USING VEHICLE-BASED SENSORS OF DRIVER BEHAVIOR TO DETECT ALCOHOL IMPAIRMENT

John Lee
University of Wisconsin-Madison
United States
Timothy Brown
University of Iowa
United States
Dary Fiorentino
Southern California Research Institute
United States
James Fell
Pacific Institute for Research & Evaluation
United States
Eric Traube
National Highway Traffic Safety Administrations
United States
Eric Nadler
Volpe Transportation Systems
United States
Paper Number 11-0322

ABSTRACT

Despite persistent efforts at the local, state, and federal levels, alcohol-impaired crashes still contribute to approximately 30% of all traffic fatalities. Although enforcement and educational approaches have helped to reduce alcohol-impaired fatalities, other approaches will be required to further reduce alcohol-related fatalities. This paper describes an approach that detects alcohol impairment in real time using vehicle-based sensors to detect alcohol-related changes in drivers' behavior.

Data were collected on the National Advanced Driving Simulator from 108 volunteer drivers. Three age groups (21-34, 38-51, and 55-68 years of age) drove through representative situations on three types of roadways (urban, freeway, and rural) at three levels of blood alcohol content (0.00%, 0.05%, and 0.10% BAC).

Driver control input, vehicle state, driving context and driver state data, individually and in combination, reveal signatures of alcohol impairment. Algorithms built on these signatures detect drivers with BAC levels that are over the legal limit with an accuracy of approximately 80%, similar to the Standardized Field Sobriety Test (SFST) used by law enforcement. Each of the three algorithms combined information across time to predict impairment. The time required to detect impairment ranged from eight minutes, for complex algorithms (i.e., support vector machines and decision trees

applied to relatively demanding driving situations), to twenty-five minutes for simple algorithms (i.e., logistic regression). Timely impairment detection depends critically on the driving context: variables specific to the particular driving situation result in much more timely impairment detection than generic variables.

INTRODUCTION

Despite persistent efforts at the local, state, and federal levels, alcohol-impaired driving crashes still contribute to approximately 30% of all traffic fatalities. The proportion of fatally injured drivers with blood alcohol concentrations (BAC) greater than or equal to 0.08% has remained at 31-32% for the past ten years [1]. Although enforcement and educational approaches have helped to reduce alcohol-impaired fatalities, other approaches merit investigation. One such approach concerns countermeasures that capitalize on the increasingly sophisticated sensor and computational platform that is available on many production vehicles. Such vehicle-based countermeasures have the potential to address alcohol-impaired driving and save thousands of lives each year.

Vehicle-based countermeasures use sensors that describe drivers' control inputs (e.g., steering wheel and brake pedal movement), vehicle state (e.g., accelerometer and lane position), driving context (e.g., speed zone information and proximity of surrounding vehicles), and driver state (e.g., eye movements and posture). Data from these sensors can be transformed, combined, and processed with a variety of algorithms to develop a detailed description of the driver's response to the roadway. These sensors and algorithms hold promise for identifying a range of driver impairments, including distraction, drowsiness, and even age-related cognitive decline. Alcohol represents a particularly important impairment that might be detected by vehicle-based sensors and algorithms.

This paper describes the development and evaluation of algorithms to detect the behavioral signature of alcohol. Such an algorithm is a central element of any vehicle-based countermeasure for alcohol-related crashes. Algorithm development depends on collecting data from impaired and unimpaired drivers. This research used data collected from three age groups of drivers (21-34, 38-51, and 55-68 years of age) driving through representative situations on three types of roadways (urban, freeway, and rural) at three levels of alcohol concentration (0.00%, 0.05%, and 0.10% BAC). The high fidelity of the National Advanced Driving Simulator (NADS) makes these data unique. Drivers' control inputs, vehicle state, driving context, and driver state were captured in representative driving situations, with precise control and in great detail. This report describes how, individually and in combination, these data reveal signatures of alcohol impairment, and how well algorithms built on these signatures detect drivers with BAC levels that are over the legal limit of 0.08%.

The overall objectives were to:

- Identify signatures of impairment and develop algorithms to detect alcohol-related impairment
- Compare robustness of metrics and algorithms

METHOD

Participants

Data were collected from 108 volunteer drivers from three age groups (21-34, 38-51, and 55-68 years of age) driving through representative situations on three types of roadways (urban, freeway, and rural) at three levels of blood alcohol content (0.00%, 0.05%, and 0.10% BAC). Table 1 summarizes participant characteristics.

To be eligible, participants were required to:
- Possess a valid US driver's license
- Have been licensed driver for two or more years
- Drive at least 10,000 miles per year
- Have no restrictions on driver's license except for vision
- To not have been taking illegal drugs or drugs that interacted with alcohol
- Not require the use of any special equipment to drive.
- Have been a moderate to heavy drinker, but not a chronic alcohol abuser.

Table 1. Participant Characteristics

	Age 21-34		Age 38-51		Age 55-68	
Variable	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Number completed	18	18	18	18	18	18
Mean age (years)	26.6	26.8	43.2	44.7	59.6	61.1
Mean height (inches)	70.7	65.5	70.6	65.4	70.1	64.8
Mean weight (pounds)	199.8	159.6	220.6	175.3	211.9	172.9
Mean body mass index	27.9	26.1	31.1	28.6	30.2	29.0
Heavy Drinkers	89%	61%	78%	50%	67%	61%

Procedure

An initial telephone interview was conducted to determine eligibility for the study. Applicants were screened in terms of health history, current health status, and use of alcohol and other drugs. The Quantity-Frequency-Variability (QFV) scale [2] was used to determine whether applicants were moderate drinkers or heavy drinkers and the Audit survey [3, 4] was used to exclude chronic alcohol abusers. Pregnancy, disease, or evidence of substance abuse resulted in exclusion from the study. Participants taking prescription medications that interact with alcohol were also excluded from the study.

Each participant participated in four sessions, the last three separated by one week. Order of target BAC levels and scenario event sequence were counterbalanced. The time of day of each of the three sessions was the same for a given participant.

On study Visit 1 (screening), each participant informed consent was obtained. They then provided a urine sample for the drug screen and, for females, the pregnancy screen. During a five-minute period following these activities, the participant sat alone in the room where subsequent measurements of blood pressure, heart rate, height, and weight were made.

Cardiovascular measures were taken and compared to acceptable ranges (systolic blood pressure = 120 ± 30 mmHg, diastolic blood pressure = 80 ± 20 mmHg, heart rate = 70 ± 20) to assess eligibility for the study. If participants met study criteria, they were then administered a breath alcohol test and verbally administered the QFV and the Audit Survey to further confirm eligibility. Participants who were not moderate or heavy drinkers on the QFV were excluded. Additionally, participants who were classified with potentially dangerous drinking patterns on the Audit Survey were excluded.

If participants met study criteria, they completed demographic surveys. These surveys included questions related to crashes, moving violations, driver behavior, drinking, and driving history. Participants viewed an orientation and training presentation that provided an overview of the simulator cab and the secondary task they were asked to complete while driving.

The task consisted of the participant turning on the CD player and sequentially advancing the CD player to two tracks provided in an auditory cue and then turning off the CD player.

Participants then completed the practice drive and completed surveys after their drive about how they felt and about the realism of the simulator. The practice drive included making a left hand turn, driving on two- and four-lane roads, and changing CDs.

During Visits 2, 3, and 4 all participants completed a urine drug screen and, for females, a pregnancy screen to confirm eligibility for the study. Participants' blood pressure and heart rate were obtained to verify study eligibility. If participants met study criteria, they then received a breath alcohol test, the QFV, and the Audit Survey to further confirm eligibility. If eligible to continue, the time and duration of last sleep, and time and contents of last meal were recorded. Age, gender, height, weight, and drinking practice were used to calculate the alcohol dose.

Participants were served three equal-sized drinks at 10-minute intervals and were instructed to pace each drink evenly over the 10-minute period. NADS staff monitored the participants periodically throughout the drinking period to ensure an even pace of drinking.

On the days when participants were dosed to achieve 0.10% and 0.05% BAC, the amount of alcohol consumed was calculated to produce a peak BAC of 0.115% or a peak BAC of 0.065%. On the 0.00% peak BAC day, the drink consisted of one part water and 1.5 parts orange juice. Each of the glasses had its rim swabbed with vodka and 10 ml of vodka was floated to produce an initial taste and odor of alcohol.

Sixteen minutes after the end of the third drink, BAC measurements were taken at two- to five-minute intervals until the target BAC (± 0.005%) was reached. Peak BAC was expected 30 minutes after the end of the third drink. Table 2 summarizes the BAC levels.

Table 2. Summary of BAC levels for the two experimental conditions.

Test Time	**0.05% BAC (*N* = 108)**			**0.10% BAC (*N* = 108)**		
	M	***SD***	**Median**	***M***	***SD***	**Median**
Pre-drive	0.053	0.005	0.054	0.098	0.009	0.102
Post-drive	0.042	0.006	0.043	0.088	0.009	0.090
Mean	0.047	0.005	0.048	0.093	0.008	0.095

When the target BAC was reached, the participants drove in the NADS. All data were collected as the BAC declined to minimize extraneous variation associated with the effect of rising and falling BAC levels and to represent the most likely situation under which alcohol-impaired driving occurs. As soon as the simulator returned to the dock and the participant exited the simulator (within 5 minutes of completing the drive), a BAC measurement was obtained, followed by an SFST. The individuals conducting the SFST were trained according to NHTSA's guidelines. The Stanford Sleepiness scale was also administered before and after each drive.

Participants were not informed of their measured BACs until their participation in the study was completed. On all experimental days, the participants were transported home after their BAC dropped below 0.03%. At the end of Visit 4, participants were debriefed and paid $250. Pro-rated compensation was provided for participants who did not complete the study.

Apparatus

The National Advanced Driving Simulator (NADS), shown in Figure 1, made it possible to collect representative driving behavior data from intoxicated drivers in a safe and controlled manner. This is the highest fidelity simulator in the United States and allowed for precise characterization of driver response. Drivers' control inputs, vehicle state, driving context, and driver state were captured in representative driving situations (see Figure 2).

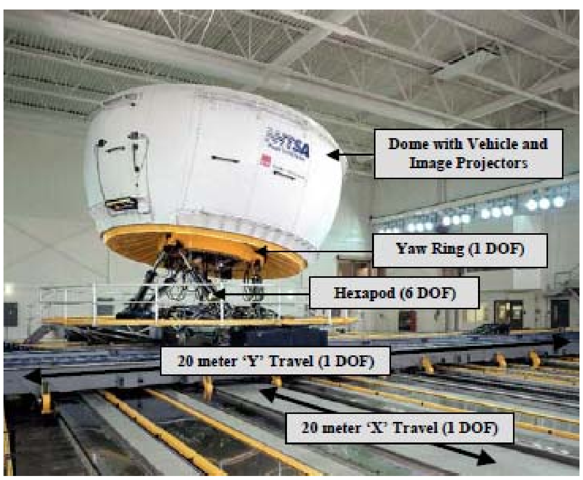

Figure 1. The NADS-1 high-fidelity driving simulator

Figure 2. An urban driving scene from the NADS-1 simulator.

Simulator Scenario

Each drive was composed of three nighttime driving segments. The drives started with an urban segment composed of a two-lane roadway through a city with posted speed limits of 25 to 45 mph with signal-controlled and uncontrolled intersections. (see Figure 3 and Figure 4) An interstate segment followed that consisted of a four-lane divided expressway with a posted speed limit of 70 mph. Following a period in which drivers followed the vehicle ahead, they encountered infrequent lane changes associated with the need to pass several slower-moving trucks (see Figure 5). The drives concluded with a rural segment composed of a two-lane undivided road with curves (see Figure 6 and Figure 7). A portion of the rural segment was gravel. These three segments mimicked a drive home from an urban bar to a rural home via an interstate. Events in each of the three segments combined to provide a representative trip home in which drivers encountered situations that might be encountered in a real drive.

Figure 3. Approach to curve in urban drive

Figure 4. Straight roadway segment in urban drive

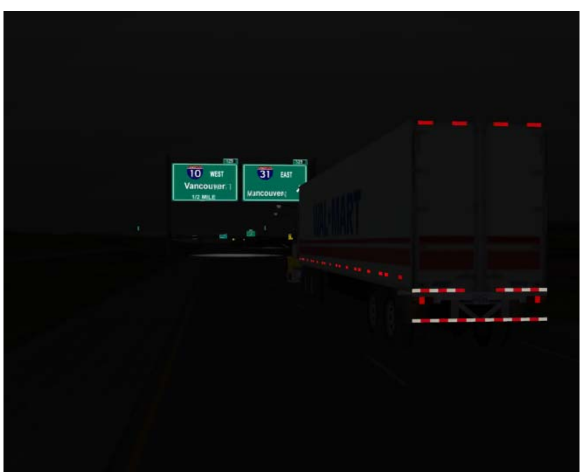

Figure 5. Passing truck on Interstate.

Figure 6. Approach to rural curve

Figure 7. Rural vertical curve.

RESULTS

Sensitivity of Scenarios to Alcohol

Analysis of common driving metrics demonstrates the sensitivity of the drive to alcohol impairment. As expected, lane position variation was particularly sensitive (see Figure 8) and speed variation (see Figure 9) was less so. Increasing BAC levels affected driving performance in an orderly manner—higher BAC levels led to a linear decrease in performance. Alcohol levels did not interact with age, gender, and roadway situation, which might have otherwise undermined the association of driving metrics and alcohol impairment.

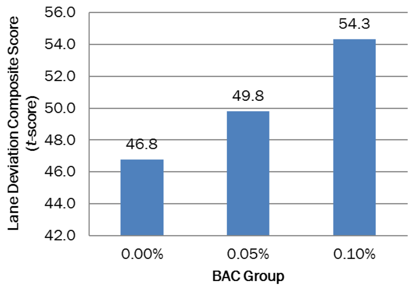

Figure 8. Lane deviation scores by BAC group

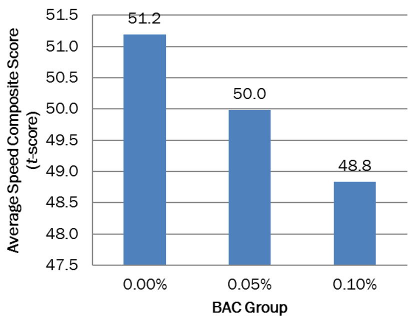

Figure 9. Average speed scores by BAC group.

Detecting Alcohol Impairment

The primary objectives for algorithm development and evaluation include:

- Develop algorithms to detect alcohol-related impairment based on behavioral signatures that vehicle-based sensors can measure
- Compare sensitivity, robustness, and timeliness of metrics and algorithms.

These objectives are addressed by describing the performance of a logistic regression algorithm that builds directly on an analysis of simple measures of driving performance—lane position variability, mean speed, and speed variability. To go beyond a linear combination of these three simple indicators of driver impairment, a decision tree algorithm was fit to individual events and to the urban, freeway, and rural segments to identify behavioral signatures of alcohol impairment. A support vector machine (SVM) was also developed for these road segments. These signatures provide a detailed description of alcohol impairment that supports more accurate detection than the three-variable logistic regression.

The objective of the following analyses was to determine whether it is possible to distinguish between drivers with BACs above 0.08% and those below 0.08%. To that end, a new variable was created (BAC Status) by dichotomizing the pre- and post-drive BACs as either both being less than 0.08% or both being at or above 0.08%. The dichotomization produced 313 valid cases. Eleven cases were eliminated because the pre- and post-drive BACs were not on the same side of the 0.08% cutoff. The median BAC for the low BAC status condition (BAC < 0.08%) was 0.037%. The median BAC for the high BAC status condition (BAC ≥ 0.08%) was 0.097%. The median differences between the conditions were 0.06%.

Three general algorithms were developed. The first was based on logistic regression and was fit using a standard least squares regression approach using the entire dataset. The two other approaches to algorithm development used support vector machines (SVMs) and decision trees, which can often outperform linear combinations of the features [5].

Originally developed by Vapnik [6], SVMs have several advantages over approaches that make assumptions of linearity and normality. The SVM approach identifies a hyperplane that separates instances with different BAC levels [7]. SVMs are particularly well-suited to extract information from noisy data [8] and avoid overfitting by minimizing the upper bound of the generalization error [9]. The C4.5 decision tree approach, developed by Quinlan, classifies data by creating a tree that divides the data using the gini index, which weights feature influence in a linear fashion [10, 11]. Adaptive boosting (AdaBoost) sequentially fits a series of classification algorithms, with greater emphasis on previously misclassified instances. It then combines the output of the classification algorithms by adjusting the importance of each classifier based on its error rate [12]. This approach is particularly valuable where a single decision tree or SVM cannot capture the complexity of the underlying relationships. Adaptive boosting was applied to both the Decision Tree and SVM, but not the logistic regression. Detailed discussion of the algorithms can be found in the NHTSA report [13].

Three criteria are used throughout to assess algorithm sensitivity: accuracy, positive predictive performance (PPP), and area under curve (AUC). Accuracy measures the percent of cases that were correctly classified, and PPP measures the degree to which those drivers that were judged to have high BAC levels actually had high BAC levels.

Performance measures such as correct detection or overall accuracy fail to provide a complete description of algorithm performance because they do not account for the baseline frequency of impairment nor differences in the decision criterion. An algorithm can correctly identify all instances of impairment simply by setting a very low decision criterion, but such an algorithm would misclassify all cases where there was no impairment. The signal detection parameter, d', avoids these problems, but its underlying assumptions include symmetry of signal and noise distributions, which are often violated. AUC is a nonparametric version of d', and represents the area under the receiver operator curve, which provides a robust performance measure that does not depend on the assumptions underlying d'. Perfect classification performance is indicated by an AUC of 1.0, and chance performance is indicated by 0.50. AUC is an unbiased measure of algorithm performance, but accuracy and PPP are more easily interpreted, so all three are used in describing the algorithms.

Three different algorithms (logistic regression, support vector machines, and decision trees) were developed to predict whether the driver was above the legal limit, using average speed, minimum speed, variability in speed, lane position and variability in lane position. The algorithms achieved an accuracy of approximately 80%, comparable to that of the SFST used by law enforcement. Each of the three algorithms combined information across time to predict impairment.

Classification accuracy was consistent with previous studies—classification accuracy exceeded 82% for all three algorithms, with the decision tree being most accurate (84.7), followed by SVM (82.3) and logistic regression (82.0). Not surprisingly, the performance discriminating between BAC levels above and below 0.08 was somewhat worse than between the more extreme range defined by the experimental conditions of 0.00% and 0.10% BAC. Table 3 shows the accuracy ranges from approximately 80.5 to 82.5%. Given that even with the SFST, the current "gold-standard" for identifying alcohol impairment, there was overlap between the BAC levels, as shown in Figure 10, the failure of the algorithms to perfectly discriminate between BACs is not surprising.

Table 3. Performance of three algorithms classifying drivers with BAC above and below 0.08% using the SFST, with confidence intervals in the parentheses.

	Accuracy	AUC	PPP
Decision tree	81.8 (5.9)	.76 (0.087)	78.4 (15.5)
SVM	80.5 (6.9)	.81 (0.072)	75.6 (17.9)
Logistic regression	82.5 (5.5)	.80 (0.062)	75.9 (13.6)

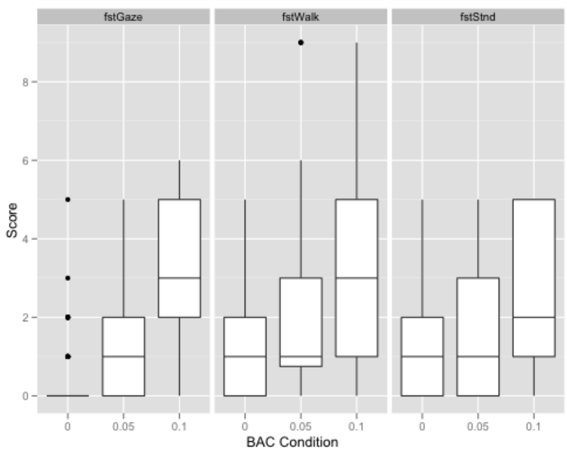

Figure 10. SFST scores show considerable overlap across BAC conditions.

The time required to detect impairment varied depending on the algorithm used. For example, when a simple algorithm (i.e., logistic regression) was used, it could take twenty-five minutes to detect impairment. However, when more complex algorithms were used (i.e., support vector machines and decision trees) for relatively demanding driving situations, impairment could be detected in as little as eight minutes. The time required to detect impairment depends on the driving context: impairment is detected more quickly when variables specific to the particular driving situation are considered (e.g. lane keeping on a rural road), rather than generic variables (e.g., number of lane departures). Timely impairment detection depends critically on the driving context: variables specific to the particular driving situation result in much more timely impairment detection than generic variables.

To illustrate this effect, the area under the curve for the decision tree algorithm are plotted in Figure 11 and show a general trend toward increasing sensitivity with longer events, but also indicate that longer events provide an increasing benefit. This figure also shows the substantial differences between events, with Urban Drive (102) and Urban Curves (106) being more sensitive than their duration would suggest, contrasting with Interstate Curves (205), which is less sensitive. As noted previously, highly precise impairment detection can occur in eight minutes if the driver encounters situations similar to Urban Curves (106) followed by Dark Rural (304). These results show that timely impairment detection depends on the types of events encountered by the driver, as well as the duration of information accumulation.

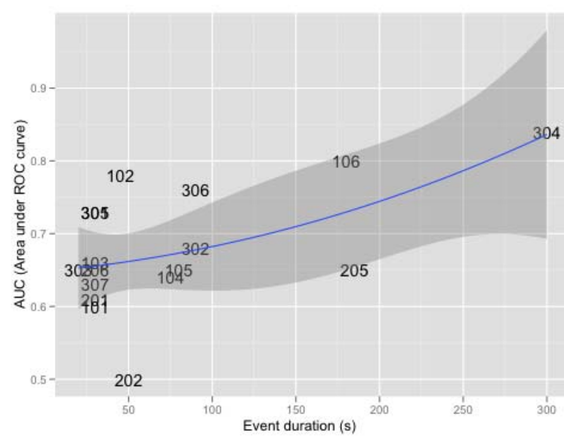

Figure 11. The sensitivity of each event as a function of its duration

Just as algorithms that consider differences in the driving context perform better, so do algorithms that consider differences between drivers. Algorithms tailored to the individual perform much better than generic algorithms that do not reflect differences between drivers. For example, an individualized algorithm might focus on a change from an individual's ability to accurately maintain lane position or speed rather than reaching a generic threshold of impaired lane keeping or speed. Although generic algorithms provide sufficient sensitivity to be useful, even very limited individualization greatly improves performance.

CONCLUSIONS

This study demonstrated that a vehicle-based system using measures of driver behavior can differentiate between drivers with BAC levels above and below 0.08% with sensitivity similar to the SFST. Because the indicators of alcohol impairment become much stronger at higher levels, the sensitivity would likely increase substantially if the algorithm was used to identify those with BAC levels over 0.15%. These outcomes strongly support the potential of vehicle-based systems to detect impaired driving which could ultimately help to prevent and mitigate alcohol-related crashes.

On the basis of this research, standard deviation of lane position and average speed were shown to be reliable measures of impairment that can be feasibly captured over a number of driving situations, and appear robust enough to be useful in future vehicle-based countermeasures. Minimum speed, as well as standard deviation of lane position and speed are useful indicators that might have particular utility in alcohol warning monitors designed to provide feedback to drivers.

A second general finding is that the driving context strongly influences impairment-detection performance. Contrary to many previous simulator studies of alcohol-impaired driving, this study used a representative series of 19 events over three types of roadway situations. These events revealed that impairment detection depends on the type of event. Because driving is a satisficing rather than optimizing activity, drivers can take many paths through low-demand situations that are all satisfactory. This variety of satisfactory responses masks impairment. The variety of events also requires a greater variety of measures to capture the relevant behavior in each event. All of these findings imply that detecting alcohol-related impairment, and impairment detection more generally, depends on the driving situation. Algorithm development needs to consider roadway situations as much as it needs to consider the drivers' perceptual, motor, and decision-making response to the impairment.

These results support the long-term research objective of using algorithms that detect impairment to provide drivers with feedback that will discourage or prevent drinking and driving. Ultimately the distraction-detection algorithms developed in this study could support a range of vehicle-based interventions to prevent alcohol-related crashes. The promising results associated with alcohol-related impairment detection also suggest other types of impairment detection might also hold promise, most notably distraction and drowsiness.

ACKNOWLEDGEMENTS

The project benefited from the contributions of Mark Gilmer and Tim Brewer from Intoximeter, who arranged for a loan of the Alco-Sensors for this study.

We would like to thank Michelle L. Reyes , from the University of Iowa, Nic Ward, from Montana State University, and Robert Dufour, from Delphi, for their contributions to this research endeavor. We would also like to thank the staff at the National Advanced Driving Simulator who put in long days over three months to collect the data for this project.

REFERENCES

1. National Highway Traffic Safety Administration. *Fatality Analysis Reporting System (FARS)*. 2009 [cited 2009 June]; Available from: ftp://ftp.nhtsa.dot.gov/fars/.
2. Cahalan, D., I. Cisin, and Crossley, *American Drinking Practices: A National Study of Driving Behavior and Attitudes*. 1969, Rutgers University Press: New Brunswick, NJ.
3. Babor, T.F., et al., *AUDIT: The alcohol use disorders identification test: Guidelines for use in primary health care*. 1992, Geneva, Switzerland: World Health Organization.
4. Conley, T.B., *Construct validity of the MAST and AUDIT with multiple offender drunk drivers.* Journal of Substance Abuse Treatment, 2001. **20**(4): p. 287-295.
5. Liang, Y., M.L. Reyes, and J.D. Lee, *Real-time detection of driver cognitive distraction using Support Vector Machines.* IEEE Intelligent Transportation Systems, 2007. **8**(2): p. 340-350.
6. Vapnik, V.N., *The nature of statistical learning theory*. 1995, New York: Springer.
7. Saarikoski, *Advances in Natural Language Processing and Applications Research in Computing.* Science, 2008. **33**: p. 107-120.
8. Byun, H. and S.W. Lee, *Applications of Support Vector Machines for Pattern Recognition: A Survey.* Pattern Recognition with Support Vector Machines, First International Workshop, SVM 2000, 2002: p. 213-236.
9. Amari, S. and S. Wu, *Improving support vector machine classifiers by modifying kernel functions.* Neural Networks, 1999. **12**(6): p. 783-789.
10. Quinlan, J.R., *Improved use of continuous attributes in C4.5.* Journal of Artificial Intelligence Research, 1996. **4**: p. 77-90.
11. Lim, T.S., W.Y. Loh, and Y.S. Shih, *A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms.* Machine Learning, 2000. **40**(3): p. 203-228.
12. Freund, Y. and R. Schapire. *Experiments with a new boosting algorithm*. in *Thriteenth International Conference on Machine Learning*. 1996. Bari, Italy.
13. Lee, J.D., et al., *Assessing the Feasibility of Vehicle-Based Sensors to Detect Alcohol Impairment*. 2010, National Highway Traffic Safety Administration: Washington, DC. p. 98.