL TO PAY FINE AND COST	9
Fta_ftp	admin	RDLSI FAIL TO POST SECURITY	1
Fta_ftp	admin	RDLSI FAILURE TO APPEAR (OBS)	3
Alcohol	alcohol	1 OFFENSE DRIVING WHILE IMPAIRED (OUT OF STATE)	5
Alcohol	alcohol	1 OFFENSE DRIVING WHILE IMPAIRED O/S	39
Alcohol	alcohol	1 OFFENSE OF DRIVING WHILE IMPAIRED	2843
Alcohol	alcohol	10 DAY CIVIL REVOCATION	2017
Alcohol	alcohol	2 OFFENSES DRIVING WHILE IMPAIRED (OUT OF STATE)	8
Alcohol	alcohol	2 OFFENSES OF DRIVING WHILE IMPAIRED	317
Alcohol	alcohol	3 OFFENSES DRIVING WHILE IMPAIRED (OUT OF STATE)	8
Alcohol	alcohol	3 OFFENSES OF DRIVING WHILE IMPAIRED	261
Alcohol	alcohol	30 DAY CIVIL REVOCATION(SUSPENSION)	3928
Alcohol	alcohol	CONSUMING ALCOHOL/DRUGS WHILE LESS THAN 21	270
Alcohol	alcohol	DRIVING WHILE IMPAIRED - MILITARY COURT	3
Alcohol	alcohol	FAIL TO COMPLY WITH CIVIL REVOCATION	39
Alcohol	alcohol	HABITUAL IMPAIRED DRIVING	34
Alcohol	alcohol	PROVISIONAL LICENSEE - USE ALCOHOL/DRUGS	49
Alcohol	alcohol	PROVISIONAL LICENSEE REFUSED CHEMICAL TEST	4
Alcohol	alcohol	RDLSI DRIVING WHILE IMPAIRED	1
Alcohol	alcohol	REFUSED CHEMICAL TEST	1085
Fatal	alcohol	REFUSED CHEMICAL TEST (ACC. W/DEATH)	26
Alcohol	alcohol	REFUSED CHEMICAL TEST ACTION SUSTAINED	25
Alcohol	alcohol	TRANSPORTING ALCOHOLIC BEVERAGES	2
Alcohol	alcohol	VIOLATION OF ALCOHOL RESTRICTION	10
Fatal	fatal	DEATH BY VEHICLE	137
Fatal	fatal	DEATH BY VEHICLE INVOLVING DWI	35

New group	group	SUS_DS	
Fatal	fatal	MANSLAUGHTER	24
Fatal	fatal	MANSLAUGHTER-2ND DEGREE MURDER INVOLVING DWI	15
P&M	incomp	ADJUDICATED INCOMPETENT	28
P&M	medical	DISAPPROVAL ADVISORY (DATE TO REAPPLY)	14
P&M	medical	DISAPPROVAL ADVISORY (NO DATE TO REAPPLY)	13
P&M	medical	DISAPPROVAL ADVISORY (PRIOR TO CONVERSION)	19
P&M	medical	DISAPPROVAL MEDICAL REVIEW BOARD	3
P&M	medical	DISAPPROVAL MEDICAL REVIEW BOARD (NO DATE TO REAPPLY)	3
P&M	medical	FAILED ROAD TEST FOR REEXAMINATION	5
P&M	medical	FAILED SIGN TEST FOR REEXAMINATION	2
P&M	medical	FAILED TO COMPLY WITH RESTRICTIONS	11
P&M	medical	FAILED TO SUBMIT CONSENT FORM OF MEDICAL REPORT	6
P&M	medical	FAILED TO SUBMIT FOR REEXAMINATION	4
P&M	medical	FAILED TO SUBMIT MEDICAL REPORT	96
P&M	medical	FAILED TO SUBMIT MENTAL/MUSC OF MEDICAL REPORT	1
P&M	medical	FAILED TO SUBMIT OCCUPATIONAL THERAPIST EVALUATION	2
P&M	medical	FAILED TO SUBMIT OTHER DOCUMENTS	12
P&M	medical	FAILED TO SUBMIT REQUIRED MED INFO (PRIOR TO CONVERSION)	12
P&M	medical	FAILED TO SUBMIT RESP/NEUROLOGIC OF MEDICAL REPORT	1
P&M	medical	FAILED TO SUBMIT SUBSTANCE ABUSE EVALUATION	18
P&M	medical	FAILED TO SUBMIT TO ROAD TEST	4
P&M	medical	FAILED TO SUBMIT VISION STATEMENT	43
P&M	medical	FAILED VISION TEST FOR REEXAMINATION	1
DWSL	nolic	1 OFFENSE OF DRIVING WHILE LICENSE SUSPENDED	1176
DWSL	nolic	1ST MOVING VIOLATION WHILE LICENSE SUSPENDED	1224
DWSL	nolic	2 OFFENSES OF DRIVING WHILE LICENSE SUSPENDED	528
DWSL	nolic	2ND MOVING VIOLATION WHILE LICENSE SUSPENDED	494
DWSL	nolic	3 OFFENSES OF DRIVING WHILE LICENSE SUSPENDED	783
DWSL	nolic	3RD MOVING VIOLATION WHILE LICENSE SUSPENDED	600
Reckless	other	2 OFFENSES OF RECKLESS DRIVING IN 12 MONTHS OOS	6
Reckless	other	2 OFFENSES OF RECKLESS DRIVING WITHIN 12 MONTHS	33
Reckless	other	2 OFFENSES OF RECKLESS/AGGRESSIVE DR 12 MO OOS	6
Reckless	other	2 OFFENSES OF RECKLESS/AGGRESSIVE DR IN 12 MO	28
NTO	other	BAD CHECK SUSPENSION	25

New group	group	SUS_DS	
Other	other	CANCELLATION OF CONDITIONAL RESTORATION	33
NTO	other	CHILD SUPPORT ISSUE (AOC)	4
NTO	other	CHILD SUPPORT ISSUE (DHR)	3
Other	other	COURT ORDER - CANNOT ISSUE LICENSE	3
Alcohol	other	COURT ORDER NOT TO OPERATE	189
Other	other	COURT SUSPENSION	39
NTO	other	DROPOUT PREVENTION SUSPENSION	335
Alcohol	other	DWLR - VIOLATION OF IGNITION INTERLOCK RESTRICTION (SUSP)	1
Fta_ftp	other	FAIL TO COMPLY WITH OUT-OF-STATE CITATION	327
Serious	other	FAIL TO STOP AND RENDER AID, HIT AND RUN	21
Alcohol	other	FAILURE TO COMPLETE A&D SCHOOL	3
NTO	other	FAILURE TO COMPLETE COMMUNITY SERVICE - SUSPENSION	3
Admin	other	FAILURE TO GIVE CORRECT INFORMATION	57
P&M*	other	FAILURE TO MEET QUALIFICATIONS	1
Serious	other	FELONY ELUDING ARREST - 2 AGGRAVATING FACTORS	7
Serious	other	FELONY ELUDING ARREST - 3 AGGRAVATING FACTORS	7
Serious	other	FELONY FORFEITURE OF LICENSING PRIVILEGES	42
Serious	other	FELONY IN WHICH A MOTOR VEHICLE WAS USED	1
Fatal	other	HIT AND RUN PERS INJ/DEATH	2
Other	other	IMPROPER USE OF OPERATOR LICENSE	2
Alcohol	other	INSTRUCTING WHILE SUBJECT TO IMPAIRING SUBSTANCE	4
Other	other	ISSUE ERROR	23
Serious	other	LARCENY OF A MOTOR VEHICLE	2
Other	other	LOSE CONTROL/LOSE YOUR LICENSE SUSPENSION	1
Admin	other	MAKING OR ALLOWING FALSE APPLICATION	4
Serious	other	MISDEMEANOR ELUDE ARREST	34
Fta_ftp	other	OUT-OF-STATE CONVICTION	89
Prov	other	PROVISIONAL LICENSEE 2 VIOLATIONS IN 12 MONTHS	274
Prov	other	PROVISIONAL LICENSEE 3 VIOLATIONS IN 12 MONTHS	77
Prov	other	PROVISIONAL LICENSEE 4 VIOLATIONS IN 12 MONTHS	15
Prov	other	PROVISIONAL LICENSEE 4 VIOLATIONS IN 12 MONTHS OOS	1
Other	other	RDLSI ACCUMULATION OF CONVICTS	1
Admin	other	RDLSI EXPIRED OR NO LIC PLATE OR DECAL (OBS)	1
Fta_ftp	other	RDLSI FAIL TO FILE DOC OR REPORT (OBS)	2

New group	group	SUS_DS	
Admin	other	RDLSI FAIL TO MAINTAIN LIAB INS	2
Fta_ftp	other	RDLSI FAIL TO MAKE REQ PAYMENT OF FEE (OBS)	1
NTO	other	RDLSI FAIL TO PAY CHILD SUPPORT	1
Fta_ftp	other	RDLSI FTA TRIAL OR COURT	6
Admin	other	RDLSI FTO MAKE REQ PAY OF TAX (OBS)	1
Alcohol	other	RDLSI ILLEGAL POSSESS OF DRUGS	1
Reckless-	other	RDLSI RECKLESS DRIVING	1
P&M	other	RDLSI UNABLE TO PASS DL TEST	1
Other	other	RDLSI WITHDRAWAL - REASON NOT KNOWN (OBS)	1
Fta_ftp	other	RESTORATION AND/OR SERVICE FEE REQ.	3
NTO	other	STOP ISSUE FOR JUVENILE	8
Alcohol	other	TRANSPORTING OPEN CONTAINER-2ND	2
Alcohol	other	TRANSPORTING OPEN CONTAINER-3RD	1
	other	UNSATISFIED JUDGEMENT	85
	other	UNSATISFIED JUDGEMENT OUT OF STATE	5
Alcohol	other	VIOLATION ABC LAWS (UNDERAGE PURCHASE/ATTEMPT)	12
NTO	other	VIOLATION OF PROBATION	134
Points	points	ACCUMULATION OF 12 POINTS WITHIN A 3 YEAR PERIOD	297
Points	points	ACCUMULATION OF 8 POINTS WITHIN A 3 YEAR PERIOD	125
Speed	speed	OUT-OF-STATE SPEEDING CONVICTION	40
Speed	speed	SPEED OVER 55 & EXCEED LIMIT > 15	190
Reckless	speed	SPEED OVER 55, EXCEEDING 15MPH & RECKLESS DR.	17
Reckless	speed	SPEED>55MPH & RECKLESS/AGGRESSIVE IN 12 MO	70
Speed	speed	SPEEDING IN EXCESS OF 80 MPH	9
Speed	speed	SPEEDING IN EXCESS OF 80 MPH - OOS	40
Speed	speed	SPEEDING IN EXCESS OF 80 MPH-OOS	40
Reckless	speed	SPEEDING OVER 55MPH & RECKLESS DR IN 12 MONTHS 00S	2
Reckless	speed	SPEEDING OVER 55MPH & RECKLESS DR IN 12 MONTHS OOS	13
Reckless	speed	SPEEDING OVER 55MPH & RECKLESS DRIVING IN 12 MON.	61
Speed	speed	SPEEDING OVER 75	11
Speed	speed	SPEEDING OVER 75 MPH	127
Speed	speed	SPEEDING OVER 75 MPH OUT-OF-STATE	157
Serious	speed	SPEEDING TO ELUDE ARREST	29
Speed	speed	TWO CONVIC. OF SPEEDING OVER 55MPH IN 12 MONTHS OOS	9

New group	group	SUS_DS	
Speed	speed	TWO CONVICTIONS SPEEDING OVER 55MPH IN 12 MONTHS	192
Serious	speed	WILLFUL SPEED COMPETITION	6

* P & M = physical and medical

Appendix 2

Detailed Results of Multiple Regression Analyses of Fault in Multi- and Single-Vehicle NC Fatal and A-injury Crashes, 2001-2005, using Induced Exposure Comparison Group

Multi-Vehicle At-Fault (MVAF) Driver Models

Model 1: MVAF crash database-only model, all drivers. In this first multivariate model (Model 1), only the crash database predictors were used to predict which drivers were at fault in the multi-vehicle crashes. This model includes all multi-vehicle crash-involved study drivers from 2001-2005, regardless of whether they were licensed/residents of North Carolina. The intent of including all drivers was to ascertain the contribution of out-of-state drivers to serious crashes in North Carolina. The final backwards logistic analysis results are shown in Table 14.

Table 14. Model 1 logistic regression results predicting Multi-vehicle Driver Fault using *crash database variables only* - all drivers included.

	n_{NAF} = 7,071, n_{MVAF} = 5,412			
Predictor	**β**	**OR**	**95% CI**	
Intercept	0.5145			
DRIVER LICENSE ISSUANCE STATE/COUNTRY IN STUDY CRASH				
NORTH CAROLINA*		1.00		
NEIGHBORING STATE	0.1702	1.19	0.97	1.44
OTHER	0.3562	1.43	1.12	1.82
DRIVER LICENSE & RESIDENCE STATE/COUNTRY COMBINED IN STUDY CRASH				
LICENSED IN RESIDENCE STATE/ COUNTRY*		1.00		
NOT LICENSED IN RESIDENCE STATE	0.2430	1.28	1.03	1.58
DRIVER COMMERCIAL DRIVER LICENSE STATUS IN STUDY CRASH				
NO CDL OR UNKOWN*		1.00		
CDL LICENSE	1.6215	5.06	1.93	13.26
DRIVER AGE GROUP IN STUDY CRASH				
16-19*		1.00		
20-29	-0.8467	0.43	0.38	0.49
30-39	-1.2334	0.29	0.25	0.33
40-49	-1.2951	0.27	0.24	0.31
50-59	-1.3174	0.27	0.23	0.31
60-69	-0.9191	0.40	0.34	0.47
70-79	-0.2475	0.78	0.65	0.94
80+	0.3682	1.45	1.12	1.86
DRIVER RACE				
WHITE*		1.00		
BLACK	-0.0314	0.97	0.89	1.06
NATIVE AMERICAN	-0.0432	0.96	0.69	1.33
HISPANIC	0.5003	1.65	1.43	1.90

	n_{NAF} = 7,071, n_{MVAF} = 5,412			
Predictor	**β**	**OR**	**95% CI**	
ASIAN/OTHER	0.6762	1.97	1.53	2.53
DRIVER BELT USE IN STUDY CRASH				
BELTED, ANY TYPE*		1.00		
NO BELT	0.8301	2.29	2.03	2.60
DRIVER PHYSICAL CONDITION IN STUDY CRASH				
NORMAL*		1.00		
IMPAIRED (ALCOHOL, DRUGS , MEDICATION)	2.5655	13.01	9.10	18.58
SLEEPY/FATIGUED	4.0263	56.05	36.83	85.31
OTHER IMPAIRMENT	2.6048	13.53	8.44	21.69
DRIVER ALCOHOL/DRUG USE IN STUDY CRASH				
NO ALCOHOL OR DRUGS*		1.00		
ALCOHOL ONLY SUSPECTED/DETECTED	0.6530	1.92	1.41	2.62
DRUGS ONLY SUSPECTED/DETECTED	1.0249	2.79	1.37	5.67
BOTH ALCOHOL/DRUGS SUSPECTED/DETECTED	0.6248	1.87	0.93	3.74
MODEL YEAR OF VEHICLE IN STUDY CRASH				
1995 OR NEWER MODEL YEAR*		1.00		
OLDER THAN 1995 MODEL YEAR	0.0822	1.09	1.01	1.17
VEHICLE TYPE IN STUDY CRASH				
CAR*		1.00		
PICKUP	-0.3683	0.69	0.63	0.76
SPORT UTILITY VEHICLE	-0.2479	0.78	0.70	0.87
VAN/MINI VAN	-0.3531	0.70	0.61	0.80
NUMBER OF PASSENGERS IN STUDY CRASH				
NONE*		1.00		
ONE	-0.1160	0.89	0.82	0.97
TWO+	-0.1450	0.87	0.78	0.96
SEASON OF YEAR OF STUDY CRASH				
WINTER*		1.00		
SPRING	0.0537	1.06	0.99	1.13
SUMMER	-0.1582	0.85	0.79	0.92
FALL	0.0196	1.02	0.95	1.10
DAY OF WEEK OF STUDY CRASH				
WEEK DAY (MON-THU)*		1.00		
WEEKEND (FRI-SUN)	0.0527	1.05	1.00	1.11
HOUR (TIME) OF STUDY CRASH				
5:00-8:59 AM*		1.00		
9:00 AM-12:59AM	0.0037	1.00	0.92	1.09

	n_{NAF} = 7,071, n_{MVAF} = 5,412			
Predictor	**β**	**OR**	**95% CI**	
1:00-4:59 PM	-0.1766	0.84	0.77	0.91
5:00-8:59 PM	-0.0931	0.91	0.84	0.99
9:00 PM-12:59PM	-0.4346	0.65	0.58	0.72
1:00-4:59 AM	-0.6586	0.52	0.41	0.66
ROAD CLASS OF STUDY CRASH				
LOCAL ROAD*		1.00		
NC/US/STATE ROAD	-0.4357	0.65	0.58	0.72
INTERSTATE	-0.4469	0.64	0.52	0.78
RURAL ROAD	-0.0081	0.99	0.89	1.10
LOCALITY TYPE OF STUDY CRASH LOCATION				
URBAN*		1.00		
MIXED	0.2683	1.31	1.19	1.44
RURAL	0.1896	1.21	1.08	1.35
DEVELOPMENT TYPE IN LOCATION OF STUDY CRASH				
RESIDENTIAL*		1.00		
COMMERCIAL/ INDUSTRIAL	-0.0383	0.96	0.89	1.04
FARM/UNDEVELOPED	0.0799	1.08	1.00	1.17
SPEED LIMIT OF STUDY CRASH ROAD				
< 55MPH*		1.00		
55MPH+	0.0849	1.09	1.01	1.17

Note. The dependent variable for these analyses coded 0 = not-at-fault driver, 1 = at-fault driver. OR = odds ratio. 95%CI = 95% confidence interval for odds ratio. * = referent category. All predictors retained in the backwards removal model if variable $p < .10$.

The results for Model 1 were surprising in that all the environment-related variables were retained as predictors in the model, even though GEE analysis with an exchangeable correlation structure was used to cluster the drivers by crash. The clustering would be expected to remove confounding on the environment-related variables, but these analyses indicate the presence of residual confounding induced by the 1-to-n clustering that needs to be accounted for in the analyses. That is, even though the GEE analysis method accounts for the correlation among drivers involved in the same crash, it appears that there is residual confounding that needs to be removed by including these environment-related variables as possible predictors in all remaining analyses.

The other results for Model 1 indicate that being at-fault in a multi-vehicle crash was associated with being licensed in a non-neighboring state or jurisdiction, and being licensed in a state that is adjacent to North Carolina. Furthermore, being licensed in a jurisdiction that was not the same as the residence jurisdiction and having a CDL license were also associated with being at fault in the crash. As far as driver-related predictors were concerned, teens ages 16-19 and adults age 80 or older were more likely to be at fault than were drivers of other ages. Interestingly, driver sex was not found to be a meaningful predictor of fault, yet Hispanic and

Asian/Other ethnicities were over-represented among the at-fault drivers. With regard to driver variables measured during the crash that may be predictive of past driving behavior, drivers who were not belted during the study crash, were sleepy/fatigued, had evidence of alcohol or drug use, or who had some other impairment (e.g., illness), were all more likely to be at fault in the multi-vehicle crashes. Further, police reporting of alcohol use alone, drug use alone, or both alcohol and drug use combined, were all also associated with being at fault. Several vehicle variables were also predictive of fault such as having a vehicle older than 1995 model year, driving a car vehicle type (as opposed to a truck, SUV, or van), and driving without any passengers.

The drivers were classified using Model 1 to predict who was actually at fault in the crashes based on both 50% and 75% probability cut-points. Table 15 shows the classification of the drivers by actual fault status and predicted fault status based on the model. The model predictions correctly classified 45.2% of at fault drivers (at the 50% cut point) and 86.6% of not-at-fault drivers. The overall result classified 68.7% of drivers correctly compared with about 51% correct classification with random assignment based on the row totals. Most of the improvement over chance was due to better prediction of not at-fault drivers (53% improved), with only about 4% improvement in classifying the at-fault drivers. In this study, the classification outcomes were used to compare efficiency of the various models developed within this study. If such models were used to predict a separate sample of driver fault status, the cut points could be adjusted depending on whether one desired to capture potentially more risky drivers or reduce the false-positive rate.

Table 15. Model 1 classification outcomes for predicting Multi-vehicle At-Fault Drivers using *crash database variables only*, all drivers included

Actual fault status[a]	Predicted fault status				
	Not-at-fault	Row %	At-fault	Row %	Total
50% probability cut-point					
Not-at-fault	6,123	**86.6**	948	13.4	7,071
At-fault	2,964	54.8	2,448	**45.2**	5,412
Total	9,087	72.8	3,396	27.2	12,483
75% probability cut-point					
Not-at-fault	6,997	**98.9**	74	1.1	7,071
At-fault	4,407	81.4	1,005	**18.6**	5,412
Total	11,404	91.4	1,079	8.6	12,483

Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off).
[a]Based on the presence of one or more driver contributing circumstances for the crash.

Model 2: MVAF Crash Database-Only Model, NC Drivers. This model development followed the same procedures as the first, except only known North Carolina drivers were included so that the licensing database predictors can also be used to predict their culpability in the next step. The purpose of this model was to see whether the predictor variables change when only North Carolina drivers are included in the dataset. The results of the backwards elimination logistic analysis are shown in Table 16.

Table 16. Logistic Regression Results Predicting Multi-Vehicle Driver Fault vs. Not-at-Fault Using Crash Database Variables Only (Model 2), North Carolina Drivers Only

	n_{NAF} = 6,612, n_{MVAF} = 5,018			
Predictor	**β**	**OR**	**95% CI**	
Intercept	0.5308			
DRIVER COMMERCIAL DRIVER LICENSE STATUS IN STUDY CRASH				
NO CDL OR UNKOWN*		1.00		
CDL LICENSE	1.8150	6.14	1.97	19.18
DRIVER AGE GROUP IN STUDY CRASH				
16-19*		1.00		
20-29	-0.8824	0.41	0.36	0.47
30-39	-1.2481	0.29	0.25	0.33
40-49	-1.3211	0.27	0.23	0.31
50-59	-1.3381	0.26	0.22	0.31
60-69	-0.9556	0.38	0.32	0.46
70-79	-0.2639	0.77	0.64	0.93
80+	0.3605	1.43	1.10	1.86
DRIVER RACE				
WHITE*		1.00		
BLACK	-0.0489	0.95	0.87	1.05
NATIVE AMERICAN	-0.0959	0.91	0.65	1.27
HISPANIC	0.5045	1.66	1.43	1.92
ASIAN/OTHER	0.6608	1.94	1.49	2.52
DRIVER BELT USE IN STUDY CRASH				
BELTED, ANY TYPE*		1.00		
NO BELT	0.8569	2.36	2.07	2.68
DRIVER PHYSICAL CONDITION IN STUDY CRASH				
NORMAL*		1.00		
IMPAIRED (ALCOHOL, DRUGS , MEDICATION)	2.5939	13.38	9.19	19.48
SLEEPY/FATIGUED	4.0625	58.12	35.60	94.88
OTHER IMPAIRMENT	2.5704	13.07	8.04	21.24
DRIVER ALCOHOL/DRUG USE IN STUDY CRASH				
NO ALCOHOL OR DRUGS*		1.00		
ALCOHOL ONLY SUSPECTED/DETECTED	0.6458	1.91	1.38	2.63
DRUGS ONLY SUSPECTED/DETECTED	1.0008	2.72	1.29	5.75
BOTH ALCOHOL/DRUGS SUSPECTED/DETECTED	0.8235	2.28	1.09	4.75
MODEL YEAR OF VEHICLE IN STUDY CRASH				
1995 OR NEWER MODEL YEAR*		1.00		

	n_{NAF} = 6,612, n_{MVAF} = 5,018			
Predictor	**β**	**OR**	**95% CI**	
OLDER THAN 1995 MODEL YEAR	0.0746	1.08	1.00	1.16
VEHICLE TYPE IN STUDY CRASH				
CAR*		1.00		
PICKUP	-0.3650	0.69	0.63	0.77
SPORT UTILITY VEHICLE	-0.2374	0.79	0.70	0.89
VAN/MINI VAN	-0.3467	0.71	0.61	0.81
NUMBER OF PASSENGERS IN STUDY CRASH				
NONE*		1.00		
ONE	-0.1238	0.88	0.81	0.97
TWO+	-0.1332	0.88	0.79	0.98
SEASON OF YEAR OF STUDY CRASH				
WINTER*		1.00		
SPRING	0.0609	1.06	0.99	1.14
SUMMER	-0.1710	0.84	0.78	0.91
FALL	0.0315	1.03	0.96	1.11
DAY OF WEEK OF STUDY CRASH				
WEEK DAY (MON-THU)*		1.00		
WEEKEND (FRI-SUN)	0.0539	1.06	1.00	1.11
HOUR (TIME) OF STUDY CRASH				
5:00-8:59 AM*		1.00		
9:00 AM-12:59AM	0.0163	1.02	0.93	1.11
1:00-4:59 PM	-0.1680	0.85	0.78	0.92
5:00-8:59 PM	-0.0635	0.94	0.86	1.02
9:00 PM-12:59PM	-0.4455	0.64	0.57	0.72
1:00-4:59 AM	-0.6237	0.54	0.41	0.70
ROAD CLASS OF STUDY CRASH				
LOCAL ROAD*		1.00		
NC/US/STATE ROAD	-0.4292	0.65	0.59	0.72
INTERSTATE	-0.2265	0.80	0.66	0.96
RURAL ROAD	-0.0170	0.98	0.89	1.09
LOCALITY TYPE OF STUDY CRASH LOCATION				
URBAN*		1.00		
MIXED	0.2661	1.30	1.19	1.43
RURAL	0.1873	1.21	1.08	1.35
DEVELOPMENT TYPE IN LOCATION OF STUDY CRASH				
RESIDENTIAL*		1.00		
COMMERCIAL/ INDUSTRIAL	-0.0356	0.97	0.89	1.04

	n_{NAF} = 6,612, n_{MVAF} = 5,018			
Predictor	**β**	**OR**	**95% CI**	
FARM/UNDEVELOPED	0.0747	1.08	0.99	1.17
SPEED LIMIT OF STUDY CRASH ROAD				
< 55MPH*		1.00		
55MPH+	0.0787	1.08	1.01	1.16

Note. The dependent variable for these analyses coded 0 = not-at-fault driver, 1 = at-fault driver. OR = odds ratio. 95%CI = 95% confidence interval for odds ratio. * = referent category. All predictors retained in the backwards removal model if variable $p < .10$.

Predictably given that most of the data consisted of North Carolina drivers, the results for Model 2 were very similar to those for Model 1. Two fewer predictor variables were retained in the model, both of which likely discriminated out-of-state drivers in Model 1. Specifically, the license jurisdiction variable was dropped, as was the created variable indicating whether license state and residence state were the same. All the other predictor variables were retained and had the essentially the same relationship to culpability as for Model 1. The classification of driver fault based on Model 2 was the highly consistent with the prior model (Table 17). The level of classification again suggests the need for additional predictor variables for the model to improve classification accuracy of multi-vehicle at-fault drivers above 45%.

Table 17. Classification Outcomes for Predicting Multi-Vehicle At-Fault Drivers Using <u>Crash Database Variables Only (Model 2)</u> with 50% & 75% Probability Cut-Offs, North Carolina Drivers Only

	Predicted fault status				
Actual fault status[a]	**Not-at-fault**	**Row %**	**At-fault**	**Row %**	**Total**
50% probability cut-point					
Not-at-fault	5,731	**86.7**	881	13.3	6,612
At-fault	2,736	54.5	2,282	**45.5**	5,018
Total	8,467	72.8	3,163	27.2	11,630
75% probability cut-point					
Not-at-fault	6,542	**98.9**	70	1.1	6,612
At-fault	4,091	81.5	927	**18.5**	5,018
Total	10,633	91.4	997	8.6	11,630

Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off). The total counts differ across models because the inclusion or exclusion of variables with missing values changes the number of observations in the models.
[a]Based on the presence of one or more driver contributing circumstances for the crash.

Model 3: MVAF Licensing Database-Only Model, NC Drivers. This model builds upon Model 2 in that only North Carolina drivers are included as observations. However, instead of using solely crash database variables, this model considered all the licensing database variables as potential predictors. Only four of the crash database variables were considered as predictors, because these variables were also available in the licensing data. These four variables are (a) CDL license status, (b) driver sex, (c) driver age, and (d) driver race.

Recall that the licensing predictors are composed of both individual types of restrictions, convictions, and suspensions, along with summary measures of the total numbers of these histories in the 5-year period prior to the crash. Hence, it was not possible to simply include all these variables at one time in a single model because of problems with part-to-whole singularity and multicollinearity. For example, the total number of convictions is completely predicted by the sum over counts of the individual types of convictions, so entering these parts and the whole count in the model would result in problems inverting the variance-covariance matrix as required for the analysis. Therefore, a somewhat modified procedure was used for the analyses including the licensing predictors.

First, all the crash database predictors and the individual-type restriction, conviction, and suspension variables were entered into the model, excluding the cumulative summary variables of the licensing history. Next, individual predictors that did not meet the < .10 backwards elimination criterion were eliminated from the model and it was re-estimated. Finally, the summary measures were added to the models to determine whether they were able to predict any additional variability in culpability. In this final step variables that did not meet the significance criterion were again removed from the model. The final backwards elimination logistic analysis results are shown in Table 18.

Table 18. Logistic Regression Results Predicting Multi-Vehicle Driver Fault vs. Not-at-Fault Using Licensing Database Variables Only (Model 3), North Carolina Drivers Only

	n_{NAF} = 7,074, n_{MVAF} = 6,086			
Predictor	**β**	**OR**	**95% CI**	
Intercept	0.9793			
DRIVER COMMERCIAL DRIVER LICENSE STATUS IN STUDY CRASH				
NO CDL OR UNKOWN*		1.00		
CDL LICENSE	1.1374	3.12	0.95	10.19
DRIVER AGE GROUP IN STUDY CRASH				
16-19*		1.00		
20-29	-0.8147	0.44	0.38	0.51
30-39	-1.1189	0.33	0.28	0.38
40-49	-1.0622	0.35	0.30	0.40
50-59	-1.1359	0.32	0.27	0.38
60-69	-0.8495	0.43	0.36	0.51
70-79	0.0867	1.09	0.90	1.32
80+	0.7829	2.19	1.70	2.81
DRIVER RACE				
WHITE*		1.00		
BLACK	0.0261	1.03	0.94	1.12
NATIVE AMERICAN	0.3151	1.37	1.05	1.79
HISPANIC	0.5763	1.78	1.53	2.06

	n_{NAF} = 7,074, n_{MVAF} = 6,086			
Predictor	**β**	**OR**	**95% CI**	
ASIAN/OTHER	0.6777	1.97	1.54	2.52
LICENSE STATUS AT STUDY CRASH				
FULLY NC LICENSED*		1.00		
SUSPENDED/LIMITED PRIVILEGE NC LICENSE	1.3357	3.80	2.94	4.91
NC UNLICENSED	0.2487	1.28	1.10	1.50
OVERALL NUMBER OF LICENSE RESTRICTIONS IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	-0.2359	0.79	0.69	0.90
TWO+	-1.4163	0.24	0.12	0.50
45MPH/NO INTERSTATE RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	1.0392	2.83	1.12	7.12
DAYLIGHT DRIVING ONLY RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	1.4585	4.30	1.61	11.45
ONLY DRIVE WHILE SUPERVISED RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	1.0425	2.84	1.54	5.23
AUTOMATIC TRANSMISSION RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	1.2486	3.49	0.80	15.25
GRADUATED DRIVER LICENSING RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	1.2510	3.49	2.34	5.21
BLOOD ALCOHOL CONTENT < .04 RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	0.8870	2.43	1.22	4.83
NUMBER OF NOT-AT-FAULT CRASHES IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	-0.6504	0.52	0.36	0.75
TWO	-0.6977	0.50	0.33	0.76
THREE+	-0.1631	0.85	0.58	1.25
NUMBER OF AT-FAULT CRASHES IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	-0.0303	0.97	0.74	1.28
TWO	1.0821	2.95	1.87	4.65
THREE+	1.8259	6.21	3.24	11.90

	n_{NAF} = 7,074, n_{MVAF} = 6,086			
Predictor	**β**	**OR**	**95% CI**	
AT-FAULT 2ND MOST RECENT CRASH INVOLVEMENT IN PRIOR 5 YEARS				
NO/NO CRASH*		1.00		
YES	-0.4648	0.63	0.43	0.93
AGE DIFFERENCE TO MOST RECENT PRIOR CRASH INVOLVEMENT IN LAST 5 YEARS				
SAME AGE*		1.00		
ONE YEAR YOUNGER	-0.5227	0.59	0.46	0.77
TWO YEARS YOUNGER	-0.7837	0.46	0.33	0.64
THREE YEARS YOUNGER	0.0615	1.06	0.86	1.31
FOUR YEARS YOUNGER	-0.3801	0.68	0.51	0.91
FIVE YEARS YOUNGER	-0.2808	0.76	0.56	1.02
NO CRASH PRIOR 5 YEARS	-0.6119	0.54	0.36	0.82
NUMBER OF SPEED-RELATED CONVICTIONS IN PRIOR 5 YEARS				
NO CONVICTIONS*		1.00		
ONE	0.1197	1.13	1.02	1.25
TWO+	0.1349	1.14	0.99	1.32
NUMBER OF OTHER MOVING CONVICTIONS IN PRIOR 5 YEARS				
NO CONVICTIONS*		1.00		
ONE	-0.3583	0.70	0.55	0.89
TWO+	0.5408	1.72	1.17	2.53
OVERALL NUMBER OF LICENSE SUSPENSIONS IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	0.3467	1.41	1.19	1.69
TWO	0.0837	1.09	0.76	1.56
THREE	0.5246	1.69	1.15	2.48
FOUR	0.5652	1.76	0.68	4.55
FAILURE TO APPEAR/ FAILURE TO PAY LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-0.3724	0.69	0.57	0.84

Note. The dependent variable for these analyses coded 0 = not-at-fault driver, 1 = at-fault driver. OR = odds ratio. 95%CI = 95% confidence interval for odds ratio. * = referent category. All predictors retained in the backwards removal model if variable $p < .10$.

The final licensing database predictor model included more variables than the crash variable models, with a total of 19 licensing predictors being retained in Model 3. The results indicated that multi-vehicle at-fault drivers were over-represented among drivers: (a) with a CDL license; (b) ages 16-19 or 80+ years-old; (c) who were Native American, Hispanic, or

Asian/Other ethnicities; (d) who had a suspended/limited privilege or no valid driver license; (e) with no license restrictions on record in the prior 5 years (versus one or two+ restrictions of any type); (f) with 45MPH/no interstate, daylight driving only, supervised driving only, automatic transmission only, graduated driver licensing, or BAC < .04 restrictions on record in the prior five years; (g) with *no* prior not-at-fault crashes in the prior five years (versus one or two); (h) with two or three+ prior at fault crashes in the prior five years; (i) who were not-at-fault in the 2nd most recent prior crash in the previous five years (versus being at-fault or not crashing twice in the prior five years); (j) who were the same age when involved in their most recent crash in the prior five years (vs. one or more years younger, with one exception, or having had no crash in the prior five years); (k) with one or two+ speed-related convictions in the prior five years; (l) with two+ "other" moving convictions in the prior five years; (m) with one or three+ suspensions of any type on record in the prior five years (versus no suspensions); and (n) who were *not* suspended for failure to appear in court or failure to pay a fine in the prior five years.

The classification of at-fault and not-at-fault multi-vehicle crash drivers based on Model 3 resulted in poorer classification for at-fault drivers (sensitivity) and not-at-fault drivers (specificity; Table 19). Using on the 50% cut point, 44% of at-fault drivers were correctly classified by the licensing predictors in Model 3 versus 45% using the crash database predictors (Model 2). The classification of not-at-fault drivers was also worse based on the licensing predictors with only 80% of such drivers correctly classified by Model 3 (vs. 87% in Model 2).

Table 19. Classification Outcomes for Predicting Multi-Vehicle At-Fault Drivers Using License Database Variables Only (Model 3) with 50% & 75% Probability Cut-Offs, North Carolina Drivers

Actual fault status[a]	Predicted fault status				
	Not-at-fault	Row %	At-fault	Row %	Total
50% probability cut-point					
Not-at-fault	5,685	**80.4**	1,389	19.6	7,074
At-fault	3,412	56.1	2,674	**43.9**	6,086
Total	9,097	69.1	4,063	30.9	13,160
75% probability cut-point					
Not-at-fault	6,909	**97.7**	165	2.3	7,074
At-fault	5,523	90.7	563	**9.3**	6,086
Total	12,432	94.5	728	5.5	13,160

Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off). The total counts differ across models because the inclusion or exclusion of variables with missing values changes the number of observations in the models.
[a]Based on the presence of one or more driver contributing circumstances for the crash.

Model 4: MVAF Crash & Licensing Database Model, NC Drivers. This model is a combination of the prior two models in that both the crash *and* licensing database predictors were simultaneously considered as potential predictors of culpability. The environment-related variables were also entered for consideration in the first step due to the fact that the clustering did not appear to remove all confounding introduced by the induced exposure method. To avoid problems with multicollinearity among the licensing database predictors, the overall totals of convictions, suspensions, and restrictions were considered in a final step. All variables with an omnibus Wald significance value less than .10 were retained in the final model. The final logistic analysis results are shown in Table 20.

Table 20. Logistic Regression Results Predicting Multi-Vehicle Driver Fault vs. Not-at-Fault Using Crash & Licensing Database Variables (Model 4), North Carolina Drivers Only

	n_{NAF} = 6,655, n_{MVAF} = 5,069			
Predictor	**β**	**OR**	**95% CI**	
Intercept	0.2693			
DRIVER COMMERCIAL DRIVER LICENSE STATUS IN STUDY CRASH				
NO CDL OR UNKOWN*		1.00		
CDL LICENSE	1.5724	4.82	1.41	16.42
DRIVER AGE GROUP IN STUDY CRASH				
16-19*		1.00		
20-29	-0.7818	0.46	0.40	0.53
30-39	-1.0916	0.34	0.29	0.39
40-49	-1.1541	0.32	0.27	0.37
50-59	-1.1294	0.32	0.27	0.38
60-69	-0.7198	0.49	0.41	0.58
70-79	-0.0975	0.91	0.75	1.10
80+	0.5541	1.74	1.34	2.26
DRIVER RACE				
WHITE*		1.00		
BLACK	-0.0817	0.92	0.84	1.01
NATIVE AMERICAN	-0.1477	0.86	0.62	1.21
HISPANIC	0.4342	1.54	1.33	1.79
ASIAN/OTHER	0.6530	1.92	1.48	2.49
DRIVER BELT USE IN STUDY CRASH				
BELTED, ANY TYPE*		1.00		
NO BELT	0.7861	2.19	1.93	2.50
DRIVER PHYSICAL CONDITION IN STUDY CRASH				
NORMAL*		1.00		
IMPAIRED (ALCOHOL, DRUGS , MEDICATION)	2.5376	12.65	8.74	18.30
SLEEPY/FATIGUED	4.0065	54.95	33.65	89.75
OTHER IMPAIRMENT	2.4090	11.12	6.94	17.82
DRIVER ALCOHOL/DRUG USE IN STUDY CRASH				
NO ALCOHOL OR DRUGS*		1.00		
ALCOHOL ONLY SUSPECTED/DETECTED	0.4899	1.63	1.20	2.22
DRUGS ONLY SUSPECTED/DETECTED	0.8710	2.39	1.14	5.02
BOTH ALCOHOL/DRUGS SUSPECTED/DETECTED	0.7265	2.07	1.01	4.25

	n_{NAF} = 6,655, n_{MVAF} = 5,069			
Predictor	**β**	**OR**	**95% CI**	
MODEL YEAR OF VEHICLE IN STUDY CRASH				
1995 OR NEWER MODEL YEAR*		1.00		
OLDER THAN 1995 MODEL YEAR	0.0640	1.07	0.99	1.15
VEHICLE TYPE IN STUDY CRASH				
CAR*		1.00		
PICKUP	-0.3485	0.71	0.64	0.78
SPORT UTILITY VEHICLE	-0.2452	0.78	0.70	0.88
VAN/MINI VAN	-0.3311	0.72	0.63	0.82
NUMBER OF PASSENGERS IN STUDY CRASH				
NONE*		1.00		
ONE	-0.1403	0.87	0.80	0.95
TWO+	-0.1096	0.90	0.81	1.00
SEASON OF YEAR OF STUDY CRASH				
WINTER*		1.00		
SPRING	0.0721	1.07	1.00	1.15
SUMMER	-0.0718	0.93	0.87	1.00
FALL	0.0260	1.03	0.95	1.10
HOUR (TIME) OF STUDY CRASH				
5:00-8:59 AM*		1.00		
9:00 AM-12:59AM	0.0130	1.01	0.93	1.11
1:00-4:59 PM	-0.1167	0.89	0.82	0.97
5:00-8:59 PM	-0.0573	0.94	0.87	1.03
9:00 PM-12:59PM	-0.4500	0.64	0.57	0.72
1:00-4:59 AM	-0.6725	0.51	0.39	0.66
ROAD CLASS OF STUDY CRASH				
LOCAL ROAD*		1.00		
NC/US/STATE ROAD	-0.2802	0.76	0.69	0.82
INTERSTATE	-0.1253	0.88	0.74	1.06
RURAL ROAD	0.1070	1.11	1.01	1.22
LOCALITY TYPE OF STUDY CRASH LOCATION				
URBAN*		1.00		
MIXED	0.1768	1.19	1.10	1.30
RURAL	0.1100	1.12	1.01	1.24
DEVELOPMENT TYPE IN LOCATION OF STUDY CRASH				
RESIDENTIAL*		1.00		
COMMERCIAL/ INDUSTRIAL	0.0129	1.01	0.94	1.09
FARM/UNDEVELOPED	0.0948	1.10	1.01	1.20
LICENSE STATUS AT STUDY CRASH				

	n_{NAF} = 6,655, n_{MVAF} = 5,069			
Predictor	**β**	**OR**	**95% CI**	
FULLY NC LICENSED*		1.00		
SUSPENDED/LIMITED PRIVILEGE NC LICENSE	0.6062	1.83	1.52	2.22
NC UNLICENSED	0.2896	1.34	1.15	1.55
DRIVE CLASS B AND C ONLY RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-0.9942	0.37	0.13	1.03
GRADUATED DRIVER LICENSING RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	0.7500	2.12	1.62	2.77
NUMBER OF NOT-AT-FAULT CRASHES IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	-0.1246	0.88	0.80	0.97
TWO	-0.1393	0.87	0.72	1.06
THREE+	0.1981	1.22	0.88	1.70
NUMBER OF AT-FAULT CRASHES IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	0.1963	1.22	1.09	1.36
TWO	0.4081	1.50	1.14	1.99
THREE+	1.1694	3.22	1.71	6.08
NUMBER OF LICENSE DEMERIT POINTS IN PRIOR 5 YEARS				
NONE*		1.00		
1-2 POINTS	0.0745	1.08	0.94	1.24
3-4 POINTS	0.2624	1.30	1.15	1.47
5-6 POINTS	0.2125	1.24	1.02	1.49
7-8 POINTS	-0.2597	0.77	0.57	1.05
9+ POINTS	0.2006	1.22	0.93	1.61
NUMBER OF NON-MOVING CONVICTIONS IN PRIOR 5 YEARS				
NO CONVICTIONS*		1.00		
ONE	-0.0331	0.97	0.74	1.26
TWO+	0.7236	2.06	1.12	3.79
NUMBER OF ALCOHOL-RELATED CONVICTIONS IN PRIOR 5 YEARS				
NO CONVICTIONS		1.00		
ONE	-0.2944	0.74	0.51	1.09
TWO+	0.5091	1.66	1.31	2.11
DRIVING WHILE SUSPENDED LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-0.4534	0.64	0.48	0.84
GRADUATED DRIVER LICENSING-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				

	n_{NAF} = 6,655, n_{MVAF} = 5,069			
Predictor	**β**	**OR**	**95% CI**	
NO*		1.00		
YES	-0.5394	0.58	0.31	1.09

Note. The dependent variable for these analyses coded 0 = not-at-fault driver, 1 = at-fault driver. OR = odds ratio. 95%CI = 95% confidence interval for odds ratio. * = referent category. All predictors retained in the backwards removal model if variable $p < .10$.

When both crash and licensing database predictors were considered for inclusion in Model 4, only 10 licensing variables and 14 crash variables (24 total) were found to meet the retention criterion. In addition, two environment-related predictors no longer met the retention criterion—day of week and roadway speed limit of the study crash. Overall, the relation of the crash licensing database variables to driver fault did not change substantively. Multi-vehicle at-fault drivers were over-represented among drivers: (a) with a CDL license; (b) ages 16-19 or 80+ years-old; (c) of Hispanic or Asian/Other ethnicity; (d) not belted during the study crash; (e) impaired by alcohol, drugs, or medication, sleep/fatigue, or some other cause (e.g., being ill or missing a limb) during the study crash; (f) with suspected/evidence of alcohol alone, drugs alone, or both alcohol and drugs during the study crash; (g) driving a model year vehicle older than 1995 in the study crash; (h) driving a car (as opposed to a truck, van, or SUV); and (i) driving without any passengers.

In terms of the license database predictors, the results of Model 4 indicated that at-fault drivers were over-represented among drivers: (a) with a suspended/limited privilege license or who were unlicensed; (b) who did *not* have a Class B and C license restriction; (c) who had a GDL license restriction; (d) who were *not* involved in any not-at-fault crashes in the prior five years (vs. being involved in only one or two); (e) who were at-fault for 1, 2, or 3+ crashes in the prior 5-years; (f) who had 3-4, 5-6, or 9+ license demerit points in the prior five years; (g) who had 2 or more *non*-moving convictions in the prior five years; (h) who had two or more alcohol-related convictions in the prior five years; and inversely associated with (i) driving while suspended and (j) GDL-related license suspensions.

The classification of at-fault and not-at-fault multi-vehicle crash drivers based on Model 4 was disappointing in that the addition of the 10 license database variables did not result in a meaningful improvement in the classification of either at-fault or not-at-fault drivers (Table 21). Based on the 50% cut point, only 47% of at-fault drivers were correctly classified in Model 4 versus 45% using only the crash database predictors (Model 2). Further, the classification of not-at-fault drivers only negligibly increased, rounding to about 87% in both models.

Table 21. Classification Outcomes for Predicting Multi-Vehicle At-Fault Drivers Using Crash & Licensing Database Variables (Model 4) with 50% & 75% Probability Cut-Offs, North Carolina Drivers

Actual fault status[a]	Predicted fault status				
	Not-at-fault	Row %	At-fault	Row %	Total
50% probability cut-point					
Not-at-fault	5,767	**86.7**	888	13.3	6,655
At-fault	2,669	52.6	2,400	**47.3**	5,069
Total	8,436	72.0	3,288	28.0	11,724
75% probability cut-point					
Not-at-fault	6,572	**98.8**	83	1.2	6,655
At-fault	4,043	79.8	1,026	**20.2**	5,069
Total	10,615	90.5	1,109	9.5	11,724

Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off). The total counts differ across models because the inclusion or exclusion of variables with missing values changes the number of observations in the models.
[a]Based on the presence of one or more driver contributing circumstances for the crash.

Single-Vehicle At-Fault (SVAF) Driver Models

Model 5: SVAF Crash Database-Only Model, All Drivers. This first model for the single-vehicle at-fault drivers is similar to Model 1 for the multi-vehicle at-fault drivers in that out-of-state and unknown licensing status drivers are also included as observations. In addition, only crash database predictors are considered for inclusion in the model. The comparisons are between drivers who were at-fault in a crash involving only their own vehicle (i.e., a single-vehicle crash) and drivers who were not-at-fault in a multi-vehicle crash and hence should represent the distribution of drivers on North Carolina roadways in general according to the logic behind induced exposure methodology. In addition, standard logistic regression (without GEEs and robust variance estimators) was used since the two groups of drivers are obviously not clustered by crash. The results of the final backwards elimination model are shown in Table 22.

Table 22. Logistic Regression Results Predicting Single-Vehicle Driver Fault vs. Multi-Vehicle Not-at-Fault Using Crash Database Variables Only (Model 5), All Drivers Included

	n_{NAF} = 7,106, n_{SVAF} = 5,648			
Predictor	**β**	**OR**	**95% CI**	
Intercept	-0.4401			
DRIVER AGE GROUP IN STUDY CRASH				
16-19*		1.00		
20-29	-1.1281	0.32	0.27	0.38
30-39	-1.6866	0.19	0.15	0.22
40-49	-1.7575	0.17	0.14	0.21
50-59	-1.7983	0.17	0.13	0.20
60-69	-2.0376	0.13	0.10	0.17
70-79	-1.6390	0.19	0.14	0.27
80+	-1.3247	0.27	0.16	0.45
DRIVER RACE				
WHITE*		1.00		
BLACK	0.1501	1.16	1.01	1.33
NATIVE AMERICAN	-0.3262	0.72	0.45	1.17
HISPANIC	0.3250	1.38	1.11	1.72
ASIAN/OTHER	0.5133	1.67	1.09	2.56
DRIVER BELT USE IN STUDY CRASH				
BELTED, ANY TYPE*		1.00		
NO BELT	2.0387	7.68	6.62	8.91
DRIVER PHYSICAL CONDITION IN STUDY CRASH				
NORMAL*		1.00		
IMPAIRED (ALCOHOL, DRUGS , MEDICATION)	3.3183	27.61	16.04	47.55
SLEEPY/FATIGUED	7.0386	1139.70	157.99	8222.30
OTHER IMPAIRMENT	4.8972	133.92	64.06	279.96
DRIVER ALCOHOL/DRUG USE IN STUDY CRASH				
NO ALCOHOL OR DRUGS*		1.00		
ALCOHOL ONLY SUSPECTED/DETECTED	1.7436	5.72	3.88	8.42
DRUGS ONLY SUSPECTED/DETECTED	1.7837	5.95	1.54	23.00
BOTH ALCOHOL/DRUGS SUSPECTED/DETECTED	1.1905	3.29	1.09	9.94
VEHICLE TYPE IN STUDY CRASH				
CAR*		1.00		
PICKUP	-0.4396	0.64	0.56	0.74
SPORT UTILITY VEHICLE	0.1235	1.13	0.98	1.31
VAN/MINI VAN	-0.6149	0.54	0.44	0.67

	n_{NAF} = 7,106, n_{SVAF} = 5,648			
Predictor	**β**	**OR**	**95% CI**	
VEHICLE INSURED IN STUDY CRASH				
INSURED VEHICLE*		1.00		
NOT INSURED	0.8775	2.40	1.60	3.62
NUMBER OF PASSENGERS IN STUDY CRASH				
NONE*		1.00		
ONE	-0.0819	0.92	0.81	1.05
TWO+	0.1131	1.12	0.97	1.29
HOUR (TIME) OF STUDY CRASH				
5:00-8:59 AM*		1.00		
9:00 AM-12:59AM	-0.1904	0.83	0.69	0.99
1:00-4:59 PM	-0.3836	0.68	0.58	0.81
5:00-8:59 PM	-0.3442	0.71	0.60	0.84
9:00 PM-12:59PM	0.2720	1.31	1.08	1.59
1:00-4:59 AM	1.1943	3.30	2.51	4.34
ROAD CLASS OF STUDY CRASH				
LOCAL ROAD*		1.00		
NC/US/STATE ROAD	-0.2698	0.76	0.61	0.95
INTERSTATE	0.9824	2.67	2.02	3.53
RURAL ROAD	0.8293	2.29	1.85	2.84
LOCALITY TYPE OF STUDY CRASH LOCATION				
URBAN*		1.00		
MIXED	0.2387	1.27	1.03	1.57
RURAL	0.4926	1.64	1.29	2.08
DEVELOPMENT TYPE IN LOCATION OF STUDY CRASH				
RESIDENTIAL*		1.00		
COMMERCIAL/ INDUSTRIAL	-1.1138	0.33	0.27	0.40
FARM/UNDEVELOPED	0.0991	1.10	0.92	1.32
SPEED LIMIT OF STUDY CRASH ROAD				
< 55MPH*		1.00		
55MPH+	0.3675	1.44	1.25	1.66

Note. The dependent variable for these analyses coded 0 = not-at-fault driver, 1 = at-fault driver. OR = odds ratio. 95%CI = 95% confidence interval for odds ratio. * = referent category. All predictors retained in the backwards removal model if variable $p < .10$.

The results of for Model 5 indicated that being at fault in a single-vehicle crash was associated very strongly with being 16-19-years-old, and also with being Black, Hispanic, or Asian/Other ethnicity. Again, it was surprising that driver sex was not found to be significantly predictive of fault. At the time of the study crash, single-vehicle at-fault drivers were more likely to be unbelted, and much more likely to be impaired by alcohol, drugs, or medication, by sleep/fatigue (in particular), or by some other cause such as an illness. The police were also more likely to report that alcohol alone, drugs alone, or both alcohol and drugs were detected/ suspected for single-vehicle at fault drivers. In terms of vehicle-related factors, the at-fault drivers were over-represented in cars and SUVs, uninsured vehicles, and when driving with 2 or more passengers.

The environment-related factors can also be meaningfully interpreted with regard to the single-vehicle at-fault drivers. Specifically, their crashes were over-represented during the hours of 9:00 pm to 1:00 a.m. and 1:00 to 5:00 a.m. While they were more likely to occur on interstates and rural roadways, they were less likely to occur on North Carolina, State, or U.S. roadways. Single-vehicle at-fault crashes were more likely to happen in rural or mixed rural/urban localities, and in residential or farm/undeveloped areas (vs. commercial/industrial areas). Finally, their crashes were also over-represented on roadways with 55MPH or higher speed limits.

The classification of the single-vehicle at-fault and multi-vehicle not-at-fault drivers based on Model 5 was much more impressive than for the multi-vehicle at-fault drivers. Based solely on the 13 crash database predictor variables, the model correctly classified 76% of the at-fault drivers using the 50% probability cut-point and correctly classified 62% using the more conservative 75% cut-point (Table 23). The model also did a good job at correctly classifying the not-at-fault drivers, with over 90% correctly classified using either cut-point. The fact that several of the variables, particularly age and driver physical condition, were *very* strongly associated with being a single-vehicle at-fault driver helps to explain the superior classification of this model compared to the multi-vehicle at-fault driver models.

Table 23. Classification Outcomes for Predicting Single-Vehicle At-Fault Drivers Using Crash Database Variables Only (Model 5) with 50% & 75% Probability Cut-Offs, All Drivers Included

Actual fault status[a]	Predicted fault status				
	Not-at-fault	Row %	At-fault	Row %	Total
50% probability cut-point					
Not-at-fault	6,479	**91.2**	627	8.8	7,106
At-fault	1,357	24.0	4,291	**76.0**	5,648
Total	7,836	61.4	4,918	38.6	12,754
75% probability cut-point					
Not-at-fault	6,925	**97.5**	181	2.5	7,106
At-fault	2,120	37.5	3,528	**62.5**	5,648
Total	9,045	70.9	3,709	29.1	12,754

Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off).
[a]Based on the presence of one or more driver contributing circumstances for the crash.

Model 6: SVAF Crash Database-Only Model, NC Drivers. Similar to the procedure used for multi-vehicle at-fault drivers, this model is a replication of the prior model, but excludes out-of-state drivers and those who could not be matched to the driver license database. The final backwards logistic regression results are shown in Table 24.

Table 24. Logistic Regression Results Predicting Single-Vehicle Driver Fault vs. Multi-Vehicle Not-at-Fault Using Crash Database Variables Only (Model 6), North Carolina Drivers Only

	n_{NAF} = 6,641, n_{SVAF} = 5,226			
Predictor	**β**	**OR**	**95% CI**	
Intercept	-0.4056			
DRIVER AGE GROUP IN STUDY CRASH				
16-19*		1.00		
20-29	-1.1432	0.32	0.27	0.38
30-39	-1.6938	0.18	0.15	0.22
40-49	-1.7690	0.17	0.14	0.21
50-59	-1.8524	0.16	0.13	0.19
60-69	-2.1000	0.12	0.09	0.16
70-79	-1.6755	0.19	0.13	0.27
80+	-1.3079	0.27	0.16	0.47
DRIVER RACE				
WHITE*		1.00		
BLACK	0.1085	1.11	0.97	1.29
NATIVE AMERICAN	-0.3398	0.71	0.44	1.16
HISPANIC	0.3151	1.37	1.10	1.71
ASIAN/OTHER	0.4013	1.49	0.94	2.38
DRIVER BELT USE IN STUDY CRASH				
BELTED, ANY TYPE*		1.00		
NO BELT	2.0596	7.84	6.72	9.15
DRIVER PHYSICAL CONDITION IN STUDY CRASH				
NORMAL*		1.00		
IMPAIRED (ALCOHOL, DRUGS , MEDICATION)	3.2751	26.45	15.21	45.97
SLEEPY/FATIGUED	6.9469	1039.90	143.67	7527.20
OTHER IMPAIRMENT	4.8476	127.43	60.66	267.71
DRIVER ALCOHOL/DRUG USE IN STUDY CRASH				
NO ALCOHOL OR DRUGS*		1.00		
ALCOHOL ONLY SUSPECTED/DETECTED	1.6842	5.39	3.63	8.00
DRUGS ONLY SUSPECTED/DETECTED	1.6758	5.34	1.35	21.08
BOTH ALCOHOL/DRUGS SUSPECTED/DETECTED	1.3710	3.94	1.17	13.21
VEHICLE TYPE IN STUDY CRASH				
CAR*		1.00		
PICKUP	-0.4381	0.65	0.56	0.75
SPORT UTILITY VEHICLE	0.1121	1.12	0.96	1.30
VAN/MINI VAN	-0.6347	0.53	0.42	0.67
VEHICLE INSURED IN STUDY CRASH				
INSURED VEHICLE*		1.00		

	n_{NAF} = 6,641, n_{SVAF} = 5,226			
Predictor	**β**	**OR**	**95% CI**	
NOT INSURED	1.1043	3.02	1.87	4.87
NUMBER OF PASSENGERS IN STUDY CRASH				
NONE*		1.00		
ONE	-0.1109	0.90	0.78	1.02
TWO+	0.0856	1.09	0.94	1.26
HOUR (TIME) OF STUDY CRASH				
5:00-8:59 AM*		1.00		
9:00 AM-12:59AM	-0.1795	0.84	0.69	1.01
1:00-4:59 PM	-0.3791	0.68	0.58	0.81
5:00-8:59 PM	-0.3155	0.73	0.61	0.87
9:00 PM-12:59PM	0.3052	1.36	1.11	1.66
1:00-4:59 AM	1.2980	3.66	2.74	4.89
ROAD CLASS OF STUDY CRASH				
LOCAL ROAD*		1.00		
NC/US/STATE ROAD	-0.2350	0.79	0.63	0.99
INTERSTATE	0.9055	2.47	1.82	3.37
RURAL ROAD	0.8543	2.35	1.89	2.93
LOCALITY TYPE OF STUDY CRASH LOCATION				
URBAN*		1.00		
MIXED	0.2246	1.25	1.01	1.55
RURAL	0.4328	1.54	1.21	1.97
DEVELOPMENT TYPE IN LOCATION OF STUDY CRASH				
RESIDENTIAL*		1.00		
COMMERCIAL/ INDUSTRIAL	-1.1296	0.32	0.27	0.39
FARM/UNDEVELOPED	0.1089	1.12	0.93	1.34
SPEED LIMIT OF STUDY CRASH ROAD				
< 55MPH*		1.00		
55MPH+	0.3710	1.45	1.25	1.67

Note. The dependent variable for these analyses coded 0 = not-at-fault driver, 1 = at-fault driver. OR = odds ratio. 95%CI = 95% confidence interval for odds ratio. * = referent category. All predictors retained in the backwards removal model if variable $p < .10$.

Though there were some slight changes in coefficient values, the results for Model 6 were entirely consistent with those for Model 5. That is, with the out-of-state and non-matching drivers removed from the data, the final logistic model contained the same set of crash database predictors and the relations to culpability were almost exactly the same. The classification results (Table 25) for Model 6 were also highly consistent with those from the prior model, again indicating much better classification than was the case for the multi-vehicle at-fault drivers.

Table 25. Classification Outcomes for Predicting Single-Vehicle At-Fault Drivers Using Crash Database Variables Only (Model 6) with 50% & 75% Probability Cut-Offs, North Carolina Drivers Only

Actual fault status[a]	Predicted fault status				
	Not-at-fault	Row %	At-fault	Row %	Total
50% probability cut-point					
Not-at-fault	6,061	**91.3**	580	8.7	6,641
At-fault	1,253	24.0	3,973	**76.0**	5,226
Total	7,314	61.6	4,553	38.4	11,867
75% probability cut-point					
Not-at-fault	6,480	**97.6**	161	2.4	6,641
At-fault	1,960	37.5	3,266	**62.5**	5,226
Total	8,440	71.1	3,427	28.9	11,867

Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off). The total counts differ across models because the inclusion or exclusion of variables with missing values changes the number of observations in the models.
[a]Based on the presence of one or more driver contributing circumstances for the crash.

Model 7: SVAF Licensing Database-Only Model, NC Drivers. This model builds upon the prior model in that only North Carolina drivers are included as observations, but the licensing database variables were considered as potential predictors instead. However, four of the crash database variables were also considered as predictors because they were also available in the licensing database: (a) CDL license status, (b) driver sex, (c) driver age, and (d) driver race. The same procedures used for Model 3 were repeated here to account for multicollinearity of the licensing database variables. The final backwards elimination logistic analysis results are shown in Table 26.

Table 26. Logistic Regression Results Predicting Single-Vehicle Driver Fault vs. Not-at-Fault Using Licensing Database Variables Only (Model 7), North Carolina Drivers Only

	n_{NAF} = 7,074, n_{SVAF} = 7,041			
Predictor	**β**	**OR**	**95% CI**	
Intercept	0.5228			
DRIVER AGE GROUP IN STUDY CRASH				
16-19*		1.00		
20-29	-0.8342	0.43	0.38	0.50
30-39	-1.2702	0.28	0.24	0.32
40-49	-1.3162	0.27	0.23	0.31
50-59	-1.5055	0.22	0.19	0.26
60-69	-1.6184	0.20	0.16	0.24
70-79	-1.2750	0.28	0.22	0.36
80+	-0.8540	0.43	0.30	0.61
DRIVER SEX				
FEMALE*		1.00		

	n_{NAF} = 7,074, n_{SVAF} = 7,041			
Predictor	**β**	**OR**	**95% CI**	
MALE	0.3460	1.41	1.31	1.53
DRIVER RACE				
WHITE*		1.00		
BLACK	-0.1361	0.87	0.79	0.96
NATIVE AMERICAN	0.3495	1.42	1.02	1.96
HISPANIC	-0.0301	0.97	0.84	1.13
ASIAN/OTHER	-0.2703	0.76	0.55	1.06
LICENSE STATUS AT STUDY CRASH				
FULLY NC LICENSED*		1.00		
SUSPENDED/LIMITED PRIVILEGE NC LICENSE	0.9299	2.53	2.09	3.07
NC UNLICENSED	0.8530	2.35	2.04	2.70
OVERALL NUMBER OF LICENSE RESTRICTIONS IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	-0.2009	0.82	0.75	0.90
TWO+	-0.4003	0.67	0.47	0.95
DAYLIGHT DRIVING ONLY RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	0.9972	2.71	0.95	7.74
OTHER LICENSE RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	0.9345	2.55	1.92	3.38
ONLY DRIVE WHILE SUPERVISED RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	0.5315	1.70	1.06	2.74
AUTOMATIC TRANSMISSION RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-1.6524	0.19	0.03	1.15
DRIVE CLASS B AND C ONLY RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-1.4537	0.23	0.05	1.02
GRADUATED DRIVER LICENSING RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	1.0292	2.80	2.03	3.85
BLOOD ALCOHOL CONTENT < .04 RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-0.7200	0.49	0.29	0.83

	n_{NAF} = 7,074, n_{SVAF} = 7,041			
Predictor	**β**	**OR**	**95% CI**	
NUMBER OF AT-FAULT CRASHES IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	0.3200	1.38	1.21	1.57
TWO	0.4371	1.55	1.17	2.04
THREE+	1.3226	3.75	2.05	6.86
AGE DIFFERENCE MOST RECENT PRIOR CRASH INVOLVEMENT IN LAST 5 YEARS				
SAME AGE*		1.00		
ONE YEAR YOUNGER	-0.1768	0.84	0.69	1.01
TWO YEARS YOUNGER	-0.2412	0.79	0.64	0.96
THREE YEARS YOUNGER	-0.1519	0.86	0.69	1.07
FOUR YEARS YOUNGER	-0.3898	0.68	0.53	0.86
FIVE YEARS YOUNGER	-0.2603	0.77	0.57	1.04
NO CRASH PRIOR 5 YEARS	-0.1172	0.89	0.75	1.06
OVERALL NUMBER OF CONVICTIONS IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	0.1685	1.18	1.07	1.31
TWO	0.1771	1.19	1.03	1.38
THREE	0.0824	1.09	0.90	1.31
FOUR	0.3499	1.42	1.08	1.86
FIVE+	0.2018	1.22	0.94	1.59
NUMBER OF OTHER SERIOUS CONVICTIONS IN PRIOR 5 YEARS				
NO CONVICTIONS*		1.00		
ONE	0.3322	1.39	1.04	1.87
TWO+	1.9671	7.15	1.91	26.70
OVERALL NUMBER OF LICENSE SUSPENSIONS IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	1.4765	4.38	3.57	5.37
TWO	2.6070	13.56	9.12	20.17
THREE	3.3371	28.14	15.64	50.62
FOUR+	4.7347	113.83	35.61	363.92
ADMINISTRATIVE LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-1.5098	0.22	0.08	0.59
DRIVING WHILE SUSPENDED LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-0.8343	0.43	0.31	0.60
FAILURE TO APPEAR/ FAILURE TO PAY LICENSE				

	n_{NAF} = 7,074, n_{SVAF} = 7,041			
Predictor	**β**	**OR**	**95% CI**	
SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-1.3108	0.27	0.22	0.34
NOT TRAFFIC RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-1.0879	0.34	0.23	0.50
OTHER LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-1.3959	0.25	0.12	0.52
POINTS-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-0.9861	0.37	0.22	0.64
GRADUATED DRIVER LICENSING-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-1.3630	0.26	0.14	0.45
RECKLESS DRIVING LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-2.1450	0.12	0.04	0.31
SPEED-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-0.8247	0.44	0.29	0.65

Note. The dependent variable for these analyses coded 0 = not-at-fault driver, 1 = at-fault driver. OR = odds ratio. 95%CI = 95% confidence interval for odds ratio. * = referent category. All predictors retained in the backwards removal model if variable $p < .10$.

The final licensing database predictor model for single-vehicle at-fault drivers was quite large, with a total of 26 predictor variables in the final Model 7. The results indicated that single-vehicle at-fault drivers were over-represented among drivers: (a) ages 16-19-years-old; (b) who were male; (c) who were Native American ethnicity (Black ethnicity was under-represented); (d) who had a suspended/limited privilege or no valid driver license; (e) with *no* license restrictions on record in the prior five years (versus one or two+ restrictions of any type); (f) with daylight driving only, supervised driving only, graduated driver licensing, or some “other” restriction on record in the prior five years (note also that automatic transmission-only, Class B & C only, and BAC < .04 restrictions were associated with *not* being at fault); (g) who had been at fault in one or more prior 5-year crashes (versus not being at fault or not having prior crashes); (h) who were the same age when involved in their most recent crash in the prior five years (vs. one or more years younger or having had no crash in the prior five years); (i) with one or more convictions of any type on record in the prior five years (versus no convictions on record); (j) with one or more serious convictions in the prior five years; (k) with one or more suspensions of any type on record in the prior five years (versus no suspensions—risk generally increases as the number of prior suspensions increases); and (l) who were *not* suspended for administrative reasons, for driving on a suspended license, for failure to appear in court or failure to pay a fine, for some non-traffic related reason, for some "other" reason, for accumulated license demerit points, for a GDL conviction, for reckless driving, or for speed convictions in the prior five years.

The classification of at-fault and not-at-fault single-vehicle crash drivers based on Model 7 (Table 27) was overall worse than that based on the crash predictor model. Specifically, (based on the 50% cut point) 60% of at-fault drivers were correctly classified by the licensing predictors versus 76% using the crash database predictors (Model 6).The classification of not-at-fault drivers was also worse based on the licensing predictors with only 77% of such drivers correctly classified versus 91% in Model 6.

Table 27. Classification Outcomes for Predicting Single-Vehicle At-Fault Drivers Using <u>License Database Variables Only (Model 7)</u> with 50% & 75% Probability Cut-Offs, North Carolina Drivers

Actual fault status[a]	Predicted fault status				Total
	Not-at-fault	Row %	At-fault	Row %	
50% probability cut-point					
Not-at-fault	5,474	**77.4**	1,600	22.6	7,074
At-fault	2,779	39.5	4,262	**60.5**	7,041
Total	8,253	58.5	5,862	41.5	14,115
75% probability cut-point					
Not-at-fault	6,694	**94.6**	380	5.4	7,074
At-fault	4,956	70.4	2,085	**29.6**	7,041
Total	11,650	82.5	2,465	17.5	14,115

Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off). The total counts differ across models because the inclusion or exclusion of variables with missing values changes the number of observations in the models.
[a]Based on the presence of one or more driver contributing circumstances for the crash.

Model 8: SVAF Crash & Licensing Database Model, NC Drivers. This final model for the single-vehicle at-fault drivers builds upon the prior models, in that it includes predictors from both the crash and licensing history databases. The same modified procedure used to account for part-whole multicollinearity and singularity among the licensing database predictors was used as was the case for Model 4 from the multi-vehicle at-fault drivers. The logistic regression results for the final backwards elimination model that resulted from this process are shown in Table 28.

Table 28. Logistic Regression Results Predicting Single-Vehicle Driver Fault vs. Multi-Vehicle Not-at-Fault Using Crash & Licensing Database Variables (Model 8), North Carolina Drivers Only

	n_{NAF} = 6,641, n_{SVAF} = 5,225			
Predictor	**β**	**OR**	**95% CI**	
Intercept	-0.7432			
DRIVER AGE GROUP IN STUDY CRASH				
16-19*		1.00		
20-29	-1.0346	0.36	0.30	0.43
30-39	-1.5477	0.21	0.17	0.26
40-49	-1.5997	0.20	0.16	0.25
50-59	-1.6389	0.19	0.15	0.24
60-69	-1.8609	0.16	0.12	0.21
70-79	-1.4221	0.24	0.17	0.35
80+	-1.1116	0.33	0.19	0.58
DRIVER RACE				
WHITE*		1.00		
BLACK	0.0557	1.06	0.91	1.22
NATIVE AMERICAN	-0.4068	0.67	0.40	1.09
HISPANIC	0.2432	1.28	1.01	1.61
ASIAN/OTHER	0.4635	1.59	1.00	2.53
DRIVER BELT USE IN STUDY CRASH				
BELTED, ANY TYPE*		1.00		
NO BELT	2.0127	7.48	6.40	8.75
DRIVER PHYSICAL CONDITION IN STUDY CRASH				
NORMAL*		1.00		
IMPAIRED (ALCOHOL, DRUGS , MEDICATION)	3.0438	20.99	12.01	36.66
SLEEPY/FATIGUED	6.9583	1051.9	145.28	7615.8
OTHER IMPAIRMENT	4.8319	125.45	59.54	264.35
DRIVER ALCOHOL/DRUG USE IN STUDY CRASH				
NO ALCOHOL OR DRUGS*		1.00		
ALCOHOL ONLY SUSPECTED/DETECTED	1.6483	5.20	3.48	7.76
DRUGS ONLY SUSPECTED/DETECTED	1.5273	4.61	1.14	18.62
BOTH ALCOHOL/DRUGS SUSPECTED/DETECTED	1.2741	3.58	1.03	12.36
VEHICLE TYPE IN STUDY CRASH				
CAR*		1.00		
PICKUP	-0.4274	0.65	0.56	0.76
SPORT UTILITY VEHICLE	0.1331	1.14	0.98	1.33
VAN/MINI VAN	-0.6165	0.54	0.43	0.68

	n_{NAF} = 6,641, n_{SVAF} = 5,225			
Predictor	**β**	**OR**	**95% CI**	
VEHICLE INSURED IN STUDY CRASH				
INSURED VEHICLE*		1.00		
NOT INSURED	0.9093	2.48	1.52	4.07
NUMBER OF PASSENGERS IN STUDY CRASH				
NONE*		1.00		
ONE	-0.1011	0.90	0.79	1.03
TWO+	0.0886	1.09	0.94	1.27
HOUR (TIME) OF STUDY CRASH				
5:00-8:59 AM*		1.00		
9:00 AM-12:59AM	-0.1711	0.84	0.70	1.02
1:00-4:59 PM	-0.3890	0.68	0.57	0.81
5:00-8:59 PM	-0.3140	0.73	0.61	0.87
9:00 PM-12:59PM	0.3085	1.36	1.11	1.67
1:00-4:59 AM	1.2905	3.63	2.72	4.87
ROAD CLASS OF STUDY CRASH				
LOCAL ROAD*		1.00		
NC/US/STATE ROAD	-0.2199	0.80	0.64	1.01
INTERSTATE	0.9210	2.51	1.84	3.43
RURAL ROAD	0.8605	2.36	1.89	2.95
LOCALITY TYPE OF STUDY CRASH LOCATION				
URBAN*		1.00		
MIXED	0.2080	1.23	0.99	1.53
RURAL	0.4247	1.53	1.19	1.96
DEVELOPMENT TYPE IN LOCATION OF STUDY CRASH				
RESIDENTIAL*		1.00		
COMMERCIAL/ INDUSTRIAL	-1.1072	0.33	0.27	0.40
FARM/UNDEVELOPED	0.1267	1.14	0.94	1.37
SPEED LIMIT OF STUDY CRASH ROAD				
< 55MPH*		1.00		
55MPH+	0.3932	1.48	1.28	1.71
LICENSE STATUS AT STUDY CRASH				
FULLY NC LICENSED*		1.00		
SUSPENDED/LIMITED PRIVILEGE NC LICENSE	0.6492	1.91	1.49	2.45
NC UNLICENSED	0.4056	1.50	1.21	1.86
DAYLIGHT DRIVING ONLY RESTRICTION IN PRIOR 5 YEARS				
NO*		1.00		
YES	1.0618	2.89	0.85	9.89
GRADUATED DRIVER LICENSING RESTRICTION IN PRIOR 5 YEARS				

	n_{NAF} = 6,641, n_{SVAF} = 5,225			
Predictor	**β**	**OR**	**95% CI**	
NO*		1.00		
YES	1.0800	2.94	2.08	4.17
NUMBER OF NOT-AT-FAULT CRASHES IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	0.0410	1.04	0.90	1.20
TWO	0.0611	1.06	0.80	1.41
THREE+	0.7493	2.12	1.31	3.42
NUMBER OF AT-FAULT CRASHES IN PRIOR 5 YEARS				
NONE*		1.00		
ONE	0.1623	1.18	1.00	1.39
TWO	0.3992	1.49	1.03	2.17
THREE+	1.5166	4.56	1.81	11.47
NUMBER OF OTHER SERIOUS CONVICTIONS IN PRIOR 5 YEARS				
NO CONVICTIONS*		1.00		
ONE	0.2111	1.24	0.79	1.94
TWO+	1.9930	7.34	1.74	30.85
ALCOHOL-RELATED LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	0.5172	1.68	1.29	2.18
RECKLESS DRIVING LICENSE SUSPENSION IN PRIOR 5 YEARS				
NO*		1.00		
YES	-1.2449	0.29	0.07	1.24

Note. The dependent variable for these analyses coded 0 = not-at-fault driver, 1 = at-fault driver. OR = odds ratio. 95%CI = 95% confidence interval for odds ratio. * = referent category. All predictors retained in the backwards removal model if variable $p < .10$.

When variables from both databases were considered as predictors of single-vehicle crash driver fault in Model 8, eight licensing variables and 13 crash database predictors were retained in the final model (21 total predictors). The relations of the crash database predictors to single-vehicle driver culpability were consistent with those observed when the licensing database predictors were not included in the model. Specifically, single-vehicle at-fault drivers were over-represented among those: (a) ages 16-19-years-old; (b) Black, Hispanic or Asian/Other ethnicity; (c) not belted during the study crash; (d) impaired by alcohol, drugs, or medication, sleep/fatigue (in particular), or some other cause (e.g., being ill or missing a limb) during the study crash; (e) with suspected/evidence of alcohol alone, drugs alone, or both alcohol and drugs during the study crash; (f) driving a car or SUV (as opposed to a truck or van); (g) driving an uninsured vehicle; and (h) driving without any passengers or 2+ passengers (vs. driving with only one passenger).

The relations of the environment-related variables were also the same as was the case for Model 6. Specifically, single-vehicle crashes were over-represented: (a) during the hours of 9:00

p.m. to 1:00 a.m. and between 1:00 to 5:00 a.m.; (b) on interstate and rural roadways, but under-represented on North Carolina, state, and U.S. roadways (relative to local roads); (c) in mixed urban/rural and rural locations; (d) in residential and farming/undeveloped areas (versus commercial or industrial areas); and (e) on roadways with 55MPH or higher speed limits. Overall, single-vehicle crash culpability seemed to be most strongly associated with young age, driver impairment from alcohol, drugs, or medication, impairment from sleepiness/fatigue, and other sources of driver impairment noted on the study crash report forms.

With regard to the license database predictors, the results of Model 8 indicated that single-vehicle at-fault drivers were over-represented among those: (a) with a suspended/limited privilege license or who were unlicensed; (b) who had a daylight driving-only license restriction in the prior 5-years; (c) who had a GDL license restriction in the prior 5 years; (d) who were involved in three or more not-at-fault crashes in the prior 5-years; (e) who were at-fault for 1, 2, or 3+ crashes in the prior 5-years; (f) who had 1 or 2+ serious moving convictions (e.g., reckless driving) in the prior five years; (g) who had an alcohol-related license suspension in the prior 5 years; and (h) who had *not* been suspended for reckless driving in the prior five years.

Similar to what occurred for the multi-vehicle at-fault drivers, the inclusion of the licensing database predictors did not much improve the overall accuracy of model classification of culpability (Table 29). With addition of the licensing database variables the model correctly classified about 1% point more single-vehicle at fault drivers using either probability cut-point. The classification of the not-at-fault drivers similarly changed only a negligible amount.

Table 29. Classification Outcomes for Predicting Single-Vehicle At-Fault Drivers Using Crash & Licensing Database Variables (Model 8) with 50% & 75% Probability Cut-Offs, NC Drivers

Actual fault status[a]	Predicted fault status				
	Not-at-fault	**Row %**	**At-fault**	**Row %**	**Total**
50% probability cut-point					
Not-at-fault	6,093	**91.8**	548	8.2	6,641
At-fault	1,207	23.1	4,018	**76.9**	5,225
Total	7,300	61.5	4,566	38.5	11,866
75% probability cut-point					
Not-at-fault	6,481	**97.6**	160	2.4	6,641
At-fault	1,895	36.3	3,330	**63.7**	5,225
Total	8,376	70.6	3,490	29.4	11,866

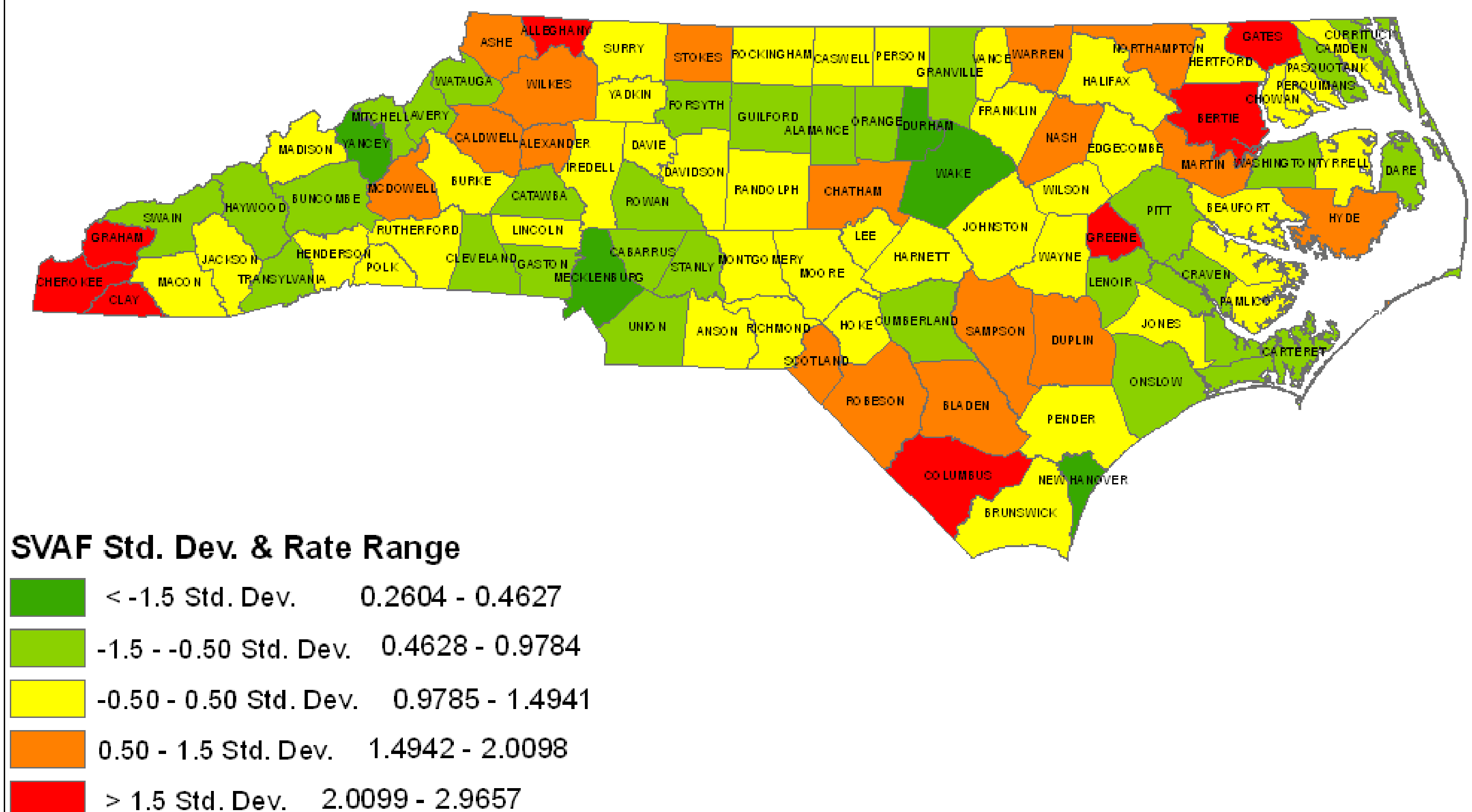
Note. A driver was predicted to have been at-fault in the fatal/serious injury crash if the probability based on the model was greater than .50 (50% probability cut-off) or .75 (75% probability cut-off). The total counts differ across models because the inclusion or exclusion of variables with missing values changes the number of observations in the models.

[a]Based on the presence of one or more driver contributing circumstances for the crash.

Appendix 3

At Fault Crash Involvement Rates by County

NC Single Vehicle Crash at Fault Involvement Rates per 1000 Population



NC Multi-Vehicle Crash at Fault Involvement Rates per 1000 Population

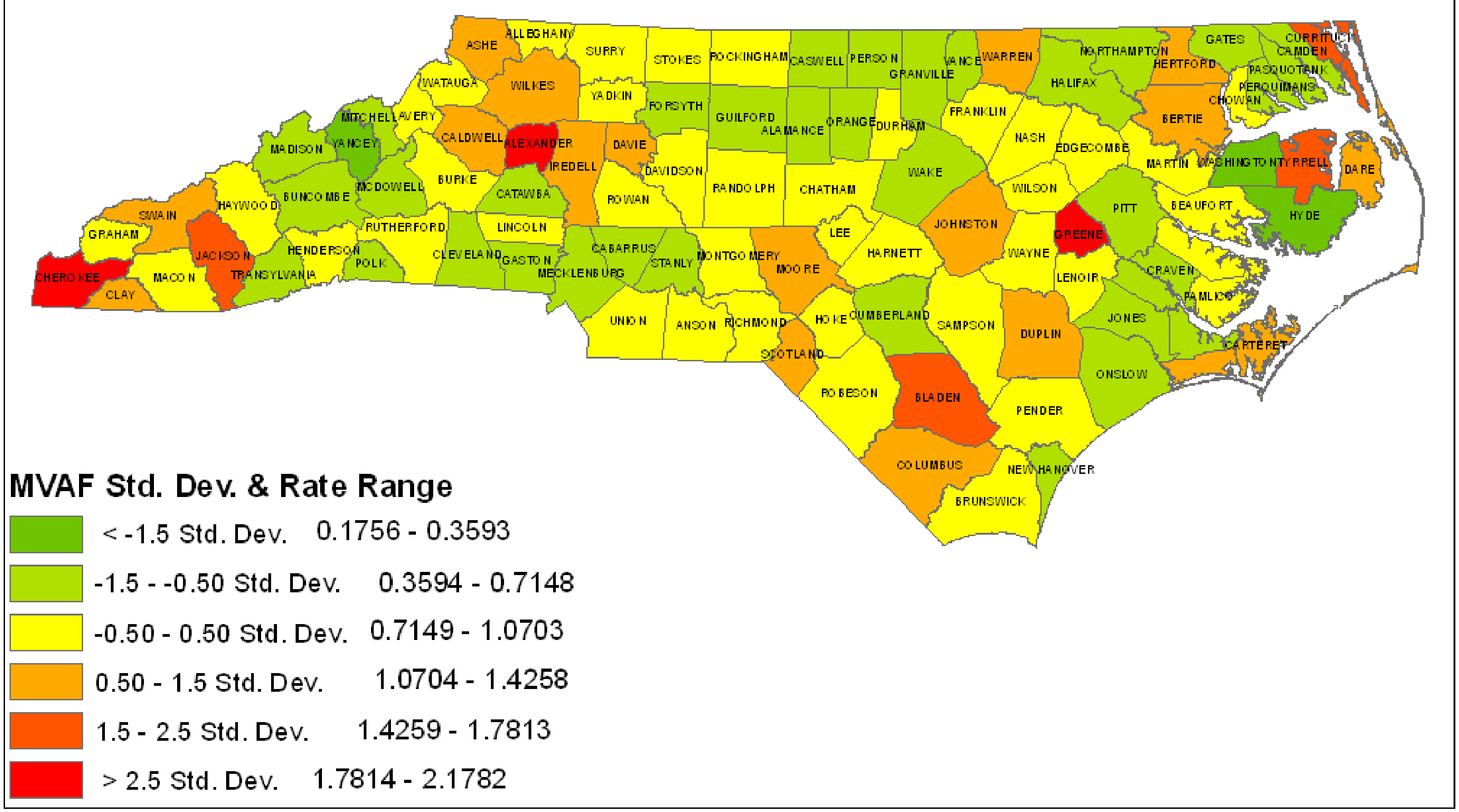


Table 30. Not-at-fault, Multi-vehicle At-fault and Single-vehicle At-fault river Crash Involvements by County.

COUNTY	NAF No.	NAF_rate	MVAF No.	MVAF_rate	SVAF No.	SVAF_rate	All_drivers	Pop. July 2003
Alamance	107	0.7853	90	0.6605	112	0.8220	309	136252
Alexander	87	2.5192	75	2.1717	61	1.7663	223	34535
Alleghany	11	1.0187	11	1.0187	24	2.2226	46	10798
Anson	24	0.9531	24	0.9531	37	1.4694	85	25180
Ashe	33	1.3162	30	1.1966	40	1.5954	103	25072
Avery	19	1.0561	15	0.8338	15	0.8338	49	17990
Beaufort	51	1.1204	48	1.0545	64	1.4060	163	45518
Bertie	24	1.2156	23	1.1649	40	2.0259	87	19744
Bladen	52	1.5919	47	1.4388	62	1.8980	161	32666
Brunswick	84	1.0267	75	0.9167	114	1.3934	273	81817
Buncombe	170	0.8010	142	0.6691	186	0.8764	498	212224
Burke	101	1.1381	90	1.0142	103	1.1606	294	88744
Cabarrus	105	0.7325	87	0.6069	78	0.5442	270	143340
Caldwell	111	1.4193	92	1.1764	133	1.7006	336	78208
Camden	6	0.7645	4	0.5097	9	1.1468	19	7848
Carteret	75	1.2344	67	1.1028	48	0.7900	190	60756
Caswell	14	0.5903	11	0.4638	31	1.3071	56	23716
Catawba	109	0.7435	97	0.6616	103	0.7026	309	146608
Chatham	62	1.1537	52	0.9676	82	1.5258	196	53742
Cherokee	62	2.4554	55	2.1782	57	2.2574	174	25250
Chowan	14	0.9794	11	0.7696	18	1.2593	43	14294
Clay	10	1.0667	11	1.1733	22	2.3467	43	9375
Cleveland	62	0.6367	60	0.6162	90	0.9243	212	97376
Columbus	76	1.3952	73	1.3401	122	2.2396	271	54473
Craven	56	0.5996	47	0.5032	60	0.6424	163	93402
Cumberland	212	0.6878	182	0.5905	164	0.5321	558	308217
Currituck	37	1.7963	34	1.6506	17	0.8253	88	20598
Dare	57	1.7112	47	1.4110	17	0.5104	121	33310
Davidson	192	1.2643	160	1.0536	172	1.1326	524	151867
Davie	49	1.3176	41	1.1024	48	1.2907	138	37190
Duplin	63	1.2404	55	1.0829	89	1.7523	207	50791
Durham	211	0.8964	182	0.7732	102	0.4333	495	235388
Edgecombe	54	1.0029	50	0.9286	65	1.2072	169	53844

COUNTY	NAF No.	NAF_rate	MVAF No.	MVAF_rate	SVAF No.	SVAF_rate	All_drivers	Pop. July 2003
Forsyth	174	0.5486	146	0.4603	148	0.4667	468	317150
Franklin	61	1.1809	53	1.0260	57	1.1035	171	51656
Gaston	97	0.5074	85	0.4446	142	0.7427	324	191183
Gates	8	0.7414	7	0.6487	32	2.9657	47	10790
Graham	7	0.8693	7	0.8693	21	2.6080	35	8052
Granville	24	0.4593	25	0.4784	46	0.8802	95	52258
Greene	37	1.8636	36	1.8132	44	2.2162	117	19854
Guilford	348	0.8079	288	0.6686	260	0.6036	896	430744
Halifax	36	0.6346	31	0.5465	68	1.1988	135	56725
Harnett	120	1.2293	103	1.0551	135	1.3829	358	97619
Haywood	49	0.8775	48	0.8596	50	0.8954	147	55838
Henderson	108	1.1424	94	0.9943	103	1.0895	305	94538
Hertford	30	1.2639	26	1.0954	31	1.3060	87	23736
Hoke	42	1.1375	38	1.0292	44	1.1917	124	36922
Hyde	1	0.1756	1	0.1756	11	1.9315	13	5695
Iredell	165	1.2385	144	1.0808	161	1.2084	470	133229
Jackson	54	1.5451	50	1.4306	47	1.3448	151	34950
Johnston	182	1.3342	159	1.1656	192	1.4076	533	136407
Jones	7	0.6879	6	0.5896	14	1.3758	27	10176
Lee	51	0.9805	47	0.9036	55	1.0574	153	52014
Lenoir	66	1.1228	57	0.9697	49	0.8336	172	58780
Lincoln	72	1.0691	58	0.8612	89	1.3215	219	67349
Macon	44	1.0225	39	0.9063	49	1.1387	132	43032
Madison	19	0.6064	18	0.5745	35	1.1171	72	31330
Martin	22	1.1015	21	1.0515	32	1.6022	75	19972
McDowell	17	0.6832	14	0.5627	38	1.5272	69	24882
Mecklenburg	602	0.8029	482	0.6428	344	0.4588	1428	749804
Mitchell	10	0.6285	10	0.6285	13	0.8171	33	15910
Montgomery	24	0.8784	22	0.8052	33	1.2078	79	27323
Moore	101	1.2928	89	1.1392	106	1.3568	296	78123
Nash	90	1.0057	79	0.8828	162	1.8102	331	89492
New Hanover	144	0.8522	114	0.6746	44	0.2604	302	168977
Northampton	12	0.5524	12	0.5524	42	1.9335	66	21722
Onslow	116	0.7405	98	0.6256	136	0.8682	350	156646
Orange	82	0.6824	71	0.5908	90	0.7490	243	120168

COUNTY	NAF No.	NAF_rate	MVAF No.	MVAF_rate	SVAF No.	SVAF_rate	All_drivers	Pop. July 2003
Pamlico	12	0.9241	10	0.7701	16	1.2321	38	12986
Pasquotank	23	0.6327	18	0.4952	22	0.6052	63	36352
Pender	45	1.0296	38	0.8694	63	1.4414	146	43706
Perquimans	5	0.4271	5	0.4271	13	1.1105	23	11706
Person	30	0.8122	26	0.7039	54	1.4620	110	36936
Pitt	83	0.5983	72	0.5190	84	0.6055	239	138726
Polk	10	0.5306	9	0.4776	25	1.3265	44	18846
Randolph	128	0.9489	121	0.8970	166	1.2307	415	134887
Richmond	36	0.7757	35	0.7541	47	1.0127	118	46410
Robeson	139	1.1083	116	0.9249	194	1.5468	449	125422
Rockingham	100	1.0827	91	0.9853	116	1.2559	307	92362
Rowan	157	1.1797	135	1.0144	121	0.9092	413	133080
Rutherford	63	0.9944	57	0.8997	84	1.3258	204	63357
Sampson	74	1.1911	63	1.0140	101	1.6257	238	62128
Scotland	50	1.4077	46	1.2950	54	1.5203	150	35520
Stanly	43	0.7307	38	0.6457	51	0.8666	132	58851
Stokes	42	0.9203	40	0.8765	76	1.6653	158	45637
Surry	78	1.0847	75	1.0429	107	1.4879	260	71912
Swain	14	1.0493	16	1.1992	13	0.9744	43	13342
Transylvania	24	0.8149	21	0.7130	24	0.8149	69	29452
Tyrrell	8	1.8877	7	1.6517	6	1.4158	21	4238
Union	168	1.1606	141	0.9741	108	0.7461	417	144747
Vance	30	0.6857	29	0.6629	60	1.3714	119	43750
Wake	412	0.5876	336	0.4792	242	0.3451	990	701177
Warren	22	1.1003	23	1.1503	39	1.9506	84	19994
Washington	3	0.2229	3	0.2229	13	0.9661	19	13456
Watauga	44	1.0294	39	0.9125	34	0.7955	117	42742
Wayne	132	1.1591	119	1.0449	127	1.1152	378	113883
Wilkes	86	1.2858	75	1.1213	127	1.8988	288	66886
Wilson	68	0.8996	65	0.8600	80	1.0584	213	75585
Yadkin	38	1.0325	32	0.8695	38	1.0325	108	36804
Yancey	5	0.2794	4	0.2235	7	0.3911	16	17896
Total/statewide	7594	0.9023	6571	0.7808	7620	0.9054	21785	8415955

COUNTY	NAF No.	NAF_rate	MVAF No.	MVAF_rate	SVAF No.	SVAF_rate	All_drivers	Pop. July 2003
mean over all counties				0.8926		1.2300		
1 std. dev.				0.3600		0.5200		
				Exceeds 1.5 std devs		Exceeds 1.5 std. devs		
				Exceeds 2.5 std. devs.				