Crash Risk of Alcohol Involved Driving:
A Case-Control Study

Final Report

September 2005

Dunlap and Associates, Inc.
110 Lenox Avenue
Stamford, Connecticut 06906-2300
203-323-8464
www.dunlapandassociatesinc.com

Foreword

This publication is being distributed by the authors in the interest of information exchange and in support of its goal to reduce mortality and morbidity brought about by misuse of alcohol and drugs by operators of vehicles in all modes of transportation. The report is based on a study funded by Contract DTNH22-94-C-05001 between the U.S. Department of Transportation, National Highway Traffic Safety Administration (NHTSA) and Dunlap and Associates, Inc. The Southern California Research Institute (SCRI) and R.C. Peck and Associates served as subcontractors and were an integral part of the project team. The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of NHTSA.

Crash Risk of Alcohol Involved Driving:
A Case-Control Study

Richard D. Blomberg
Raymond C. Peck*
Herbert Moskowitz**
Marcelline Burns**
Dary Fiorentino**

Final Report

September 2005

Dunlap and Associates, Inc.
110 Lenox Avenue
Stamford, Connecticut 06906-2300
203-323-8464

***R.C. Peck and Associates**
****Southern California Research Institute**

ABSTRACT

A case-control study was conducted in Long Beach, CA and Fort Lauderdale, FL to examine the relative crash risk associated with drivers' blood alcohol concentrations (BACs). Data were obtained for drivers involved in 2,871 crashes of all severities. Two control drivers for each crash driver were sampled a week after the crash at the same location, on the same day of the week and at the same time of day. For both groups of drivers, a research team recorded observations, administered a brief questionnaire and obtained breath specimens for BAC measurements. Of the 14,985 sampled drivers who were potentially available for testing, 91.7% of crash drivers and 97.9% of control drivers provided breath specimens. When drivers who fled the crash scene are included in the number of potentially available drivers, the percentage that provided a breath specimen reduced to 81.3%. Relative risk models were generated with logistic regression techniques with and without covariates such as driver age, gender, marital status and ethnicity. The models without adjustment for the covariates show elevated relative risk beginning at 0.05 – 0.06% BACs with an accelerating increase in risk at BACs greater than 0.10%. With adjustment for covariates and bias due to missing data (non-tested hit-and-run drivers, refusals, and incomplete responses), risk was elevated at a slightly lower BAC and the risk curve was steep. Statistically significant risk occurred at 0.04% BAC and small, non-significant elevations occurred at BACs closer to zero. Relative risk models were also produced for age groups and alcohol consumption levels.

ACKNOWLEDGMENTS

A field study of the complexity and breadth of the current undertaking cannot be accomplished with out the support of numerous individuals and organizations. The authors are grateful for the invaluable inputs made to the definition of the study protocol by the expert panel consisting of:

Dr. Richard E. Allsop
University of London
Centre for Transport Studies

Dr. Paul Gruenewald
Prevention Research Center
Berkeley, CA 94704

Dr. Paul M. Hurst
Research Consultant
Wellington, New Zealand

Captain John Moffat
Seattle Police Department
(currently Director of the Washington Traffic Safety Commission)

Dr. Olga J. Pendleton
Pen-Hock Statistical Consultants

Gordon Stephen Smith, M.D.
Associate Professor
Division of Public Health
Department of Health Policy and Management
Johns Hopkins School of Hygiene and Public Health

Dr. Robert B. Voas
Pacific Institute for Research and Evaluation

Dr. Franklin Zimring
School of Law
University of California

Richard Zylman
Researcher

The data collection phase of the effort was successful largely through the dedicated efforts of the Long Beach, California and Fort Lauderdale, Florida police departments and the locally-recruited field researchers. Specific thanks is directed to:

Long Beach Police Department
Supervisors:
Lieutenant Jack Starbird
Sergeant Rick Colbert

Data Collection Officers:
Gina Bowen-Gilbert, Ray Dennison, Merle McGee, and Edward Reinhold.

Fill-ins:
Corporal Russ Clark, and officers Tyrone Anderson, Robert Mendoza, and Cliff Mifflin.

Monitoring and reporting dispatch activity throughout the project:
Sergeant Hanz Kohnlein.

Research Site Manager for Southern California Research Institute:
Dary Fiorentino

Field Researchers:
Michael King, Steve Kair, Juan Garcia, Mehul Patel, Floyd Allen, Kim Manatt, Debrah Horwath, John Hogan.

Fort Lauderdale Police Department
Michael D. Brasfield, Chief

Supervisors:
Major Bruce G. Roberts
Captain Leonard Schneider
Sergeant Frank Scarpino

Data Collection Officers:
Steve Littlefield, Gregory P. Ridenour, Hector R. Martinez, Mary Gillis, Phil Girrbach, Quintin Waters, Christine Nelson-Lott.

Research Site Manager for Dunlap and Associates, Inc.:
Kristopher T. Korbelak

Field Researchers:
Nicole Vosdoganes, Patty Scarpino, Donna Vaughan, Sherry Acosta, Debbie Ruggiero, Galina Shistik, Alisa Cappuzzello.

National Highway Traffic Safety Administration
The Contracting Officer's Technical Representatives for the National Highway Traffic Safety Administration (NHTSA) were John F. Oates, Jr. and Dr. Paul J. Tremont.

National Institute on Alcohol Abuse and Alcoholism
The Contracting Officer's Technical Representative for the National Institute on Alcohol Abuse and Alcoholism (NIAAA) was Dr. Susan E. Martin.

In addition the authors express their sincere thanks to participating project staff members from the involved organizations:

Dunlap and Associates, Inc.
Richard D. Blomberg was the Dunlap and Associates, Inc. project director and the prime contract director for the study.

Edward W. Bishop, John W. Hamilton and Paul Turcotte helped design the data collection protocol and select the field equipment.

Mark D. Lardner managed the key entry of the data and the resulting database as well as relieving as Fort Lauderdale site manager.

Kristopher T. Korbelak, in addition to managing the Fort Lauderdale site, participated in the data management and report preparation.

R. C. Peck and Associates
Raymond C. Peck directed the R.C. Peck and Associates subcontract effort and led the data analysis task.

Michael Gebers was responsible for preparing all of the SAS® and SPSS® code for executing the inferential analyses and assisted in the preparation of the technical results section of this report.

Debbie McKenzie and Ellie Peck assisted in the preparation of the analysis outputs and report.

Southern California Research Institute (SCRI)
Dr. Herbert Moskowitz directed the SCRI subcontract effort and was principal investigator.

Dr. Marcelline Burns and Dary Fiorentino managed SCRI's efforts and played major roles in the design of the questionnaire and field data collection forms.

Although this effort clearly could not have been successful without the aid of these cited individuals, numerous others associated with each study site and the approximately 14,000 drivers who agreed to be interviewed, the findings and conclusions herein are solely the responsibility of the authors.

This Page Intentionally Left Blank

Table of Contents

List of Tables

List of Figures

This Page Intentionally Left Blank

EXECUTIVE SUMMARY

This case-control study examined the relative crash risks associated with drivers' blood alcohol concentrations (BACs). The measure, relative crash risk, is defined as the ratio of the proportion of crash drivers to the proportion of control drivers in a BAC classification, compared to a similarly formed ratio of drivers with 0.00% BACs.

The specific objectives of the study were to:

- Determine the relative crash risk of drivers at various BACs compared to drivers with zero BACs while controlling for other factors (e.g., age, gender, drinking patterns)
- Determine the relative crash risk of groups of drivers (e.g., youth, males, heavy drinkers) at various BACs compared to similar groups with zero BACs.

Background.

A relative risk function derived from epidemiological data collected in Grand Rapids, Michigan during the 1960's contributed to the current legal limits for driving (Borkenstein et al., 1964). It is possible, however, that changes in driving, alcohol consumption, or drinking-driving practices have occurred during the decades since that study was conducted, and that the changes altered the relative risks associated with the combination of alcohol and driving. Regulatory agencies and legislators need current information about the crash risks associated with driving at BACs above zero. A study to update the epidemiological evidence was conducted with currently available measurement apparatus, robust experimental designs, and advanced multivariate statistical methods.

Research Approach.

- During 12-month periods, data were collected in two locations to acquire samples of $\geq$ 1,300 crashes at each site.
- Data were obtained from crash-involved and non-crash-involved (control) drivers during the hours 1600 - 0200 in Long Beach, California and 1700 - 0300 in Fort Lauderdale, Florida. The evening and nighttime hours were selected for sampling, because drivers who have been drinking are most likely to be on the roadways at those times.
- A two-person research team, a police officer and an interviewer, collected the data from crash and control drivers.
- Drivers involved in crashes and non-crash involved control drivers were matched for crash location, day, time, and travel direction. Control drivers were obtained one week after a crash at the same site, traveling in the same direction, on the same day of the week and at the same time of day. Using random selection procedures, drivers in the flow of traffic were stopped and asked to participate in the study.

- Interviews with crash and control drivers were conducted with a standard questionnaire to obtain data about the covariates of drinking and driving.

- Drivers' BACs were measured at roadside by obtaining breath specimens with evidential quality portable/preliminary breath testers (PBTs). Hospitalized and fatally-injured drivers' BACs were obtained from the analyses of blood specimens by police, hospitals, and coroners.

- Systematic observations by the research teams, together with BAC estimates obtained with passive alcohol sensors (PAS), provided additional information about drivers' use of alcohol. These data served as the basis for estimates of the bias resulting from the refusal of some drivers to provide breath specimens.

- A complete, single case for the statistical analysis of BAC and crash risk included a crash driver and two control drivers.

Statistical Analysis.

Relative risk curves describing the relationship between BAC and crash risk were derived from a sequence of univariate and multivariate logistic regression analyses. After establishing simple univariate models, more complex models were derived in which the relative risk estimates were adjusted for potentially confounding covariates (e.g., age, education) and other sources of bias, such as refusal to provide a BAC specimen and non-recovered hit and run cases. Most of the models were based on the total sample and all crashes but separate analyses were also done for the two sites (Long Beach and Fort Lauderdale), type of crash and other subgroups of interest (e.g., age and alcohol consumption patterns). The statistical analyses produced answers to the following questions:

- How does crash risk increase as a function of a driver's BAC?

- Is the relationship between BAC and crash risk materially altered by controlling potential sources of bias and confounding covariates?

- Is the BAC-crash relative risk curve similar for different crash types?

- Is the BAC-crash relative risk curve similar for the two study sites (Long Beach and Fort Lauderdale)?

- Does the BAC-crash relative risk curve vary by driver characteristics, such as age, gender and alcohol consumption patterns?

- How accurately can individual crash involvement be predicted from knowledge of a driver's BAC and other characteristics?

Results.

A total of 2,871crashes (Long Beach 1,419; Fort Lauderdale 1,452) yielded 4,919 crash and 10,066 control drivers (14,985 total). In total, 93.5% of the drivers who were contacted at crash scenes agreed to participate. An additional 603 fled the scene of their crash. One hundred four of those hit-and-run drivers were located within two hours of the crash, and 94 (90.4%) provided breath specimens. Those who were not located or refused to participate reduced the participation rate of crash-involved drivers to 83.1% and the percentage who provided usable breath specimens to 81.3%. Non-crash drivers participated as controls at a higher rate of 97.9%.

Since the data obtained in California and Florida were similar, most of the logistic regression analyses were performed for the total data set. The analyses showed elevated relative risk with increasing BAC and a strongly accelerated risk at BACs greater than 0.10%. The influence of covariates on the magnitude and shape of the curve was relatively modest. Underestimates of crash risk resulted from biases, which arose from three sources:

- Differences in non-participation rates between the crash and control groups–Alcohol positive crash drivers refused to participate and failed to complete the interview more often than other drivers

- Non-apprehended hit-and-run drivers–BACs of apprehended hit-and-run drivers were much higher than BACs of drivers who did not flee.

- Missing covariate data due to non-participation and incomplete interviews.

Re-weighting adjustments based on observational data and PAS estimates of BACs were used to correct for these biases. The resulting adjusted model of relative risk showed greater risks at all BACs with large increases at high BACs.

Table ES-1 on the following page displays calculated relative risks from: a model with no covariates (column 2), a model with demographic covariates (column 3), and a model adjusted for all three major biases (column 4). Column 5 is from a model developed by Allsop (1966) for the Grand Rapids Study data and is comparable to the no-covariate model for this study.

Logistic regression analyses were performed for subgroups of the driving population (e.g., youth, heavy drinkers), but each subgroup was small. It is unclear whether the small samples account for unexpected findings such as the absence of increased risk for young drivers at low BACs. Although the issue cannot be resolved without larger numbers in the subgroups, the absence of an Age x BAC interaction suggests that the effects of age and BAC on crash risk are largely an additive function.

The study data demonstrate that case-control studies need to control or adjust for differential non-participation and non-random missing data, particularly for the loss of data from hit-and-run drivers. The effect of these sources is evident in a comparison of the third and fourth columns of Table ES-1. Note the magnitude of the underestimation of relative risk at very high BACs. For example, the unadjusted relative risk of 26.60 at BACs $\geq$ 0.25% becomes 153.68 when it is adjusted for these biases.

Table ES-1. Relative Risk Models

	Relative Risk			
BAC	**No Covariates**	**Non-reactive Demographic Covariates**	**Final Adjusted Estimate**	**Grand Rapids***
0.00	1.00	1.00	1.00	1.00
.01	.91	.94	1.03	.92
.02	.87	.92	1.03	.96
.03	.87	.94	1.06	.80
.04	.92	1.00	1.18	1.08
.05	1.00	1.10	1.38	1.21
.06	1.13	1.25	1.63	1.41
.07	1.32	1.46	2.09	1.52
.08	1.57	1.74	2.69	1.88
.09	1.92	2.12	3.54	1.95
.10	2.37	2.62	4.79	5.93
.11	2.98	3.28	6.41	
.12	3.77	4.14	8.90	4.94
.13	4.78	5.23	12.06	
.14	6.05	6.60	16.36	10.44
.15	7.61	8.31	22.10	
.16	9.48	10.35	29.48	21.38
.17	11.64	12.74	39.05	
.18	14.00	15.43	50.99	
.19	16.45	18.31	65.32	
.20	18.78	21.20	81.79	
.21	20.74	23.85	99.78	
.22	22.07	25.99	117.72	
.23	22.51	27.30	134.26	
.24	21.92	27.55	146.90	
.25+	20.29	26.60	153.68	

*From reporting of Grand Rapids Study data in Table 25 (a) of Allsop (1966).

When adjusted for non-participation and missing data bias, the results suggest that the small dip observed at 0.01– 0.03% BAC in the unadjusted risk calculations (columns 2 or 3 in the table above) may be an artifact of sampling errors and small biases. The magnitude of the adjusted risk elevations are too small in relation to the standard errors of the model, however, to reject the hypothesis of no increase or even slight decreases in risk at 0.01 – 0.03% BAC. Regardless of the direction of the risk change at these low BACs, however, the size of the relative risk deviation from unity is sufficiently small to be of no practical consequence.

In a re-analysis of the Grand Rapids data, Hurst, Harte and Frith (1994) also showed that the relative risk curve changed substantially and a decrease or dip in risk at low BACs disappeared with an adjustment for drivers' drinking frequency. Although the present study failed to replicate the phenomenon and artifact noted by Hurst et al., the present analysis, as discussed above, did show the existence of a dip in the model with no covariates that disappeared when adjusted for the sources of bias discussed above.

In conclusion, the study results confirm a notable dose-response relationship beginning at 0.04% BAC and increasing exponentially at ≥ 0.10% BAC. The final adjusted relative risk function (column 4 in the table) is graphed in Figure ES-1 below. It illustrates the extraordinary magnitude of the crash risk at high BACs.

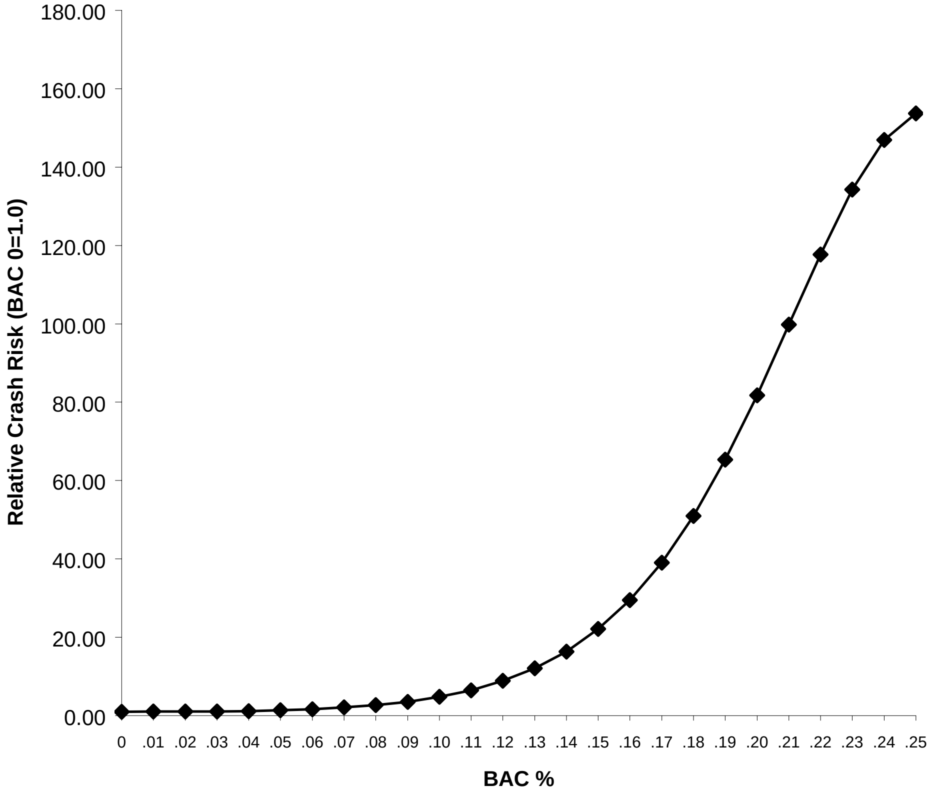

Figure ES-1. Final Adjusted Relative Risk Estimate

This Page Intentionally Left Blank

1. INTRODUCTION

1.1 Study Rationale

Alcohol impairment of driving skills, which was identified as a traffic safety problem early in the 20th century, continues to be a major highway safety issue. Although traffic fatalities decreased 30% during the period 1989 - 1999, the numbers remain unacceptably large. Data from roadside surveys and accident investigations, together with findings from controlled laboratory experiments, demonstrate that many human performance skills, including skills essential for safe driving, begin to decline at blood alcohol concentrations (BACs) very close to zero. The risk of crash increases as a driver's BAC increases. NHTSA (2004) reported that 15,626 traffic fatalities (41% of total) during 2002 were alcohol-related, and 32% of the fatally injured drivers had BACs that were $\geq$0.08%.

How and to what extent does alcohol elevate the risk of crash? At what BAC does significant impairment of driving skills begin? Do driver characteristics, such as age, gender, driving experience, and drinking history influence the magnitude of the alcohol-related deficits? These questions illustrate the type of information that will assist policy makers in making sound legislative decisions about BAC limits for drivers, special provisions or restrictions for youthful and aging drivers, and enforcement policies.

A comprehensive assessment of the crash risk associated with alcohol requires both experimental and epidemiological data. Data from controlled laboratory experiments explicate the mechanisms of alcohol- related changes in performance. Epidemiological data that compare BACs of crash-involved and non-crash-involved drivers are necessary to interpret laboratory findings in term of crash risk.

Epidemiological data from a landmark study conducted by Borkenstein et al. (1964) have been used in the assessment of crash risk due to alcohol for almost four decades. Since those data were collected, however, driving and drinking environments have changed, and the changes possibly have altered the risk of a crash. This study, using modern apparatus for BAC measurement, advances in statistical methods, and a robust experimental design, was undertaken to determine whether extant epidemiological data remain valid.

1.2 Study Objectives

This study examined the relative risk of crash associated with drivers' BACs. Relative risk is defined as the ratio of the percentage of crash drivers to the percentage of control (non-crash) drivers in a BAC classification, compared to a similarly formed ratio of drivers with zero BACs.

Study objectives were to quantify:

- The relative crash risk of drivers at various BACs compared to drivers with zero BAC while controlling for covariates such as age, gender, and drinking patterns. Covariates can introduce a spurious, non-causal association due to their concomitant association with BAC.

- The relative crash risk of driver subgroups such as youth, males, and heavy drinkers at various BACs compared to such groups at zero BAC.

- The relative crash risk associated with BACs below 0.05%.

The study was conducted at two sites where samples of ≥1,300 crashes could be obtained. Multiple sites facilitated the large data collection effort and provided the potential for within-study replication of results.

2. BACKGROUND

A literature review and input from a panel of experts helped guide the development of this study. This section reviews the background literature with respect to some of the key issues that were germane to the choice of methods. A more extensive discussion of generic issues appears in Appendix A.

2.1 Very Early Studies

The first discussion of a relationship between alcohol consumption and motor vehicle collisions to be published in an American scientific journal appeared as an editorial in the Quarterly Journal of Inebriation (1904). The editor had received a communication about 25 fatal crashes of automobile wagons in which 23 occupants died and 14 suffered injuries. Nineteen of the drivers had used alcohol within an hour of the crash. The author of the communication commented that driving automobile wagons was a more dangerous activity for drinkers than driving locomotives. Drinking by on-duty railroad employees had been prohibited since 1843 (Borkenstein, 1985).

Although the editorial suggested a relationship between alcohol and crashes, it lacked quantitative analysis. In fact, quantification of the role of alcohol and a determination of the number of alcohol-related crashes did not occur for another half century.

2.2 Evanston, Indiana Study of Injury Crashes

Holcomb (1938) published the first study in which the BACs of crash and non-crash drivers were compared. Under the leadership of Herman Heise, urine specimens had been collected over a three-year period from 270 hospitalized drivers. Forty six percent of the specimens contained alcohol, and almost 14% of the drivers had been over 0.15% BAC. Holcomb pointed out, however, that it was theoretically possible that 46% of the entire driving population would have had similar alcohol levels, and interpretation of the finding would require a comparison group. To establish such a group, breath specimens were obtained from 1,750 drivers using the Drunkometer, a recently developed instrument. Drivers were sampled during evening hours in eight locations, half of which were near taverns. Age, gender, time of day and day of the week data were collected for both hospitalized and control drivers. Although variations of these factors with BAC were examined, comparisons of the two groups did not adjust for them. Alcohol was present in 12% of the 1,750 control drivers. Two percent of the group was above 0.10% BAC and 0.4% was above 0.15%.

Although alcohol appeared to be highly over-represented in injured drivers, the crash data collection was not protected from bias, and the two groups differed by age, gender, day of the week and hour of the day. Holcomb did not create a relative risk curve or calculate the relationship of BAC and crash probability. In a later analysis, Hurst (1971) used ancillary data from Holcomb, corrected for time of day, and estimated relative risk of an injury crash as 3.0 at 0.06%, 4.0 at 0.09% and 10.0 at 0.12% BACs.

2.3 Toronto Study of all Crashes

Noting that the data collected by Holcomb "did not allow a statistically valid comparison" between crash and non-crash drivers, Lucas, et al. (1955), emphasizing that traffic conditions for the two driver groups had to be as closely alike as possible, improved both methodological procedures and data analysis in a Toronto study.

Between the hours 1830 – 2230, Monday – Saturday, during the period December 1951 – November 1952, police and research personnel went to as many traffic collisions as possible to obtain breath specimens. For each crash driver, at least four non-crash drivers of cars of the same vintage passing the crash scene at approximately the same time were asked for breath specimens. All drivers were asked four questions and were given a questionnaire to complete at a later time. Sixty percent returned the questionnaire.

The investigators calculated estimates of relative crash risk as a function of BAC with data from 433 crash and 2,015 control drivers. Instead of using the frequency of crashes for drivers at zero BAC as the referent group for the analysis, however, they used the frequency for the 0.00% to 0.05% BAC interval. Since this changed the relative risk curves vis-à-vis a 0.00% base, a realistic estimate of the relative risk produced by alcohol was not obtained.

Estimates of relative risk from Holcomb's data are significantly higher than those from Lucas et al. Several variables may have contributed to the difference. Notably, Lucas, et al. sampled drivers involved in *crashes* whereas Holcomb sampled drivers involved in *injury crashes*. In the Holcomb study, BACs of injured drivers were measured with urine specimens, and BACs of control drivers were measured with breath specimens. In the Lucas, et al. study. BACs of both crash and control drivers were measured with breath specimens. In both studies, the specimens were collected in plastic bags for analysis at a later time, a method known to increase variability.

Lucas et al. (1955) also obtained four *or more* drivers for each crash driver, which over- represented some sites and may have skewed the influence of the crash site. Despite its limitations, however, the study was a methodological advance that produced the first estimate of crash probability for all crashes as a function of BAC.

2.4 New York City Study of Fatal Crashes

McCarroll and Haddon (1962) conducted a case-controlled study in New York City during June – October in 1959 and 1960. Crash drivers were 43 fatally injured passenger car drivers. Data were obtained from police reports and coroner reports. In 1960, six control drivers (258 total) for each crash driver were sampled at the crash sites on the same day of the week at the same time of day and within a few weeks of the calendar week of the crash. The drivers responded to questions and provided breath specimens for alcohol analysis. Missing information for a fatally injured driver resulted in the deletion of the driver and his/her controls from analyses of the missing item.

Bivariate comparisons of interview data and coroner/police data were performed with the BAC distributions. The investigators concluded that alcohol was associated with both fatal crash probability and crash responsibility. Covariate data were not used to adjust the

BAC – fatal crash probability, possibly because the cell categories would have been too small for statistical analysis. Nonetheless, the use of a matched control group for controlling exposure and bias represented a methodological advance.

The investigators did not calculate relative risk, but the data can be used for that purpose. If the frequency of crash and non-crash drivers at 0.00 BAC is set as a base comparison, the calculations yield the estimates that appear in the table below. Extremely high BACs were heavily over-represented with 75% of the fatalities at BACs $\geq$0.250%. In comparison, only 16% of fatalities nationally were that high during the years 1979 – 1990. (Zobeck, et al., 1992)

BAC	Relative risk estimate
0.020 - 0.099%	1.04
0.100 - 0.249%	2.62
0.250 - 0.399%	>176.8

2.5 Grand Rapids, Michigan Study

At the time it was conducted, the Borkenstein et al. study (1964, 1974) was the largest controlled study of alcohol-involved collisions, and its findings contributed to alcohol and driving policy decisions for several decades. A number of methodological characteristics and statistical issues of that study were particularly important to the design of the present effort including:

- The investigators attempted to sample all crashes in Grand Rapids, Michigan between July 1, 1962 and June 30, 1963. During that period, police reported crashes involving 9,353 drivers. The police department did not respond to 2,764 minor crashes, but the drivers were instructed to submit a delayed report. Haddon (1964) suggested that this procedure may have produced a bias against low BACs.

- The Borkenstein et al. study was not a matched case-control study. The control group (8,008 drivers) was created by sampling four drivers at 2,000 crash sites selected at random from a pool of 27,000 crashes that had occurred during the three preceding years. Four control drivers were interviewed at all sites regardless of the number of drivers involved in the crashes. They were sampled at the time of day and day of the week of the selected prior crash, rather than at the time and day of the collision in which the crash driver was involved. The direction of traffic for sampling control drivers was randomly chosen. In total, the BACs of 5,985 collision drivers and 7,590 control drivers were measured.

- Although the research teams attempted to obtain breath specimens and interviews with all drivers, some drivers in both groups refused to cooperate, some crash drivers fled, and some potential control drivers avoided contact. During peak crash periods, it was not possible to go to all crashes.

Statistical analyses of the Grand Rapids data and the study conclusions have been closely scrutinized. It is important to note that the authors did not compute the relative risk of *involvement* in a crash as a function of BAC. Rather, they generated a figure purported to indicate the relative risk of *causing* a crash as a function of BAC. To accomplish this, they created a BAC distribution of crash-causing drivers.

The causation group was created in several steps. First, it was assumed that the 622 drivers involved in single vehicle crashes were responsible for the crashes, and all were included in the crash-causing group. Their BACs were known from breath measurements at the crash scene. Second, it was assumed that half of 5,366 drivers (2,683) involved in multiple vehicle crashes caused the crashes and that half were not at fault. Third, the BACs of 2,683 not at-fault drivers and BACs of control drivers were assumed to be identical. The BAC distribution for 2,683 at-fault multiple vehicle crash drivers was created by subtracting the BAC distribution of not-at-fault drivers from the BAC distribution of all multiple-vehicle crash drivers. Fourth, a total "at fault" driver BAC distribution was created by adding the BAC distributions of the 622 single vehicle crash drivers and 2,683 multiple vehicle at-fault drivers, creating a total of 3,305 at-fault drivers. Finally, a comparison of the BAC distributions of at-fault drivers and the control group created Chart XV of the Grand Rapids Study (Borkenstein, et al., 1974), the "relative probability of causing an accident."

The creation of the crash group only from the distribution of at-fault drivers resulted in a larger proportion of drivers with alcohol than contained in the total group of crash-involved drivers. A review of the Grand Rapids Study, reported in a British government monograph (Allsop, 1966), includes a relative risk calculation for crash involvement as a function of BAC. The relative risk curve of drivers "causing" a crash (Borkenstein, et al., 1974, Chart XV) rises at roughly double the rate for the complete, original sampled group (Allsop, 1966, Table 1). The estimates of the reductions in total crashes that would be achieved by preventing driving at various BACs are, therefore, roughly triple the estimates by Allsop (compare Borkenstein et al., 1974, Chart XV with Allsop, 1966, Figure 3).

The Borkenstein causation curve shows the relative risk at 0.01% - 0.04% BACs to be less than the risk at zero BAC. This Grand Rapids dip, as it is known, has generated controversy and analysis. Several explanations for the observed risk pattern are possible. The dip could be an artifact of the crash/control univariate comparison, which did not control for various crash-determining covariates. In several analyses with bivariate techniques, the dip has been eliminated. For example, in an analysis of BAC, drinking frequency and relative crash involvement risk, the dip disappeared (Allsop, 1966).

By gathering interview data from both driver groups, the Grand Rapids investigators made a valuable contribution to the study of the risk associated with drivers' use of alcohol. The obtained information included BAC, age, gender, educational level, ethnicity, marital status, occupation and drinking frequency. The finding from univariate analyses that crash and control groups differed on nine of these variables demonstrates that they are associated with crash involvement. It must be remembered, however, when reviewing the results of the Grand Rapids Study that the relative risk of *involvement* in a crash as a function of BAC was not computed. Rather, the study generated a figure that displays the relative risk of *causing* a crash as a function of BAC. Among the limitations of this approach is that fault determinations are subject to bias and unreliability.

2.6 Other Studies

None of several studies conducted since the Grand Rapids Study has been a true case-control effort. Although relative risk analysis was performed for data obtained in a series of roadside surveys (Biecheler, 1970, 1971, 1974, 1983), the crash and control drivers were not matched, and it was not a case-control study. The Vermont Study (Perrine, M. W., Waller, A. J. and Harris, L. S., 1971) focused on 113 fatally injured drivers. Since control subjects were obtained at sites and times other than where the 113 fatalities occurred, however, it was also not a true case-control design.

A study in Adelaide, South Australia by McLean, Holubowycz and Sandow (1980) was quasi-case-controlled. Injured drivers, passengers or pedestrians formed the crash group. Control drivers were sampled two years later, matching the crash driver's age and sex, the time, day and week of the crash, crash site, and travel direction. The relative risk curve derived from the data fell between the curves for the Grand Rapids Study of all crashes and the Vermont Study of fatally injured drivers.

2.7 Lessons from the Literature

The review of the previous studies described above and in Appendix A identified the following necessary characteristics for the design of the present study:

- Employ a true case-control design with the control group matched as closely as possible to the crash group.

- Obtain quantitative BAC data for both crash and control groups.

- Obtain as many covariates of drinking and driving behavior as possible with standardized questions and observations.

- Devise procedures to minimize drivers' refusals to respond to questions or provide breath specimens.

- For refusing drivers, collect as much relevant data as possible, such as Passive Alcohol Sensor (PAS) readings and observations, to estimate the bias produced by their non-participation.

- Obtain as much information as possible for drivers who flee the scene of crashes to minimize the potential for bias as a result of missing data.

- Use a multivariate analytical approach in the assessment of relative risk. Include covariates and use appropriate reference groups for crash versus control comparisons.

This Page Intentionally left Blank

3. METHOD

This section describes the study approach, methods of data acquisition, selection of study sites, and methods of data management and analysis that were developed to insure valid and reliable data.

3.1 Approach

The study approach consisted of the following elements:

- Data were collected from two study sites each of which could yield a sample of at least 1,300 crashes. The project team selected one site in the east to be managed by Dunlap and Associates, Inc. and one in the west to be managed by the Southern California Research Institute (SCRI). Sampling for a minimum of 12 months at each site precluded possible seasonal biases.

- Crashes were the sampling unit. The occurrence of a crash determined the drivers that were the focus of data collection for both the crash and control groups.

- Drivers were the unit of analysis. Analysis of relative risk was on a driver-by-driver basis to allow calculations of relative risk at various BACs.

- The study design was a true case-control paradigm. The crash-involved and control drivers from each sampled crash were matched as closely as possible on covariates such as location, time of day, day of week and direction of travel.

- Two-person research teams, each including a police officer and a researcher, operating from a police patrol car collected the data.

- Specimens for BAC measurements were obtained with evidential quality portable/preliminary breath testers (PBTs) and from police, hospital, and coroner analyses of blood specimens.

- Information collected by standard observations and PAS readings created a database on all potential subjects whether or not they agreed to participate. These data were used to assess bias produced by refusals.

- A brief questionnaire administered verbally by the interviewers obtained data from crash and control drivers on covariates of drinking and driving.

After the basic approach had been defined, an expert panel was assembled to assess the research team’s preliminary decisions. The panel consisted of nine preeminent researchers with experience in similar studies, field data collection and analogous data analysis tasks. The panel met on January 26-28, 1995 in Washington, D.C. and generated valuable guidance that was incorporated into the final study design.

The study team, with input from the NHTSA and NIAAA Contracting Officer's Technical Representatives (COTRs), finalized the data collection, management and analysis protocols. The protocols were pre-tested to the extent possible prior to the initiation of data collection. Project staff monitored data collection throughout the project to insure the consistent application of the protocols.

3.2 Site Selection

Study objectives did not include assessment of the incidence of alcohol in crash-involved drivers. Rather, the study data were to be collected at two sites solely for the purpose of determining the relative risk of a crash as a function of the involved drivers' BACs. Since drivers on the roadway during evening and weekend hours are most likely to have been drinking, crash statistics from potential sites were examined to determine whether the required minimum of 1,300 crashes could be sampled during those times. To facilitate project management and quality control of data, site selection focused on jurisdictions with a single police agency.

Population served as a reasonable surrogate for number of crashes for the western site search. A California Highway Patrol (CHP) database provided crash statistics for eight cities with populations of 200,000 or more. From those data, it was determined that a population of 400,000 would provide a sufficient number of crashes.

For the eastern site search, population proved not to be a reliable predictor of crash frequency. Actual crash data obtained from possible sites in several states were therefore scrutinized to assess whether a sufficient number of evening-hour crashes could be expected to occur during the study period.

The selection process in both areas also considered roadway characteristics. Since it would have been unsafe to conduct study activities on freeways or other high-speed roadways, it was necessary to find a jurisdiction in which the criterion number of crashes would occur on city streets patrolled by the cooperating local police agency.

A final site selection criterion was the willingness of law enforcement agencies to commit staff and equipment to the study. Although the agencies were reimbursed for project-related costs, they were required to guarantee the availability of officers, vehicles, and dispatch services, allow full access to crash-related documents, and provide office space for an on-site coordinator.

Selection activities moved most quickly in California. The police department in the City of Long Beach (LBPD) met all study criteria and accepted an invitation to participate in the study. Data collection began on June 21, 1997 and ended on September 10, 1998.

Several locales in the eastern part of the country were examined. Fort Lauderdale, Florida met study criteria, and the Fort Lauderdale Police Department (FLPD) was willing to participate. Sampling began on September 5, 1998 and ended on September 4, 1999.

Table 1 summarizes the basic characteristics of the two study sites. The two cities are similar in area, both are beach cities with warm climates, both have substantial industry, and the median incomes are similar. The population of Long Beach is significantly higher than the permanent population of Fort Lauderdale, although the latter has many winter-only residents. The per capita *rate* of crashes is higher in Fort Lauderdale due to the larger number of crashes and its smaller population, This difference between sites is believed to be due at least in part to different reporting limits and investigation criteria as well as to a higher flow of non-resident traffic through Ft. Lauderdale. The sample acquisition rate in Fort Lauderdale was somewhat higher than in Long Beach.

Table 1. Characteristics of Study Sites

Characteristic	Long Beach	Fort Lauderdale
Population (2000 census)	430,298	152,397
Land Area	50.0 mi² (129.5 km²)	31.4 mi² (81.2 km²)
Median Income (HUD 1998 Estimate)	$49,800	$49,200
Total Crashes	6,020 (1997) (LBPD)	10,058 (7/1/97-6/30/98) (FLPD)
Sampling Start	6/21/97	9/5/98
Sampling End	9/10/98	9/4/99

3.3 Questionnaire Development

A questionnaire was developed to obtain information from both crash and control drivers about their drinking practices and the major correlates of alcohol-involved crashes. Its content was drawn from an extensive review of the alcohol survey literature and the experiences of previous alcohol and highway safety studies.

The Office of Management and Budget (OMB) reviewed the questionnaire, as required, and advised that it should not impose an unreasonable time burden on participants. To comply with the time constraint, as well as to minimize time-related refusals by drivers, the questions were limited to insure that the interview could be completed in five minutes or less.

The questionnaire was developed in English and translated into Spanish, Cambodian and Creole. Although there is a sizable Creole-speaking population in Fort Lauderdale, that version of the questionnaire was never needed. The Cambodian version was used eight times in Long Beach. The Spanish version was used for 133 interviews with 121of those in Long Beach.

Appendix B presents the final questionnaire containing 22 questions and two observations. The “B Code” observation on the first page assessed behavioral signs of alcohol use (N=No Suspicion, S=Suspicion, O=Obvious Signs). The “C Code” observation on the last page assessed the extent of communication problems (N=none; M=moderate; S=severe). It was anticipated that drivers would become increasingly cooperative as interviewer-driver rapport developed during the interview; for that reason, sensitive items such as the request for a breath specimen for BAC measurement were placed at the end of the interview.

Each question's topic is described below.

1. Driving variables are strong predictors of crashes. Pre-testing indicated that people were comfortable reporting an average weekly mileage, and they believed they could do so accurately. Other relevant questions (miles driven, crash history, citations) were not asked due to the time limit for the entire questionnaire.

2. Trip origin was included at the suggestion of Dr. Paul Gruenewald, a member of the project's expert panel. He reported that research by the Prevention Research Center had found trip origin to be a reliable surrogate BAC measure.

3. Age data were obtained for the assessment of relative risk by age. Drivers' birth dates provide more precision than numeric ages.

4. – 9. Marital status, education, employment status, occupation, and ethnicity have been shown to correlate with the probability of crash both with and without alcohol. Item 7 allowed the time of crash or control sampling to be related to the driver's workday.

10-12. Current sleep status and cumulative sleep deficit variables are related to a possible alcohol/sleep deprivation interaction, as suggested by data from studies conducted during the last decade.

13. A comparison of drivers' reports of alcohol use within two hours with actual BACs provides a measure of the drivers' veracity.

14. Alcohol use within 30 minutes of a breath test can leave residual mouth alcohol and cause erroneous readings. Responses to the question enabled interviewers to assess the likelihood that mouth alcohol contributed to the measured BAC.

15. Alcohol beverage preferences correlate somewhat with extent of alcohol use. This non-threatening question also served to initiate the inquiries about drinking behaviors.

16. Alcohol use in the past 28 days, which has been validated as a quantity-frequency measure of alcohol use (Gruenewald and Nephew, 1994), served as a parsimonious method for obtaining data for estimates of tolerance to alcohol effects.

17. Prior DUI arrest information was obtained for an examination of the relationship between prior arrests and subsequent alcohol-involved crashes.

18. Driving within two hours of drinking, as reported by drivers, was viewed as an index of aberrant drinking and driving behavior.

19. Many drugs, both licit and illicit, impair driving skills. Study objectives required information about whether a crash or control driver may have been drug-impaired.

20. Breath specimens for quantitative measurement of BAC[1] provided key data.

21. A second test was requested if the driver did not give a sufficient breath specimen or if the PBT malfunctioned. Occasionally, the test was repeated for quality control.

22. Impaired drivers were not allowed to drive. They were transported to their residences by taxi or a sober passenger, or were escorted by project staff. A record of the actual method was added for the Fort Lauderdale sampling to provide more process information.

3.4 Data Collection

A complete case consisted of data collected after a crash that occurred within the LBPD or FLPD patrol area together with the data from non-crash-involved drivers passing the crash site at the same time of day one week later.[2] Research teams included a sworn officer and a researcher/interviewer who traveled together by police car to the scene of traffic crashes. On scene, the officers made initial contact with drivers and requested their cooperation with the research. The interviewers administered the questionnaire and collected breath specimens with a PBT. The interviews were conducted privately.

A research team (generally the same team that sampled a crash) returned to the site of the crash one week later on the same day of the week and at the same hour of the day. To obtain control data, they stopped two non-crash-involved drivers for each crash driver, administered the same questionnaire, and obtained breath specimens with a PBT. Throughout the research, the teams' priorities in interactions with citizens at crash and control sites were safety, privacy, data accuracy, and data confidentiality.

As mentioned earlier, the focus of the study was relative risk due to alcohol, not the incidence of alcohol-related crashes. Thus, there was no reason to sample crashes during most daytime hours when alcohol use by drivers is low. To conserve resources, sampling times were limited to evening and nighttime hours when the proportion of crashes involving alcohol is highest.

Data collection was scheduled in Long Beach for regular LBPD shift hours. Officers work four 10-hour shifts per week from 4:00 PM (1600) to 2:00 AM (0200). Five research teams worked 17 weekly ten-hour shifts. LBPD supervisors assigned, scheduled, and supervised the officer members of the teams. Four officers (three men, one woman) were assigned full-time as Teams 1 to 4. Team 5 officers were drawn from the LBPD DUI unit. Three teams collected data on Thursday, Friday, and Saturday nights, the days with the most alcohol related crashes. Two teams worked on the remaining days. Sampling was suspended on major holidays. The weekly schedule appears in Table 2.

SCRI hired, trained, scheduled, and supervised the Long Beach field interviewers who collected data working with LBPD officers as research teams. Four college students

---

[1] Project breath tests were made with the Alco-Sensor IV® Portable Breath Tester (PBT) produced by Intoximeters, Inc. This unit is highly accurate and produces readings of evidential quality.

[2] Control sampling was scheduled whenever possible for the same day of week one week after the crash. This target was achieved for the vast majority of cases.

(two men, two women) served as full time interviewers. A fifth man served as a member of Team 5 and also as Site Coordinator. In the latter position, he scheduled interviewers and on a daily basis checked all data records for accuracy and completeness. An on-call interviewer filled in when regular interviewers were absent. Both of the women and one man in the initial group of interviewers resigned and were replaced during data collection. One of the men hired as a replacement held an undergraduate degree, and the second man was completing his last year of college. All interviewers spoke English, one spoke Spanish, and another spoke Hindi.

Table 2. Long Beach Teams Data Collection Schedule

Team	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1				1600 to 0200	1600 to 0200	1600 to 0200	1600 to 0200
2	1600 to 0200	1600 to 0200				1600 to 0200	1600 to 0200
3			1600 to 0200	1600 to 0200	1600 to 0200	1600 to 0200	
4		1600 to 0200	1600 to 0200	1600 to 0200	1600 to 0200		
5	1600 to 0200						

FLPD supervisors assigned, scheduled, and supervised the police officers assigned to the project. The officers were dedicated to the project, worked from 5:00 PM (1700) to 3:00 AM (0300), and responded to routine police calls only in the event the department experienced a work overload or they happened to be the closest to the scene. Seven officers (five men, two woman) were selected for the project. Five officers worked 17 weekly 10-hour shifts. Four were assigned full time to the study on teams 1 to 4, and the fifth was a full-time backup. He was available at any time to respond to a crash or to join the research site coordinator to conduct control sampling when the other teams were tied up. The two additional officers were backups and also served two times each month on a third Wednesday night team. These regularly scheduled shifts kept the alternates up to date on research procedures. A minimum of two teams collected data on all study days with three teams deployed on Wednesday. Sampling was suspended on major holidays. The weekly Fort Lauderdale schedule appears in Table 3.

Table 3. Fort Lauderdale Teams Data Collection Schedule

Team	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1			1700 to 0300	1700 to 0300	1700 to 0300	1700 to 0300	
2				1700 to 0300	1700 to 0300	1700 to 0300	1700 to 0300
3	1700 to 0300	1700 to 0300			1700 to 0300	1700 to 0300	
4	1700 to 0300	1700 to 0300	1700 to 0300				1700 to 0300
5			1700 to 0300				

Initially, four individuals (all women) were recruited and trained as full-time field interviewers to work four shifts per week. All had other jobs or were students, and the schedule proved to be too demanding. Therefore, three additional researchers (also all women) were trained as researchers.[3] They worked several regular shifts and served as backups, making the workload manageable. Most were graduate students and medical students, some had military experience, and others had used breath alcohol instruments during work for the police department. All spoke English, one was fluent in Russian, and two were fluent in Spanish. All seven remained with the project throughout data collection.

A full-time Dunlap staff member was the research site manager. He scheduled all teams, reviewed incoming data to ensure completeness and accuracy, and often filled in as a field researcher.

3.5 Crash Sampling

A case entered the study when a research teams reached the crash scene while the involved drivers were still available. The research teams at both sites were available in police cars throughout their 10-hour shifts. They monitored the police radio, and the FLPD teams scanned on-board mobile data terminals. When informed of a crash, a team proceeded immediately to the location if the crash met the study's sampling criteria.[4]

Crashes were sampled if they occurred on a city street, alley or parking lot but not if they occurred on a freeway. A research team responded to crashes of all severities and all types if one vehicle in the crash was a passenger car, pickup truck, van, minivan or sport utility vehicle (SUV). Teams did not respond if the only vehicles involved were motorcycles, mopeds, taxis, bicycles, buses, trucks (other than vans and pickups), construction equipment, or emergency vehicles.

As a team drove to a crash site, the field interviewer readied the study forms by entering basic information such as date, hour and weather conditions. The Long Beach study officers were in charge of the investigation of most crashes, whether or not they were the first responders. The Fort Lauderdale teams did not perform the primary crash investigation unless they were the first responders and there was no other unit available to assume that responsibility. As first responder, the Fort Lauderdale team officer began the crash investigation but handed it over to a second responder. This freed the team to react to another crash as soon as study data had been collected. When not the first responder, the Fort Lauderdale team members identified themselves to the investigating officer, collected the study data as soon as possible, and moved on to wait for another crash.

The specific procedures for data collection varied somewhat as a function of crash severity. Upon arrival, the officer began the appropriate level of investigation, and the interviewer began the *Crash Site Observation Form* (Appendix C). This form was used to record observations of the crash and its location that could be obtained without input from

---

[3] The fact that all of the Fort Lauderdale researchers were women was purely coincidental. Recruiting solicited qualified individuals without regard to gender.

[4] In Fort Lauderdale, teams proceeded to crashes that were known to meet the study's criteria or when there was uncertainty if the crash was appropriate. The teams did not respond only when it was certain that the particular event was not suitable for sampling, e.g., the call was for a tractor-trailer that ran off the road. In Long Beach, the teams generally responded to all crashes.

drivers or investigating officers. The observations provided information about weather, lighting, road surface condition, roadway type, land use, type of crash, and number of involved vehicles, pedestrians and bicycles. In addition, observations of vehicle and a driver gender, approximate age, and ethnicity were recorded even though the questionnaire included these items. This provided data for a comparison of the characteristics of those who refused with those who agreed to participate.

3.5.1 Property Damage Only (PDO) Crashes

This category, in which none of the involved parties required medical attention, accounted for the largest number of sampled crashes. At PDO crashes, the team officer informed drivers at the earliest appropriate time that a research project was in progress and asked if they would speak with the interviewer. During the initial contact with a driver, the officer obtained an estimate of the driver's BAC with a Passive Alcohol Sensor (PAS).[5] Those estimates were the only BAC data obtained from drivers who chose not to speak with the interviewer.

The officer directed drivers who agreed to participate to the interviewer and then moved to a position out of hearing range from which the activity could be monitored. This position provided privacy but permitted the officer to intervene if necessary. To insure drivers' safety and comfort during the interview, adjustments were made as necessary for weather conditions and other environmental factors.

Interviewers approached each driver with a standard greeting: [6]

> *Hello. I am ____, of ________________. We are conducting a research study for the United States Department of Transportation. The information you give us will help make our roads safer.*
>
> *I need about five minutes of your time. I'll ask some questions about driving, and I'll record your answers, not by your name but by a number. What you tell me will be entirely confidential. There is no way that anything we talk about could be connected with you.*
>
> *It's really important that we get this information from you. Can you give me five minutes?*

---

[5] The device was the Sniffer™ P.A.S.® III manufactured by Public Service Technologies. This unit is a metal flashlight, which houses a pump that draws air through a fuel cell to measure alcohol content. The pump inlet must be held within 5 - 7 inches of the subject's face with the pump activated as he/she speaks. A series of lighted bands indicate the alcohol level with a higher number of illuminated bands indicating a higher BAC. The Sniffer™ is not an evidential quality instrument that provides precise BAC measurements, but it provided information in addition to the officers' detection or suspicion of alcohol.

[6] All standard scripts were translated into Spanish and printed on laminated forms that non-Spanish speaking interviewers could present to drivers who spoke only Spanish.

Interviewer training emphasized the need to convey this opening information clearly to prospective subjects. If the driver agreed to continue after the introduction, the interviewer offered a copy of the Certificate of Confidentiality for inspection (Appendix D). This certificate, issued by the U.S. Department of Health and Human Services, guaranteed that information obtained for the study would not be released even to a court under a subpoena and was therefore totally confidential.

Interviewers were instructed to administer the questionnaire in a professional, non-judgmental, and polite manner. Training emphasized the importance of tone of voice, facial expressions, and general body posture. Interviewers followed the exact wording and specified order of the questions. Deviations were not allowed. Questions were repeated without elaboration whenever a driver asked for clarification. A breath specimen was requested after the questions had been asked:

> *OK, we just have one more thing to do.*
> *This is a Preliminary Breath Test. I'll need you to take a deep breath, bring the mouthpiece to your lips, and steadily blow into it. I will tell you when to stop blowing and begin breathing normally. The duration of the test is about 5 seconds. Are you ready?*

After a breath specimen had been obtained, the interviewer thanked the driver and instructed him or her to await the officer's conclusion of the investigation.

3.5.2 Injury Crashes

If at least one driver or passenger required medical attention, the event was an injury crash. If the drivers involved in injury crashes were treated by emergency medical service (EMS) personnel at the scene, data were obtained by following the procedures described above for PDO crashes. If injured drivers were transported for treatment, the research teams attempted to obtain data at the hospital. In Long Beach, a small self-adhesive dot affixed by EMS personnel to the patient's paperwork alerted hospital staff that the driver was a study subject. They then were prepared when the research team arrived to direct them to the driver's location within the facility. If an injured driver was admitted, the hospital staff drew a blood specimen, and the research team obtained the results of the blood alcohol analysis.

In Fort Lauderdale, all local hospitals were notified of the data collection effort prior to study initiation. After a crash, the research team went to the hospital and obtained the location of the driver within the hospital from staff. If medical staff approved, the team interviewed the driver and requested a breath specimen. Since the officers on the team were usually not in charge of the crash investigation, they were not permitted to request blood tests from the hospital. The site coordinator therefore attempted to obtain the results of any blood alcohol tests done by the hospital as part of treatment or requested by the investigating officer. Whenever a BAC was obtained by a method other than the project's PBT, the results were recorded on the Non-Project BAC Form (Appendix C).

Blood specimens were not routinely obtained from non-admitted patients by hospitals at either site. In Long Beach, the team officer requested a blood specimen as permitted under California law for drivers believed to be under the influence of alcohol (DUI). In Fort

Lauderdale, the team officer was not in charge of a DUI investigation and could not, therefore, request an evidentiary test. If the hospital did not order an alcohol test, the officer attempted to obtain voluntary cooperation for a breath specimen with a PBT.

3.5.3 Fatal Crashes

One or more of the drivers or passengers sustained fatal injuries in a small number of sampled crashes. In those cases, the team followed the procedures as previously described to obtain data from the surviving drivers. Coroner's reports, including the analysis of blood or other body fluid for BAC, were obtained for deceased drivers.

3.5.4 Hit-and-Run Crashes

In a significant number of cases, one or more drivers fled the scene of a crash. When this happened, the research team first obtained data from any drivers who did not flee. They then used information obtained from witnesses to try to locate the hit-and-run driver. If the driver was apprehended within two hours of the crash, the team requested an interview and breath specimen following regular study procedures.

3.5.5 Intoxicated Crash-Involved Drivers

Because drivers participated voluntarily and were guaranteed confidentiality, their BACs measured for the study could not be used for enforcement. Nevertheless, safety concerns dictated that drivers with elevated BACs not be allowed to drive away from the scene.

Relatively few crash-involved, intoxicated drivers escaped detection. If the officer investigating a crash suspected that a driver was DUI, that suspicion triggered an investigation that included invoking implied consent and a request for a chemical test for alcohol. When this occurred, the project team used the evidential test and did not request an additional breath specimen. This avoided legal problems that might have arisen from a claim that the project test satisfied the driver's duty to provide a sample under the implied consent statute. Although the teams used evidential quality PBTs, the devices did not meet all of the certification requirements for evidential specimens in the study jurisdictions.

3.6 Control Sampling

The research teams returned to each crash site one week after the crash to sample control drivers who were matched to the crash-involved drivers on location, time of day, day of week, season of year and direction of travel. It was not possible to match drivers on the age, gender and other demographic and socioeconomic characteristics of the crash driver, but two control drivers were sampled for each crash driver. For example, a single vehicle crash required two control drivers, but a crash involving two drivers required four control drivers. Two controls per crash driver did not impose an unreasonable burden on the driving public. Also, that number could be obtained within the limits of sampling time at virtually all locations. If a larger number had been obtained, the sampling times might have extended far beyond the time of the crash.

Vehicles driven by potential control subjects were stopped in locations deemed to be safe and as close as possible to the crash site. They were selected from traffic flowing in the same direction(s) as the crash driver(s). To control for selection bias, the first qualified vehicle to cross an imaginary line at the defined sampling start time was stopped. If several vehicles were following closely and the officer deemed it dangerous to stop the lead vehicle, the last car of the group was selected.

Sampling was limited to the 30 minutes preceding and 30 minutes following the actual time of a crash. Prior to a sampling occasion, the research team estimated the time needed based on the number of drivers and the traffic flow at the site and then scheduled their activities to obtain half of the controls before the crash time and half after it. If events or conditions interfered with scheduled sampling, it was re-scheduled for a full week later to maintain the correct day-of-week.

3.6.1 Control Subject Processing

After a vehicle stop, the officer greeted the driver:

> *Good evening. Sir/Madam, you have done nothing wrong.*
>
> *I am stopping citizens this evening to ask them to cooperate with a research study being conducted for the U.S. Department of Transportation. I would like you to speak with an interviewer but you certainly are not required to do so and are free to leave whenever you wish.*
>
> *Could we have a few minutes of your time?*
> *If you agree to do this, it will take less than 10 minutes.*

While waiting for the driver's response, the officer observed the driver and vehicle carefully. Gender, age, and vehicle type were recorded on the *Control Site Observation* form (Appendix E) and thus were available later for comparisons of refusing drivers and participating drivers. The officer also obtained a BAC estimate with a PAS.

If a driver declined to participate, he/she was guided back into traffic without further comment. Although the information collected by the officer up to that point was recorded, the contact was not a counted as a completed stop, and sampling continued until the required number of drivers were obtained or the allotted time expired.

If the driver agreed to participate, the officer directed the driver to move the vehicle to a safe location and then signaled the interviewer to approach. The interview proceeded in the same manner as for crash-involved drivers. Upon completion, the officer assisted the driver to enter the flow of traffic.

3.6.2 Intoxicated Control Drivers

The control stops were made solely to collect study data, not on suspicion of DUI. Nonetheless, the measured BACs of a number of the drivers were ≥0.08%. These drivers were offered four options:[7]

- A non-impaired passenger with a valid driver's license would be allowed to assume driving responsibility.

- The impaired driver would be transported to his/her home by taxi with the fare paid by the project. The driver's vehicle would be locked and parked in a safe place.

- The person could drive away, but the police would be notified and he/she would likely be arrested for DUI. No driver ≥ 0.08% BAC drove away.

- If a Fort Lauderdale driver lived relatively nearby, the researcher could elect to drive him/her home with the team officer following close behind.

3.7 Data Management

The project activities produced large sets of the data forms shown in Appendices B, C and E. At the end of each work shift, completed forms were checked by the researcher and by the site coordinator to confirm that they were accurate, complete and legible. When a case was complete with all the forms for both crash and control drivers, the coordinator delivered it to Dunlap (Fort Lauderdale) or to SCRI. (Long Beach). After another review, the forms were entered into a database at Dunlap.

Data entry personnel used software that mimicked the paper forms. Numeric and multiple-choice items were range-protected so that only allowable values could be entered. Cases were re-keyed by a different data entry person to verify the accuracy of entries. When a difference was found, the Dunlap database manager determined the correct value.

After entry and verification, all data were listed and checked for apparently discrepant values. Questionable entries were crosschecked with the paper case. When the entire data set had been entered, a set of tabulations for each variable by site was prepared using the SPSS® software package. These were examined for outliers that were then checked by hand back to the paper forms to determine whether they were legitimate.

As the final step in the data management process, selected tabulations were repeated using the SAS® package that was to be used for most of the statistical analyses. These were compared with the SPSS®-generated tabulations to verify that the database had been accurately transferred between the systems.

---

[7] FLPD research officers strongly recommended not driving if the subject reached a .05 BAC, but allowed the participant to drive if he or she was under the .08 limit.

3.8 Data Analysis

Data analysis focused on three areas. First, analyses were performed to describe the database, assess its internal consistency and determine its ability to support inferential approaches. This first step determined the extent of missing values for each variable and compared the data from the two sites. Appendix F presents many of the results of these examinations.

The second set of analyses used logistic regression techniques to calculate relative risk due to alcohol for the entire dataset. Various univariate and multivariate approaches produced estimates for the entire driving population. The third set of analyses examined relative crash risk as a function of BAC for various subgroups of the sample (e.g., age and drinking frequency). Section 4 presents the results of these analyses along with more details on the statistical approach.

This Page Intentionally Left Blank

4. RELATIVE RISK ANALYSES

This section presents an overview of the statistical analyses of study data to calculate relative risk and describes the steps of the parameter estimation process. Specific methodological details are interspersed with the results.

4.1 Analysis Strategy

As discussed earlier, after data entry, a set of analysis files were created using the SPSS® software package. These original SPSS® files were then converted to files for the SAS® software for the relative risk analyses. The accuracy of the conversions was verified by generating frequency distributions and cross-tabular arrays in SAS® for comparison with similar SPSS® tables. Most of the data analysis then was executed with the SAS® PHREG procedure (SAS Institute, 1996), which can perform a conditional logistic regression analysis for a wide variety of matched case-control designs.[8] The analyses proceeded in the following sequence:

(1) analysis of missing data and potential selection bias due to subject non-participation and non-recoverable hit-and-run drivers;
(2) univariate logistic regression of the unadjusted relationship between BAC and relative crash risk;
(3) analysis of site x BAC interactions;
(4) multiple logistic regression analyses of the BAC – crash risk relationship adjusted for various subsets of covariates, such as age, gender, drinking patterns and socioeconomic status;
(5) evaluation of selected BAC x covariate interactions;
(6) evaluation of the BAC – crash risk relationship for three crash subtypes: late night (10 PM or later), single vehicle, and severity level (PDO vs. injury/fatality);
(7) adjustment of relative risk curves for nonparticipation bias;
(8) assessment of the accuracy of the classification of drivers by case-control status by logistic regression equations;
(9) calculation of confidence intervals for the relative risk estimates;
(10) calculation of relative risks for specific subgroups of a priori interest.

4.2 Logistic Regression Model Development And Case-Control Contrasts

Two control drivers were matched to each crash-involved driver, but not all drivers agreed to participate, and some drivers (hit-and-runs) who fled the scene could not be located. Three alternative case-control designs, therefore, include differing numbers of crash and control drivers:

---

[8] This design was executed in SAS® by specifying “TIES=Breslow” in the model statement and configuring the input data so that the matched sets (1 crash driver and 2 controls) were separate strata (SAS Institute, 1995). This model statement correctly conditions for the fact that two controls are matched to each crash driver thereby creating two contrasts of the covariate vector within each strata. All analyses were done using the SAS® events/trials syntax mode**.** For a full description of the logistic regression theory and approach, the interested reader is directed to references such as Hosmer and Lemeshow (2000).

- A conditional 1:2 design included only those cases in which the crash driver(s) and two control drivers participated.

- An M x N conditional design included all crashes with at least one crash driver and one control driver. This design produced a larger sample, but it did not preserve the pair wise matching on direction of travel or the 1:2 crash driver to control driver ratio.

- An unconditional design included all drivers who participated in the study and produced the largest sample. It did not preserve the pair wise matching and did not require that a given crash driver have a control driver match or visa versa.

Within each of the above designs, multiple analyses differed by inclusion or exclusion of the hit-and-run drivers who were apprehended and by varying covariate sets. Figure 1 graphically portrays the configurations and sample sizes generated by the three models. Note that since BAC was required for a relative risk determination, the three designs included only drivers who provided breath specimens.

All three designs are subject to self-selection bias as a result of non-random factors associated with the refusal or inability of drivers to participate. Preliminary logistic regressions examined the relative risk curves produced by each strategy and, based on those results, most analyses used the 1:2 conditional design. This design best conforms to the method of sampling drivers, and its preservation of matched linkages provided more precision and accuracy for estimating the BAC - crash risk relationship**.**

4.3 Variable Selection and Chosen Significance Level

To formulate the regressions, a modified backward elimination procedure forced all variables into the equation and then eliminated them one at a time until only variables with $p \leq .20$ were retained. At this point, a sequential option was utilized to delete the BAC polynomial terms one at a time and then as a set. The individual polynomial powers and all powers as a set were evaluated by changes in the log likelihood chi-square and score statistic. Graphical plots and tabulations of the relative risk curve for the BAC variable were generated with Excel®. An equation constructed using the conventional $p \leq .05$ retention criterion produced similar results.

The $p \leq .20$ significance level for retaining covariates maximized statistical precision and allowed adjustment of the BAC – crash risk relationship for potentially confounding covariates whether or not they met the conventional $p \leq .05$ significance level. In this study, BAC was the independent variable and the other variables were considered covariates analogous to an analysis of covariance design. The purpose of the covariates was to increase the precision of the crash risk estimates and to adjust the estimates for bias due to confounding variables. The procedure did not identify all covariates that might be related to BAC, but it controlled covariates that were jointly related to both BAC and crash risk.

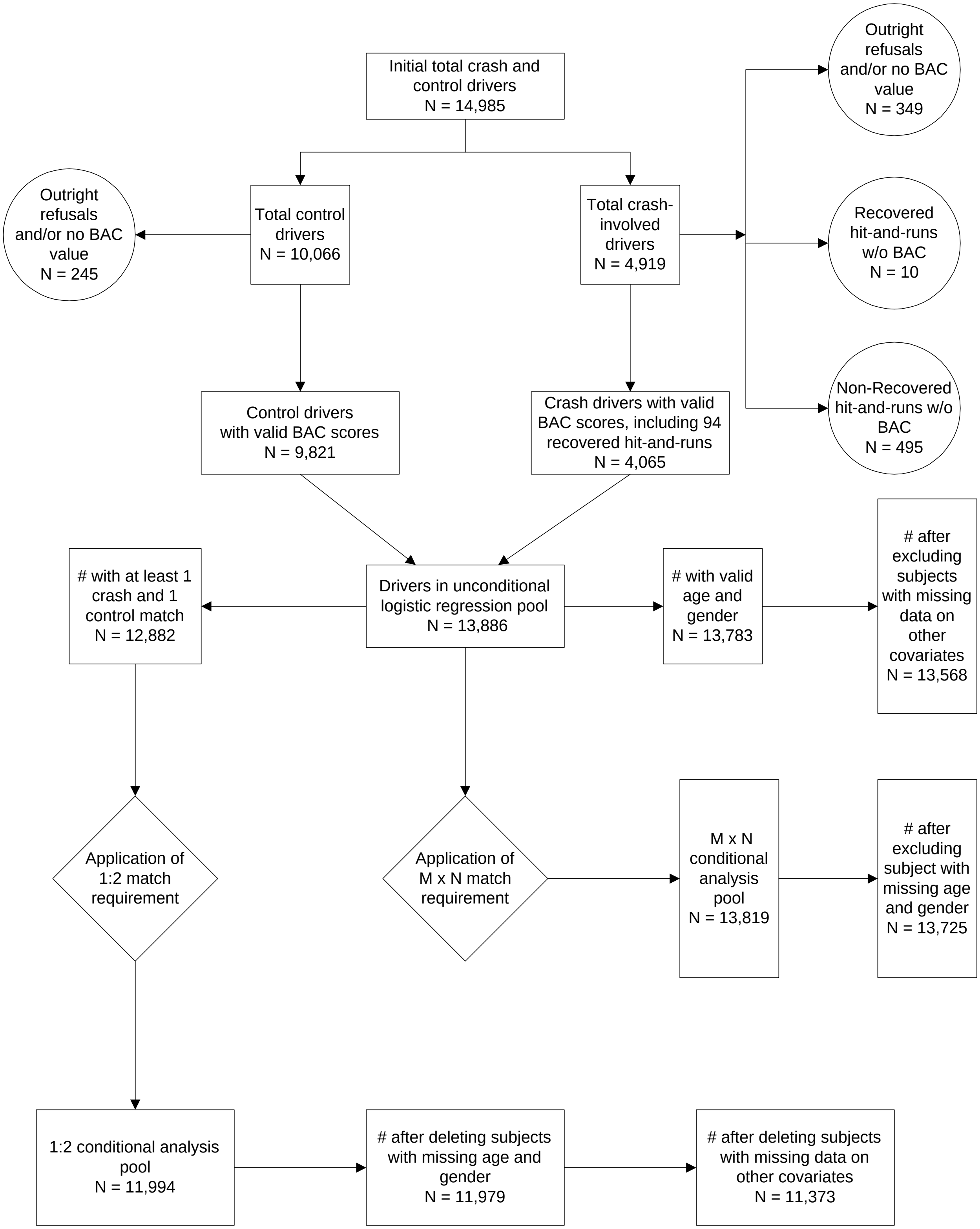

Figure 1. Flow of sample sizes of crash and control drivers resulting from missing data and logistic regression model requirements

4.4 Source and Selection of Covariate Pools

Covariates were selected for a series of logistic regressions on the basis of a priori, substantive, and methodological considerations. The primary source for the covariate data was the subject interview form (Appendix B). Some analyses also included covariate data from the crash and control observation forms (Appendices C and E). The following regressions were computed:

- Univariate logistic regressions that used BAC as the sole independent predictor variable of crash-control status.
- Multiple logistic regressions that used age and gender as covariates.
- Multiple logistic regressions that used age, gender, and alcohol consumption as covariates.
- Multiple logistic regressions that used age, gender, and other demographic and socioeconomic variables as non-reactive covariates.[9]
- Multiple logistic regressions that used all available covariates, including alcohol consumption and sleep patterns.
- Multiple logistic regressions that used all covariates except direct alcohol-related measures (e.g., alcohol consumption patterns, prior DUI convictions, self-reported frequency of driving after drinking).

4.5 BAC Data

BACs were displayed on the PBTs and recorded on the data forms and in the initial SPSS® database with three digits. Two-digit BAC estimates obtained with the PAS instrument, which were available for 35% of the drivers who refused to provide breath specimens, were used to assess non-participants' BACs and to generate BAC imputations. Comparison of subsets of the logistic regressions with two and three digit BACs showed almost identical equations and relative risk curves. To provide uniformity, therefore, two digit BAC data were used in all analyses.

The elapsed time from a crash to the collection of a BAC specimen differed between crashes. In most cases, however, the time interval was short, and variations were minimal. BACs were therefore not adjusted for the time differences, because the accuracy of the adjustments would have served to add uncertainty.

Non linearity of relative risks. Relationships involving binary outcome variables (proportions) are inherently non-linear**.** A review of BAC relative risk studies suggested that the non-linearity in the study data would be too extreme to be removed through a logistic transform, and a preliminary inspection of the data confirmed this expectation. The BAC –

---

[9] A covariate is defined herein as "non-reactive" if it is not a potential cause of crashes through its effect on BAC. For example, gender and age are non-reactive covariates because they do not directly cause any particular BAC or crashes but drinking frequency would be a reactive covariate because it has a functional relationship to BAC and crash risk.

crash risk relationship remained non-linear on the log odds scale, but it was adequately modeled with the introduction of power terms. Most of the analyses required a cubic polynomial model (BAC, BAC^2, BAC^3) for adequate fit. A fourth power (quartic) produced marginally significant improvement in some instances, but the small gain was not sufficient to offset the added complexity. Although all analyses centered the BAC values to reduce multicollinearity (Darlington, 1990), it remained high between some of the polynomial components, and that condition led to increased standard errors of the parameter estimates (Cohen and Cohen, 1983).

Interaction terms. To determine whether the BAC – crash risk relationship varied as a function of driver characteristics such as age, gender and drinking patterns, a priori selected interactions were evaluated by forming the respective two-way product terms in the SAS® data step commands (SAS, 1995). The interactions were evaluated with the Wald chi-square and the score and likelihood ratio tests (SAS, 1995; Hosmer and Lemeshow, 2000). A similar procedure tested the significance of site x BAC interactions. The structure of significant interactions was summarized graphically by plotting the logits and relative risks from the regression coefficients for main effect and interaction parameters.

Adjustment for self selection bias and missing data. Data were missing due to:

- Drivers' refusals to provide specimens for BAC measurement, refusals to participate in the interview, or refusals of both the specimen and the interview;
- Non-apprehended hit-and-run drivers;
- Driver refusals or inability to answer questions;
- Data collection and recording errors.

Data were available to reduce or assess bias arising from refusals and non-recovered hit-and-runs from two sources: 1) BAC estimates obtained with the PAS; and 2) breath specimens obtained from hit-and-run drivers who were apprehended within two hours (see Table F-2 in Appendix F). A simple weighting procedure was used to account for the proportion of hit-and-run drivers who were not apprehended. The validity of the adjustment rests on the assumption that the BAC distribution for non-apprehended drivers was the same as the distribution for those who were located and tested. Although this assumption is not testable and may not be entirely accurate, this approach is far superior to ignoring the non-tested hit-and-run drivers altogether.

Crash subtypes. The triggering event for sampling was a crash of any severity reported by or to the police and responded to by a research team. The crashes are categorized in terms of three characteristics: 1) time of day (late night vs. earlier); 2) number of vehicles (single vs. multiple); and 3) severity (injury/fatality/property damage only). Each of these crash types was evaluated by including the respective BAC x crash type product term in a logistic regression model to test for the significance of the interaction. To illustrate the meaning of the interactions, a finding of statistical significance for BAC x crash type would indicate that late night crashes (10 PM or later) are related to BAC differently than crashes that occur earlier in the day, and that crashes at different times have different relative risk curves.

Classification accuracy. The prediction of crash or control group membership from BAC and the regression equations was evaluated by the Cox-Snell R^2 statistic, which is a log likelihood analog to the percent of variance explained (R^2) in ordinary least squares regression (Tabachnick and Fidell, 2000; SPSS, 1997):

$$\text{The Cox-Snell } R^2 = 1 - \frac{L_{(o)}}{L_{(B)}}^{2/N}$$

where $L_{(o)}$ is the likelihood for the model with only a constant, $L_{(B)}$ is the likelihood for the model containing the specified covariates and N is the sample size. One limitation to the Cox-Snell R^2 statistic is that it cannot achieve a maximum value of 1.0, even for a model that perfectly predicts an outcome measure. However, similar constraints also operate to attenuate the maximum attainable R^2 in ordinary least square regression (Peck, McBride and Coppin, 1971).

The classification matrix option in the SAS® LOGISTIC routine provided an additional measure of predictive accuracy. This option used the logistic regression function to predict the group membership probability (control vs. crash) of each driver for comparison with his or her actual status in the form of a 2 x 2 classification matrix.

4.6 Relative Risk Results

This section discusses additional factors taken into account in calculations and presents relative risk results. The results are presented by increasing complexity, beginning with the results of univariate risk calculations and proceeding to the results of multivariate analyses.

4.6.1 Participation Rates and Potential Biases

Table 4 summarizes the rates of participation and refusal by control and crash drivers. Data for drivers who participated but did not complete the interview and/or did not provide a valid breath specimen are tabled in the row labeled Partial. The table does not include data for hit-and-run drivers.

Drivers refused to participate as control subjects at low rates at both sites (Long Beach 2.90%; Fort Lauderdale 1.34%). Crash-involved drivers refused at higher rates with an average of 7.65% for both sites. A log linear analysis indicated that the difference between crash and control drivers was statistically significant ($p < .0001$). The refusal rates varied by site with 4.9% of crash drivers refusing at Long Beach and 10.2% at Fort Lauderdale (site x case-control status interaction $p < .0001$). Appendix G, Table G-1 contains the log linear results.

The crash and control drivers’ characteristics determine the extent of the bias arising from refusals to participate. It is unlikely that refusals are a random variable. Sobriety checkpoint data indicate that drivers who will not provide a breath specimen are more likely to have been drinking and to be impaired (Carlson, 1979). The PAS estimates of BACs, together with interviewers’ observations for a majority of participants and many non-participants, permitted a direct examination of the bias.

Table 4. Agreement to Participate on Initial Approach by Case-Control Status and Site
N=14,382

Level of agreement to Participate	Long Beach		Fort Lauderdale		Total	
	Crash	Control	Crash	Control	Crash	Control
Complete	1932 (94.1%)	4837 (96.6%)	1972 (87.2%)	4984 (98.5%)	3904 (90.5%)	9821 (97.6%)
Partial	22 (1.07%)	24 (0.48%)	60 (2.70%)	8 (0.16%)	82 (1.90%)	32 (0.32%)
Refusal or unable	100 (4.87%)	145 (2.90%)	230 (10.2%)	68 (1.34%)	330 (7.65%)	213 (2.12%)
Total	2054 (100.0%)	5006 (100.0%)	2262 (100.0%)	5060 (100.0%)	4316 (100.0%)	10066 (100.0%)

Note: Excludes all 603 hit-and-runs whether or not they were recovered

The PAS readings signal the presence of alcohol with bars on a display of nine light emitting diodes (LEDs). Table 5 summarizes the PAS data by three levels: 0 bars illuminated, 1 – 2 bars illuminated, ≥3 bars illuminated. Since a ≥3 score indicates the detection of alcohol, scores in the range 3 – 9 were collapsed into one group to simplify the analysis (Appendix G, Table G-2).

Table 5. Participation Rates by Case-Control Status and PAS Reading
(*N*= 14,382)

		Crash			Control		
		Complete	Partial	Refused	Complete	Partial	Refused
PAS level (No. of bars)	**0**	3016 (87.8%)	34 (60.7%)	40 (52.6%)	8393 (92.6%)	21 (87.5%)	79 (83.2%)
	1 & 2	47 (1.37%)	2 (3.58%)	2 (2.64%)	145 (1.60%)	2 (8.33%)	6 (6.32%)
	3 or more	373 (10.9%)	20 (35.7%)	34 (44.7%)	530 (5.84%)	1 (4.17%)	10 (10.5%)
	Total	3436 (100.0%)	56 (100.0%)	76 (100.0%)	9068 (100.0%)	24 (100.0%)	95 (100.0%)

Note: Excludes all hit-and-runs and subjects with missing or invalid PAS scores

The results indicate that refusals were associated with the presence of alcohol, particularly among crash drivers. Note that the PAS reading was ≥3 for 44.7% of the crash drivers who refused to participate, compared to 10.9% of crash drivers who participated fully. A log linear analysis (Appendix G, Table G-3) confirmed that non-participants' PAS scores were significantly higher (p<.0001) and that the association was stronger for crash drivers (case-control status x participation, p<.03). Given the substantial relationship (R = .82) of PAS estimates with drivers' BACs, the PAS data were subsequently used to predict the BACs of drivers who refused to be tested. The BAC – crash-risk relationships were re-analyzed after imputing the BAC of the refusals.

Table 6 summarizes interviewers' ratings of drivers' intoxication levels. To minimize the problem of small cell numbers, partial participants have been combined with refusals, and drivers suspected of impairment have been combined with drivers with obvious signs of impairment. The results are similar to and support the PAS results. Drivers who displayed signs of impairment were more likely to refuse to participate, a tendency that was much stronger among crash drivers.

Table 6. Distribution of Perceived Impairment Signs by Case-control Status and Level of Participation*
***N* = 13,889**

Participation Level	Crashes (*N* = 4084)**		Controls (*N* = 9803)	
	No Signs (%)	Suspicious or obvious (%)	No Signs (%)	Suspicious or obvious (%)
Complete	88.2	11.8	95.0	5.0
Refused or Partial	53.0	47.0	84.0	16.0

*Excludes non-recovered hit-and-runs and subjects with missing or invalid codes.
**Recovered hit-and-runs who gave alcohol specimens are included as participants.

Table 7 presents the participation status of crash drivers, including hit-and-run. The relatively high incidence of hit-and-run drivers created a second source of bias. Note that 603 hit-and-run drivers (499 not recovered + 104 recovered) represent 12.2% of all crash drivers. On intuitive grounds alone, a high incidence of alcohol is expected among drivers who flee the scene of a crash, and the data in Table 8 confirm that expectation. Almost 70% of 94 hit-and-run drivers who were breath tested were alcohol positive, and 62.7% of their BACs were $\geq$0.08%.

Table 7. Participation and Recovery Incidence of Crash Drivers by Site

Participation Status	Long Beach	Fort Lauderdale	Total
Complete	1932	1972	3904
Partial	22	60	82
Refused on approach	100	230	330
Hit-and-run – recovered	77	27	104*
Hit-and-run –not recovered	293	206	499
Total	2424	2495	4919

*Ten of these subjects did not provide a BAC specimen.

Table 8. Distribution of Recovered Hit-and-Runs and Total Crash Cases, by BAC (Both Sites Combined)

BAC	Recovered hit-and-run drivers with BAC values		All crash-involved drivers with BAC values except hit-and-runs	
	N	%	*N*	%
0.00	29	30.9	3290	82.9
.01-.04	3	3.2	176	4.4
.05-.07	3	3.2	97	2.4
.08-.09	5	5.3	53	1.3
.10-.19	33	35.1	233	5.9
.20+	21	22.3	122	3.1
Total	94	100.0	3971	100.0

Exclusion of hit-and-run drivers from the crash group creates a serious bias in relative risk estimates. The problem is more serious than that created by refusals due to the larger number of drivers, a larger proportion of high BAC drivers, and restriction of the bias to the crash group.

A reasonable adjustment of the bias can be based on data obtained from hit-and-run drivers who were apprehended and breath tested. This method extrapolates from the BAC distribution of the recovered hit-and-runs to all hit-and-runs and recalculates the relative risks. This weighting method assumes that recovered hit-and-run drivers and non-recovered hit-and-run drivers are samples from a common underlying population. Like most ad hoc assumptions, it may not be entirely valid, but a much larger bias would exist if no such adjustment were made.

The final source of known bias arises from missing covariate data. When participants refused to complete the interview or could not answer many of the questions, data are missing for several or most of the covariates. Interviewer omissions and recording errors account for an additional small amount of missing data. Overall, data are missing for one or more covariates for approximately 5% of the sample. Age and gender data were missing from 0.7% of the drivers with valid BACs. Values are missing for 2.2% of the drivers with valid BAC, age, and gender data on one or more of the alcohol consumption questions. For 0.6%, data are missing for all of the alcohol consumption questions.

The most rigorous methods for imputing missing values (Breslow, 2000; Rao, 1996) base estimates on the relationships among covariates in the complete data matrix using the EM algorithm or multivariate propensity score procedures (Rubin, 1987; D'Agostino and Rubin, 2000). These approaches are very labor intensive and become problematic if many of the subjects have missing scores on numerous covariates.

For this study, the severity of the bias depended on the amount of missing data and whether missing data were related to crash/control status and BAC. The magnitude of this bias was evaluated by performing a multiple logistic regression of crash/control status on age, gender, BAC and presence of missing data for alcohol consumption covariates (0 = none; 1 = some). Table 9 summarizes the results.

Table 9. Unconditional Logistic Regression Analysis of the Effects of Missing Covariate Data

***N* = 13,874**

Variable	df	Coefficient	Wald chi-square	p	Odds ratio
Intercept	1	-.34	21.4	< .0001	-
Gender (female)	1	-.21	28.4	< .0001	.81
Age under 21	(referent)	-	-	-	1.00
Age 21-24	1	-.46	29.7	< .0001	.63
Age 25-34	1	-.67	89.3	< .0001	.51
Age 35-44	1	-.75	105.3	< .0001	.47
Age 45-54	1	-.79	97.0	< .0001	.45
Age 55+	1	-.45	29.9	< .0001	.64
Missing data (none = 0)	1	.41	9.0	< .003	1.51
BAC linear	1	-5.61	4.8	< .03	-
BAC quadratic	1	230.8	28.4	< .0001	-
BAC cubic	1	-615.5	15.8	< .0001	-

Note: Cox-Snell R^2 = .053

Note from the regression coefficient (B) and odds ratio (exp[B]) for the missing covariates that after adjusting for age, gender and BAC, subjects with missing data were more likely to be members of the crash group (OR = 1.51). Data were missing for less than 5% of the crash drivers, however, and the linear correlation between the missing data dummy variable and BAC was modest (*r* = .26). These results suggest that the bias introduced by deleting subjects with missing covariate values would slightly attenuate the relative risk curve. This bias was removed through a simple weighting procedure.

4.6.2 BACs

Three digit BACs from the PBTs as stored in the SPSS® data file were rounded to two digits in SAS®. To determine whether this rounding procedure introduced distortion or attenuated meaningful variance, two unconditional logistic regressions of BAC on case-control status were performed. One used three-digit BACs, and the other used two-digit BACs. The parameter values of the two equations were very similar with almost identical cubic polynomial fits. The BAC means and standard deviations were almost identical, and each produced the same Cox-Snell R^2 (.041). A plot of the relative risk curves showed that the two equations produced virtually identical relative risks for all levels of BAC.

4.6.3 Univariate Analyses

Table 10 shows the raw distribution of crash and control drivers by BAC, the crude unadjusted relative risk (% of crash-involved drivers divided by % of control drivers), and the relative risk indexed to the zero BAC group.[10] The relative risks are estimated by the odds ratio, which provides an unbiased estimate when applied to case-control data where the outcome (crashes) is relatively rare (Breslow, 2000). The odds ratios are not adjusted for covariates, nor do they reflect the effects of subsequent transformation to a multiplicative scale through the logistic link function.

Table 10. Crude Relative Risk and Percentage Distribution of Crash Cases and Controls by BAC (Recovered Hit-and-Runs Included)
***N* = 13,886**

BAC	**Controls**		**Crash-involved**		**Ratio of crash % to control %**	**0 – indexed relative risk**
	N	**%**	***N***	**%**		
0.00	8716	88.75	3319	81.65	.92	1.00
.01-.02	405	4.12	111	2.73	.66	.72
.03-.04	253	2.58	68	1.67	.65	.71
.05-.06	158	1.61	62	1.53	.95	1.03
.07-.08	112	1.14	64	1.57	1.38	1.50
.09-.10	50	0.51	57	1.40	2.75	2.99
.11-.12	44	0.45	46	1.13	2.53	2.75
.13-.14	32	0.33	57	1.40	4.30	4.68
.15-.16	17	0.17	54	1.33	7.67	8.34
.17-.18	17	0.17	66	1.62	9.38	10.20
.19-.20	8	0.08	39	0.96	11.78	12.80
.21-.22	3	0.03	40	0.98	32.21	35..01
.23-.24	0	0.00	23	0.56	NC*	NC*
.25-.26	1	0.01	24	0.59	57.98	63.03
.27-.28	2	0.02	18	0.44	21.74	23.63
.29-.30	0	0.00	8	0.20	NC*	NC*
.31-.32	0	0.00	5	0.12	NC*	NC*
.33+	3	0.03	4	0.10	3.22	3.50
All BACs**		100.0		100.0		

*Not computable due to 0 frequency in controls. **Totals do not add to 100% due to independent rounding

---

[10] The indexing is accomplished by setting the risk at zero BAC to a value of 1 and adjusting all other ratios accordingly.

Table 10 shows a dip in relative risks at 0.01% - 0.04% BACs in both the Long Beach and Fort Lauderdale data. The dip is similar to that seen in the Borkenstein (1964; 1974) data. The relative risk values maximize at about 0.25% BAC and then begin to decline. Relative risk estimates are much less stable at high BACs, because the sample sizes are very small.

Three different statistical models (see Section 4.2) were used to generate logistic regression equations:

1. The conditional 1:2 logistic regression included only those crash events in which a crash driver was matched to two control participants (*N* = 11,994 drivers).

2. The conditional M x N logistic regression included all crash events in which at least one crash driver and one control were sampled (*N* = 13,819 drivers).

3. The unconditional logistic regression included all drivers with valid BAC values since the model did not require each case or control driver to have a match (*N* = 13,886 drivers).

Polynomial regression models were used for each of the above designs. For the conditional models, each analysis was performed with and without the inclusion of the recovered hit-and-run drivers thereby producing four separate regressions. Because of sparse data at very high BACs, complete BAC distributions would have introduced instability in the regression models.[11] The distributions were therefore capped at 0.25%, and drivers with higher BACs were given that value. All five regressions required a cubic polynomial equation for adequate fit. In some cases, a fourth order (quartic) term was marginally significant, but the contribution of the quartic term was negligible.

Table 11 contains the equation for the 1:2 conditional analysis (hit-and-run drivers included).

Table 11. 1:2 Conditional Logistic Regression of BAC – Crash Risk Relationship, Recovered Hit-and-Runs Included

***N* = 11,994**

BAC parameter	Df	Coefficient	S.E.	Wald chi-square	P
BAC linear	1	-7.01	2.79	6.31	.012
BAC quadratic	1	266.1	40.29	30.4	< .0001
BAC cubic	1	-730.4	174.9	17.4	< .0001

Cox-Snell *R2* = .038; Likelihood ratio chi-square = 459.4, df = 3, p < .0001.

---

[11] The number of drivers with BACs of .25 or greater was 65, and the maximum observed BAC value was .51. The median BAC value of all drivers with BACs $\geq$ .25 was .27.

The cubic polynomial term was statistically significant based on the Wald chi-square test ($p < .0001$). The Wald test is a simultaneous test for evaluating parameters in which each effect in the model is adjusted for all other effects and is conservative when applied to large parameter values such as those obtained for the quadratic and cubic polynomial terms (SPSS, 1997). The result was confirmed by the stepwise likelihood ratio chi-square and the score statistic, both of which are more appropriate than the Wald test for evaluating sequential and hierarchical structures (Tabachnick and Fidell, 2000; Hosmer and Lemeshow, 2000). Table 12 summarizes these tests.

Table 12. Stepwise Decomposition of Polynomial Components for 1:2 Conditional Model

Order of polynomial	Likelihood ratio chi-square	Score statistic	Cox-Snell *R2*
Linear	401.1	416.4	.032
Quadratic	443.4	432.6	.037
Cubic	459.4	466.8	.038
Quartic	460.1*	468.5*	.038

*Incremental increase not significant, p > .15.

In higher order polynomials, the magnitude and significance level of the lower order terms does not reflect their importance relative to the higher order terms. For example, a model containing only the linear term for the regression produced a Cox-Snell R^2 of .032, thereby capturing 84% (.032/ .038) of the variance explained by all three terms combined (Table 12).

Table 13 shows the results for the M x N regression. Although the parameter values for the equations are slightly different, the equations have the same cubic structure and the magnitude of the coefficients is similar to the 1:2 model. The Cox-Snell R^2 is smaller (R^2 .033 vs. R^2 .038).

Table 13. M x N Conditional Logistic Regression of BAC-Crash Risk Relationship, Recovered Hit-and-Runs Included

(*N* = 13,819)

Order of Polynomial	Df	Coefficient	S.E.	Wald chi-square	p
BAC linear	1	-6.12	2.70	5.12	.024
BAC quadratic	1	238.0	45.6	27.2	< .0001
BAC cubic	1	-637.9	163.7	15.2	< .0001

Cox-Snell R^2 = .033; Likelihood ratio chi-square = 453.6, *df* = 3, $p < .0001$.

The polynomial structure makes it difficult to visualize or make inferences about the relative risk function by inspection of the regression coefficients. Each BAC is a composite of three intercorrelated numbers – BAC, BAC^2 and BAC^3. To assess the shape of the relative risk curve, each BAC has been multiplied by the polynomial values and then expressed as exponentiated deviations from the logit of the zero BAC group in Table 14, and Figures 2 and 3.

Table 14. Univariate Relative Risks as a Function of BAC and Case-control Design, Cubic Polynomial Models

BAC	**Recovered hit-and-runs included**			**Hit-and-runs excluded**	
	1:2 conditional	**M x N conditional**	**Unconditional**	**1:2 conditional**	**M x N conditional**
0.00	1.0	1.0	1.0	1.0	1.0
.01	.91	.92	.89	.90	.91
.02	.86	.88	.84	.86	.87
.03	.87	.88	.86	.86	.87
.04	.91	.92	.87	.90	.90
.05	.99	1.00	.94	.97	.97
.06	1.12	1.12	1.05	1.09	1.07
.07	1.31	1.29	1.22	1.26	1.23
.08	1.57	1.53	1.45	1.50	1.43
.09	1.93	1.84	1.77	1.82	1.71
.10	2.41	2.26	2.21	2.24	2.08
.11	3.06	2.81	2.79	2.81	2.56
.12	3.92	3.52	3.56	3.54	3.17
.13	5.04	4.45	4.56	4.49	3.96
.14	6.49	5.63	5.86	5.70	4.95
.15	8.32	7.10	7.50	7.20	6.17
.16	10.59	8.91	9.52	9.03	7.65
.17	13.31	11.07	11.94	11.17	9.39
.18	16.45	13.57	14.74	13.59	11.36
.19	19.91	16.35	17.80	16.19	13.52
.20	23.49	19.29	20.95	18.80	15.75
.21	26.90	22.19	23.93	21.18	17.90
.22	29.76	24.81	26.41	23.07	19.75
.23	31.69	26.85	28.02	24.18	21.07
.24	32.31	28.01	28.47	24.28	21.75
.25+	31.42	28.07	27.56	23.27	21.52
N	11,994	13,819	13,886	11,700	13,725

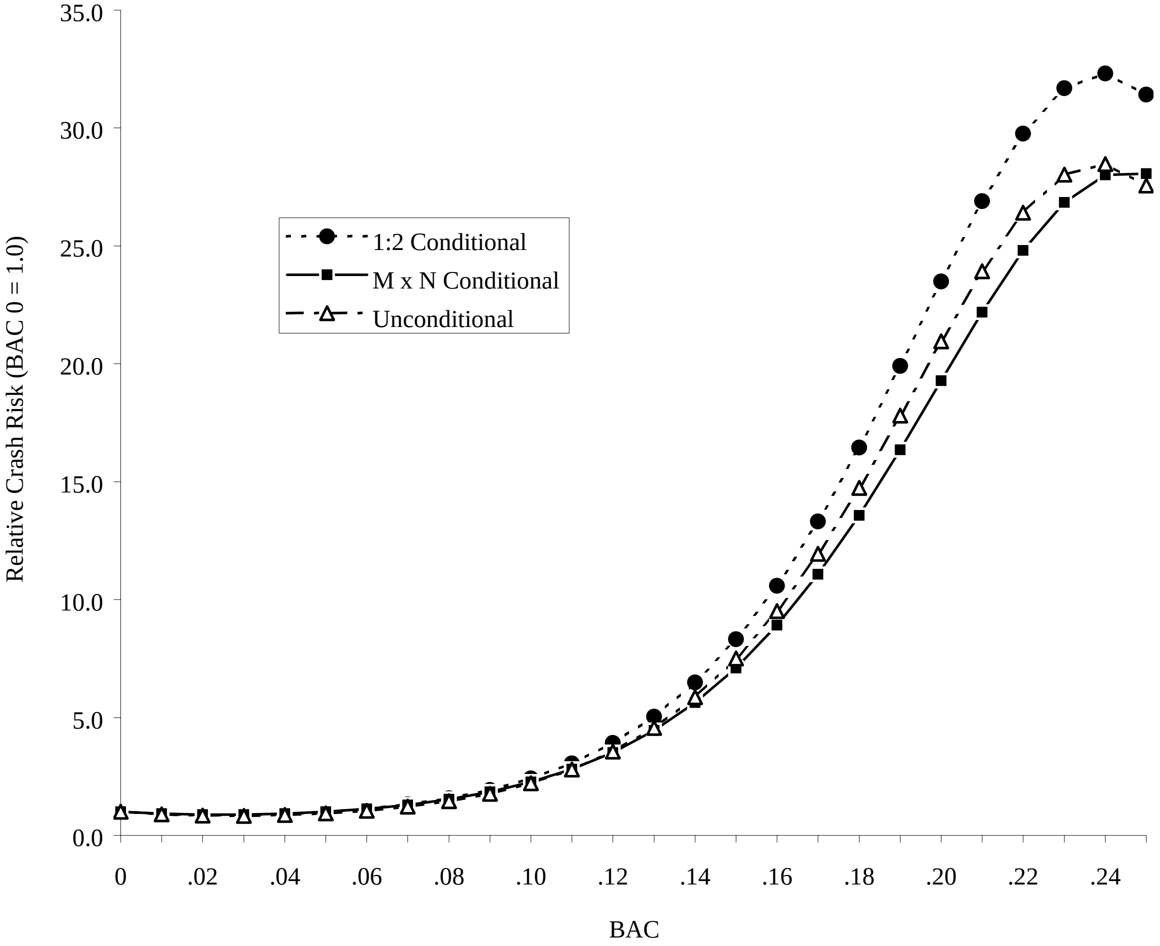

Figure 2. Univariate relative risks as a function of BAC level and case-control design – cubic polynomial models with hit-and-runs included

The univariate risk curves demonstrate the following:

- Exclusion of recovered hit-and-run drivers substantially lowers relative risk, particularly at higher BACs (compare Figures 2 and 3).
- A dip in relative risk appears for BACs below 0.05%. Crash odds for drivers at 0.02% - 0.03% BACs are 12% - 17% lower than for drivers at 0.00% BAC.
- The 1:2 conditional model produces the steepest risk curve, but differences between the three models are not dramatic.

BAC was not adjusted for any covariate in the preceding univariate models. As a result, it is possible that the relative risk curves are confounded by variables that are jointly associated with crash-involvement and BAC. The confounding could either attenuate or inflate the BAC – crash risk relationship. Also, the relative risk estimates are subject to self-selection bias due to differences in participation rates and hit-and-run drivers.

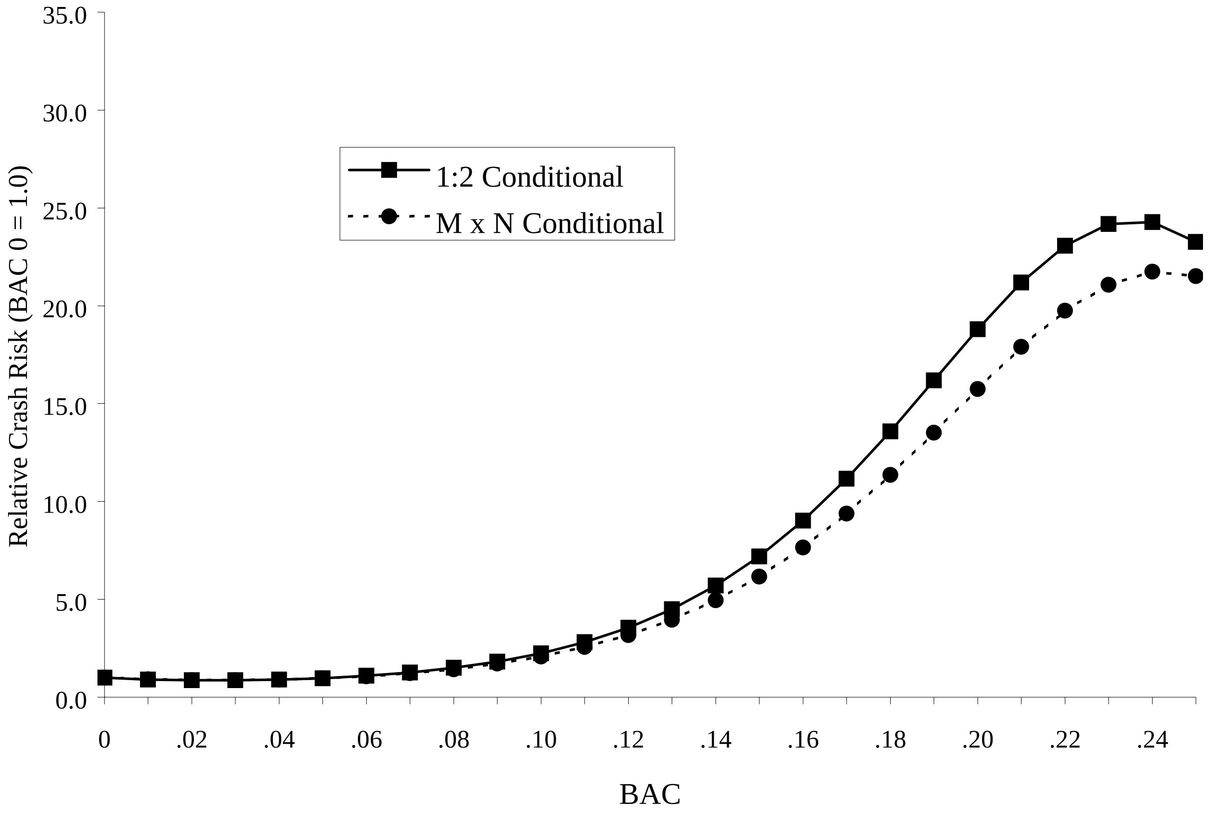

Figure 3. Relative risks as a function of BAC level after hit-and-runs deleted

The potential for biasing effects can be illustrated with the hit-and-run drivers. The distribution of recovered hit-and-runs by BAC (Table 8) indicates that they were more likely than crash-involved drivers who did not flee to have an elevated BAC. Figure 4 and Table 15 display the net increase in relative risk as a function of including recovered hit-and-run drivers. The differences, which are negligible at low BACs, become large at BACs approaching and exceeding 0.15%. At 0.25% BAC and higher, the relative risks for the three conditional models are 8.15, 8.56 and 4.50 points higher with inclusion of recovered hit-and-runs. Only 15.6% of all hit-and-runs were recovered and tested, and it can be reasonably assumed that the relative risk at high BACs would be much greater if adjusted for the remainder (84.4%).

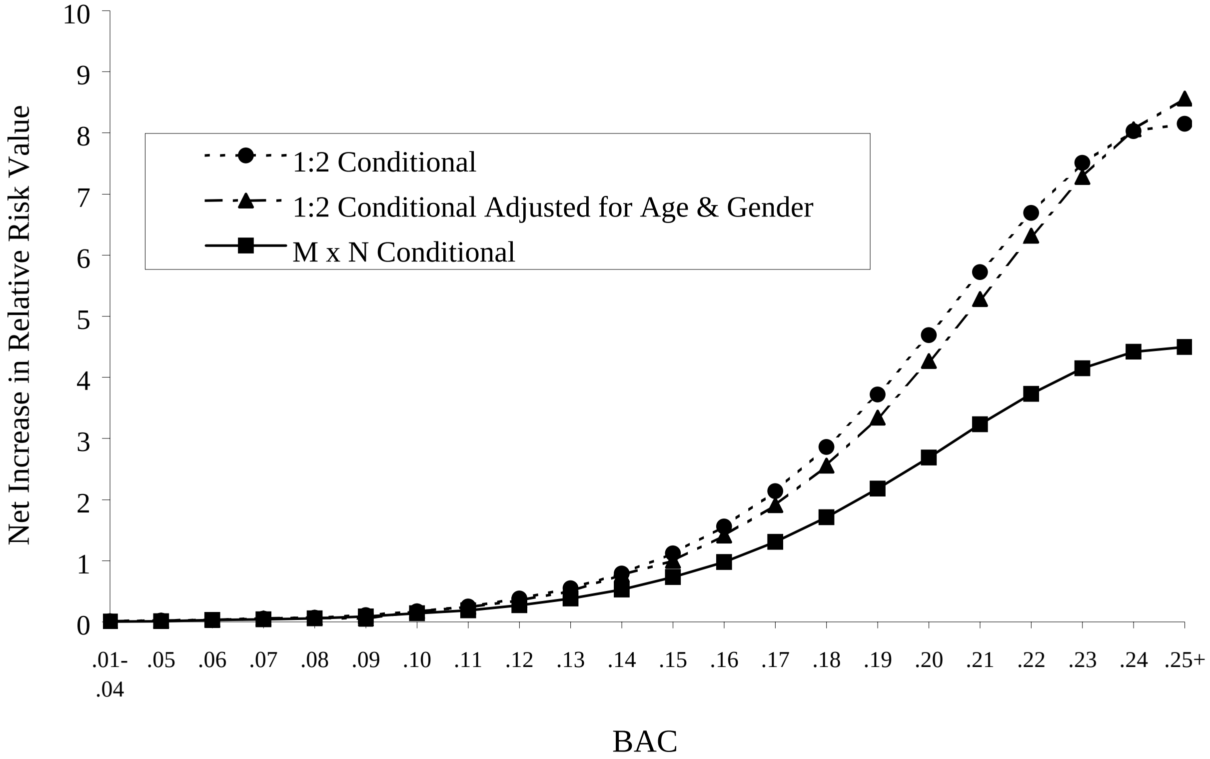

Figure 4. Changes in relative risk (odds ratios) as a function of inclusion of recovered hit-and-run drivers

Table 15. Relative Risk Points (Odds Ratios) Changes as a Function of Inclusion of Recovered Hit-and-Run Drivers

BAC level	1:2 conditional	1:2 conditional adjusted for age & gender	M x N conditional
.01-.04	.01	.01	.005
.05	.02	.02	.01
.06	.03	.03	.03
.07	.05	.06	.04
.08	.07	.07	.06
.09	.11	.05	.09
.10	.17	.17	.14
.11	.25	.24	.19
.12	.38	.35	.27
.13	.55	.50	.38
.14	.79	.77	.53
.15	1.12	1.00	.73
.16	1.56	1.41	.98
.17	2.14	1.91	1.31
.18	2.86	2.56	1.71
.19	3.72	3.34	2.18
.20	4.69	4.27	2.69
.21	5.72	5.28	3.23
.22	6.69	6.32	3.73
.23	7.51	7.28	4.15
.24	8.03	8.06	4.42
.25+	8.15	8.56	4.50

4.6.4 Site, Age and Gender – Main Effects and Interactions

The analyses provide no evidence that the relationship between BAC and crash risk differed by study site (Table 16). A significant site x BAC interaction would have demonstrated that the BAC - crash risk relationship differed between Long Beach and Fort Lauderdale. None of the polynomial components produced a significant interaction with site, however, and the simultaneous test of the three interaction terms did not approach significance (*p* > .50). Although not significan**t,** tests of interaction, particularly in the presence of multicollinearity, tend to have low sensitivity (McClelland and Judd, 1993). For this reason, the separate relative risk curves for the two study sites are examined in a subsequent section.

Table 16. Test of Site by BAC Interaction in Conditional 1:2 Logistic Regression Model

(*N* = 11,994)

Variable	Df	Coefficient	S.E.	Wald chi-square	p
BAC linear	1	-7.91	4.69	2.85	.092
BAC quadratic	1	266.1	79.6	11.2	.0008
BAC cubic	1	-663.3	284.5	5.4	.02
BAC x site (LB = 0; FL = 1)	1	2.07	5.82	0.13	.72
BAC^2 x site	1	-2.01	100.0	0.0004	.98
BAC^3 x site	1	-140.6	361.2	0.15	.70

Increase in likelihood ratio chi-square due to inclusion of interaction terms = 459.6 - 459.4 = 0.2, $df = 3$, $p > .50$.

Table 17 summarizes the analysis of the BAC – crash relationship adjusted for the main effects of age and gender. Gender was coded as a binary dummy variable (0 = female; 1 = male), and age was coded as a six-category nominal variable with drivers aged under 21 serving as the omitted referent group. The regression coefficients represent the unit change (log odds or “logit”) in the crash probability of each group compared to the referent group after adjusting for correlation with other covariates. The odds ratio shown in the last column is the exponentiated (e^x) form of the regression coefficient in which no relationship (regression coefficient = 0) would result in an odds ratio of 1.0.

The results show a .168 reduction in the probability of crash-involvement for males compared to females. That is, males have a relative crash risk expectancy that is $e^{-.168} = .846$ that of females. If the inverse of this odds ratio is calculated, the .846 odds ratio can be expressed as a proportional increase (1.00/.846 = 1.18), indicating that females have an 18% greater crash risk. It is possible that this counterintuitive finding is an artifact of the missing hit-and-run drivers (recovered hit-and-run drivers were 71.2% male) and non-participants (61% male).

Each age group was found to be associated with a significantly lower crash risk than the youngest age group (under 21). For example, the crash risk for drivers aged 45-54 years is only .43 as high as the risk for drivers under 21 years. Given that gender and age have been shown in many studies to be related to crash risk, it is reasonable that the inclusion of these variables increased the predictive power of the model (Cox-Snell statistic: R^2.046, Table 43; $R^2 = .038$, Table 38 without age and gender).

A meaningful odds ratio interpretation for the individual polynomial terms of BAC is not possible. The results do show the same significant cubic polynomial structure as obtained when age and gender were not controlled, although the coefficients differ somewhat. Table 18 and Figure 5 show the relative risks.

Table 17. Conditional 1:2 Logistic Regression of BAC-Crash Risk Relationship Adjusted for Age and Gender, Main Effects Only

***N* = 11,979**

Variable	*df*	Coefficient	S.E.	Wald chi-square	*P*	Adjusted Odds Ratio
Gender (females = 0)	1	-0.168	043	15.5	< .0001	.846
Age under 21 (referent)	-	-	-	-	-	1.00
Age 21-24	1	-0.451	.093	23.7	< .0001	.637
Age 25-34	1	-0.698	.079	79.0	< .0001	.498
Age 35-44	1	-0.772	.081	91.1	< .0001	.462
Age 45-54	1	-0.845	.089	90.7	< .0001	.430
Age 55+	1	-0.497	.090	30.3	< .0001	.609
BAC linear	1	-4.00	2.82	2.01	0.16	-
BAC quadratic	1	231.1	48.8	22.4	< .0001	-
BAC cubic	1	-627.3	177.4	12.5	.0004	-

Cox-Snell R^2 = .046; Likelihood ratio chi-square = 594.8, *df* = 9, *p* < .0001.

Comparing the relative risks in Table 18 to those produced when BACs were not adjusted for age and gender (Table 14), the risk slopes are somewhat steeper and the dip at low BACs is smaller. The relative risk is 1.00 for drivers at 0.04% BAC (0.035% - 0.044% interval). At 0.25%+, the highest BAC, the relative risk is 36.91 (hit-and-runs included) compared to an odds ratio of 31.42 in Table 14. Also note that the slight non-monotonicity in the relative risk values that can be seen in Table 14 at 0.025%+ BACs does not occur in the age-gender adjusted curve. The effect of including recovered hit-and-runs is similar to the effect that can be seen in Table 14 and Figure 3. The relative risk curves diverge as BACs increase due to high BACs among hit-and-run drivers.

The results shown in Table 18 are based on an additive model in the logits (i.e., no age x BAC, gender x BAC or age x gender x BAC interaction). The assumption that an additive model was appropriate was tested for selected interactions involving age, gender and site (Table 19).

The significant overall interaction (*p* < .01) is almost entirely attributable to the only interaction that exceeded conventional significance levels (site x gender *p* < .0001). This indicates that the gender – crash risk relationship differed at the two sites. A further examination of the distributions of gender and crash and control drivers by site (see Table F-12 in Appendix F) clarified the structure of the interaction. In Long Beach, females were significantly more likely to be members of the crash group (*p* < .0001). In Fort Lauderdale, there was a slight but non-significant trend in the opposite direction (*p* = .44). Since gender, site, and the site x gender interaction did not interact with BAC, however, this finding is not germane to the determination of BAC-mediated relative risk. It is sufficiently curious, however, to warrant further comment.

Table 18. Age and Gender Adjusted Relative Risks 1:2 Conditional Logistic Regression Cubic Polynomial Model

BAC	Hit-and-runs included	Hit-and-runs excluded
0.00	1.0	1.0
.01	.94	.93
.02	.92	.91
.03	.94	.93
.04	1.00	.98
.05	1.10	1.08
.06	1.25	1.22
.07	1.47	1.41
.08	1.75	1.68
.09	2.14	2.09
.10	2.66	2.49
.11	3.34	3.10
.12	4.24	3.89
.13	5.40	4.90
.14	6.89	6.17
.15	8.76	7.75
.16	11.08	9.67
.17	13.86	11.95
.18	17.11	14.55
.19	20.74	17.40
.20	24.63	20.36
.21	28.51	23.23
.22	32.06	25.74
.23	34.89	27.61
.24	36.61	28.55
.25+	36.91	28.35

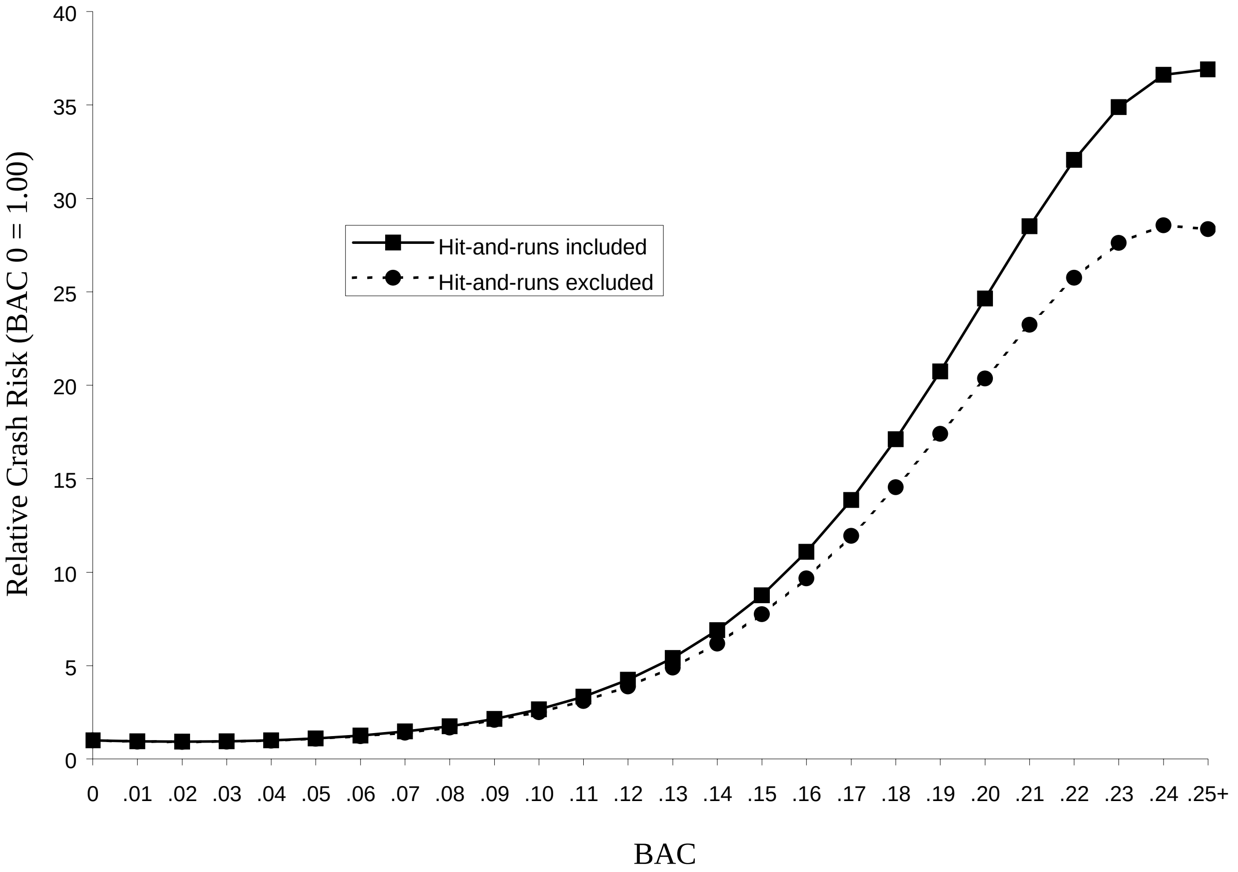

Figure 5. Age and gender-adjusted relative risks with and without hit-and-runs

An over-involvement of females in crashes in the Long Beach data is both counterintuitive and inconsistent with other research data. The most obvious explanation is that most of the hit-and-run drivers were male. If the 293 non-recovered Long Beach hit-and-run drivers were allocated to the control and crash groups in proportion to their representation in the recovered group, the gender relationship would reverse. Similarly, proportional allocation of the 206 non-recovered Fort Lauderdale hit-and-runs would cause males to be significantly over-represented in the crash group. Also, those who refused to participate were more likely to have been male crash drivers, which created a similar bias.

No other interaction was significant (Table 19), but each of the age group x BAC interactions approached significance. The interaction terms are all negative and of similar magnitude. This finding suggests that drivers under age 21 have a steeper linear relative risk gradient than older drivers, but that the older age groups do not differ much among themselves with respect to the BAC - crash risk relationship.

Table 19. Summary of Two-Way and Three-Way Interaction Tests, BAC (linear), Age, Gender and Site, Conditional 1:2 Model

***N* = 11,994**

Interaction parameter*	Coefficient	Chi-square	*p*-value**
Age 21-24 x site	-0.18	.75	.39
Age 25-34 x site	.14	0.57	.45
Age 35-44 x site	.016	0.007	.93
Age 45-54 x site	-.04	0.04	.84
Age 55+ x site	-.23	1.23	.27
Site x BAC	-10.9	0.83	.36
Age 21-24 x BAC	-16.2	2.60	.11
Age 25-34 x BAC	-14.5	2.28	.13
Age 35-44 x BAC	-16.3	2.89	.09
Age 45-54 x BAC	-15.3	2.40	.12
Age 55+ x BAC	-18.1	3.13	.077
Site x gender* **	0.37	17.5	.0001
Gender x BAC	1.08	.115	.70
Site x age 21-24 x BAC	10.0	.64	.42
Site x age 25-34 x BAC	11.4	.92	.34
Site x age 35-44 x BAC	13.2	1.23	.27
Site x age 45-54 x BAC	8.15	0.45	.50
Site x age 55+ x BAC	13.9	1.21	.27
Gender x site x BAC	-1.66	.19	.66

*Main effects and BAC polynomials included in model but not shown in order to simplify display. The main effect of site was constrained to be 0 by the matched case-control design. The interactions with BAC were limited to the linear component.

**Likelihood ratio chi-square for net significance of interaction terms = 38.9, *df* = 19, *p* < .01.

***Female coded 0 and Fort Lauderdale 1 in this and all other analyses.

4.6.5 Covariate Adjusted Equations

A series of logistic regressions began with relatively small covariate pools and culminated in an equation with all covariates. A more conventional procedure would have been a single analysis with the total covariate pool, but the covariates were not included in the data collection for a single reason. Some were collected based on a priori interest, and others were expected to elucidate the BAC – crash risk relationship. Traditional covariates such as age and gender have been shown to be related to alcohol consumption, driving while impaired and crash risk and frequently are controlled in cohort and case-control studies through matching or statistical adjustment. Demographic variables (education, ethnicity, socioeconomic status) are potential confounders of the alcohol – crash risk relationship.

Driving mileage is directly related to crash risk and must be controlled in any study of this nature.

Alcohol consumption variables are causally related to BAC, and past studies have shown them to be mediators of the BAC – crash risk relationship (Hurst, Harte and Frith, 1994). Their inclusion as covariates, however, presents some subtle complexities. First, the random and non-random measurement error of self-reported alcohol consumption can introduce distortion, particularly if the errors are non-random in relation to alcohol consumption and outcome status (Cohen and Cohen, 1983; Breslow, 1994). Second, the BAC – crash risk relationship may be adjusted for part of its own cause. If, for example, tolerance develops with heavy drinking and attenuates crash risk, this fact would be functionally related to both BAC and crash risk. Conversely, if some heavy drinkers are more impaired at given alcohol levels, more prone to take risks, or more likely to exercise poor judgment, these facts would be related to both BAC and crash risk. Under these hypothetical scenarios, alcohol consumption covariates would adjust the BAC – crash risk relationship for variables that are functionally related to the occurrence of alcohol-related crashes and BAC.

Other covariates with plausible causal connection to drinking patterns and BAC include sleep patterns, safety belt use, DUI offenses, and trip origin. Such relationships can be better handled with path analysis, hierarchical regression and structural equation models, all of which were beyond the scope of the present analysis. Instead, equations were produced that include and exclude potentially reactive covariates to determine whether they affect calculated relative risk. In some instances, an inspection of the regression coefficients and relative risk curves as a function of the included covariate set is informative. In all instances, performing the analyses both ways, which likely bounds the true state, provides a more complete picture of the underlying phenomena. The following logistic regression covariate pools were examined:

- BAC alone and BAC adjusted for age and gender.
- Age, gender, and driver responses to alcohol consumption questions.
- All non-reactive demographic covariates, excluding responses to alcohol consumption questions and covariates that are potentially affected by alcohol consumption (e.g., sleep patterns, DUIs, self-reported driving after drinking).
- Total pool of covariates, including alcohol consumption responses and sleep patterns.
- Total pool of covariates, excluding alcohol consumption responses.
- Total pool of covariates, including time of crash to evaluate time x BAC interaction.
- Total pool of covariates, including crash severity to evaluate of severity x BAC interaction.
- Total pool of covariates, including number of vehicles (single vs. multiple) to evaluate single vehicle x BAC interaction.

All equations are based on subjects with valid scores on all covariates, reducing the sample size of drivers to N= 11,373. The univariate BAC and age/gender-adjusted equations presented earlier were based on a larger N and were less subject to bias. Presentation of each reduced N equation, however, facilitates comparisons across covariate pools and provides insight into biases created by missing data.

BAC, age and gender. The BAC parameter estimates for two regressions, one without covariates and the other including age and gender, are shown in Table 20. Both show the same cubic polynomial structure as obtained in the previous equations, and the value of coefficients for the polynomial terms are quite similar to those shown in Table 11 and Table 17. These equations, however, are subject to small biases due to the deletion of subjects who did not have valid scores on all of the covariates. The Cox-Snell R^2 for the model with age and gender is .044; the Cox-Snell R^2 for the model with the larger sample was .046 (Table 17).

Table 20. 1:2 Reduced Sample Conditional Logistic Regression of BAC-Crash Risk Relationship Adjusted and Unadjusted for Age and Gender
***N* = 11,373**

	Adjusted		Unadjusted	
BAC parameter	**Regression coefficient**	***p* value of Wald chi-square**	**Regression coefficient**	***p* value of Wald chi-square**
Linear	-4.72	.13	-7.65	< .05
Quadratic	234.3	< .001	266.1	< .001
Cubic	-654.4	< .001	-750.8	< .001

Age, gender and alcohol consumption. The following variables were entered into a logistic regression model: gender, gender x location interaction, age group, preferred alcohol beverage, number of drinks per drinking day in past month (categorical groupings 0 - 8+), number of drinking days in past month (continuous variable), a binge drinking index (number of times six or more drinks were consumed during past month) and a deception index.

The deception index was derived from drivers whose denial of drinking or driving after drinking was contradicted by their measured BACs. Drivers with positive BACs who denied ever driving after drinking were coded 1, and all others were coded 0.[12] Because only drivers with positive BACs could receive a value of 1, this coding convention results in some non-orthogonality between BAC status and deception.

The structure of the alcohol questions was based on a sophisticated mathematical model of the probability of continued drinking once drinking is initiated (Gruenewald and Nephew, 1994). The entire model was not used for this study. Rather, the responses to the survey questions were computer-transformed to traditional quantity-frequency scales.

The results of the logistic regression combining age, gender and alcohol consumption patterns are summarized in Table 21. The binge-drinking index did not remain in the final

---

[12] All others would be positive BAC drivers who were truthful and all those with 0.0 BAC.

equation ($p \geq .20$), and it also failed to show a significant relationship in any subsequent equations.

Table 21. Logistic Regression – BAC, Age, Gender and Alcohol Consumption Variables
(*N* = 11,373)

Covariates	Regression Coefficient	*Df*	Wald chi-square	*p*	Adjusted odds ratio
Gender	-.283	1	19.9	<.0001	.75
Gender x site	.366	1	17.1	<.0001	1.44
Age					
Under 21 (ref.)	--	--	--	--	1.00
21-24	-.411	1	17.6	< .0001	.66
25-34	-.657	1	61.4	< .0001	.52
35-44	-.776	1	80.9	< .0001	.46
45-54	-.822	1	76.2	< .0001	.44
55+	-.484	1	25.4	< .0001	.62
Number of drinking days past month	-.018	1	17.8	< .0001	.98
Number of drinks per setting	--	--	--	--	--
0 (ref.)	--	--	--	--	1.00
1	.057	1	.48	.49	1.06
2	-.018	1	.037	.85	.98
3-4	-.276	1	6.94	.008	.76
5-7	-.451	1	6.46	.011	.64
8+	-.413	1	.50	.48	.66
Deception (1=yes)	.636	1	26.8	< .0001	1.89
Favorite alcohol beverage	--	--	--	--	--
None (ref.)	--	--	--	--	1.00
Beer	-.161	1	3.93	.048	.85
Liquor	-.311	1	10.53	.001	.73
Wine	-.164	1	3.63	.057	.85
BAC	-8.72	1	4.80	.029	--
BAC^2	303.3	1	25.3	< .0001	--
BAC^3	-860.2	1	17.3	< .0001	--

Cox-Snell $R^2 = .055$; likelihood ratio chi-square = 630.1; $df = 20$; $p < .0001$.

Gender, age group, and gender x site interaction were highly significant. The direction and magnitude of the relationships do not differ materially from those obtained without the alcohol consumption pattern variables.

Several alcohol consumption variables shown in Table 21 are highly significant predictors of crash-control status ($p < .001$). Compared to control drivers, crash drivers more often describe themselves as non-drinkers, particularly of hard liquor ($p < .002$), report fewer drinking days per month ($p < .0001$), and report an average of one or fewer drinks per setting.

For all other drink categories (2 to 8+ drinks), drinkers were less likely to be crash involved than those reporting themselves to be non-drinkers, and the differences were significant for the 3 to 4 and 5 to 7 drink categories. Eight or more drinks per setting was associated with the lowest crash involvement expectancy (odds ratio = .66), but the difference is not statistically significant, perhaps due to the small number of drivers in this category.

The alcohol deception variable proved to be a highly significant predictor. Deceptive drivers were 1.89 times more likely to be crash involved ($p < .0001$). This finding is critically important to an interpretation of study results, because it demonstrates that denial of drinking or denial of ever driving after drinking alcohol occurred more often among crash-involved drivers than among controls. Because the deception index could be calculated only for drivers with positive BACs, it provides no evidence about the truthfulness of drivers (86%) who had zero BACs.

These findings seem to indicate that drinkers are substantially less likely than non-drinkers to be crash involved and that crash risk declines further as drinking increases, but there are potential biases that must be recognized. Probably the most important limitation is that the data do not reflect the alcohol consumption of non-recovered hit-and-runs, non-participants (refusals), and those who did not answer the alcohol-consumption questions when interviewed. Also, alcohol consumption is often underreported, particularly among heavy drinkers and alcoholics.

The BAC parameters in Table 21 show the same statistically significant cubic polynomial pattern as the previous equations, but all coefficients have increased in magnitude. The Cox-Snell R^2 for the model is .055. The relative risk curve resulting from the coefficients for the BAC parameters is presented later in this section along with those of the other logistic regression models.

Non-reactive demographic covariates. Table 22 shows the logistic regression equation with all non-reactive demographic covariates. The variable pool for this analysis (with all covariates that were not potential direct causes or effects of alcohol consumption and BAC) included demographic variables (age, gender, ethnicity, socioeconomic indices) and driving habit information. Table 22 and subsequent tables contain global significance tests of the covariate domains containing more than two levels (e.g., marital status, ethnicity, etc.) These global tests indicate the statistical significance of a given classification when all groups or levels that define the categorical variable are considered as a set.

Type of occupation, average weekly mileage, number of passengers, and non-English speaking were not significant at $p > .20$ and were not retained in the demographic equation or in any of the subsequent logistic regressions including the all covariates analyses. There is no evidence that these variables are independently associated with crash risk in the context of this study. Prior DUI convictions, favorite alcoholic beverage, use of medication/drugs and safety belt usage also did not meet the criterion for retention ($p \leq .20$).

Weekly mileage did not show a significant relationship with crash risk, possibly due to the site – time matched sampling design. More miles per week indicate the drivers are more likely to be on the road at a given location and time of day, and those variables were controlled by the design.

Table 22. Logistic Regression Including BAC and All Non-Reactive Demographic Covariates ($p \leq 20$)

N = 11,373

Covariates	Regression coefficient	*Df*	Wald chi-square	*p*	Adjusted odds ratio
Marital status	--	4	11.1	.026	--
Single (ref.)	--	--	--	--	1.00
Cohab	-.134	1	1.19	.27	.87
Divorced or separate	.033	1	.17	.68	1.03
Widowed	.383	1	5.88	.015	1.47
Married	-.074	1	1.90	.17	.93
Education	--	6	24.9	< .0001	--
Grammar (8 or less)	.427	1	17.9	< .0001	1.53
9-11	.291	1	9.42	.002	1.34
High school (12)	.175	1	6.28	.012	1.19
13-15	.073	1	1.18	.28	1.08
16 (BA) (ref.)	--	--	--	--	1.00
17-18	.067	1	.53	.47	1.07
19+ (doctorate)	.200	1	2.93	.087	1.22
Ethnicity	--	4	37.9	< .0001	--
Caucasian (ref.)	--	--	--	--	1.00
African American	-.084	1	1.57	.21	.92
Asian	.051	1	.28	.59	1.05
Latino	-.226	1	11.0	.0009	.80
All others	.313	1	14.5	< .0001	1.37
Employment status	--	6	44.9	< .0001	--
Full time (ref.)	--	--	--	--	1.00
Part-time	.115	1	2.30	.123	1.12
Unemployed	.541	1	29.1	< .0001	1.72
Retired	.188	1	2.57	.109	1.21
Student	.290	1	8.72	.003	1.34
Home maker	.0033	1	.0005	.98	1.00
Other/disabled	.467	1	11.67	.0006	1.60
Vehicle type	--	6	30.5	< .0001	--
4-door sedan (ref.)	--	--	--	--	1.00
2-door coupe	-.237	1	19.6	< .0001	.79
Sport	-.101	1	.58	.44	.90
Pick up	.058	1	.59	.44	1.06
Convertible	-.438	1	4.52	.033	.65
Van/wagon	.044	1	.35	.55	1.05
All others	-.021	1	.05	.82	.98
Gender (0=female)	-.331	1	26.1	< .0001	.72
Gender x site	.345	1	15.0	< .0001	1.41
Age group	--	5	50.8	< .0001	--
Under 21 (ref.)	--	--	--	--	1.00
21-24	-.311	1	8.97	.003	.73
25-34	-.521	1	28.7	< .0001	.59
35-44	-.638	1	38.0	< .0001	.53
45-54	-.721	1	40.8	< .0001	.49
55+	-.492	1	16.0	< .0001	.61
BAC	-4.97	1	2.46	.115	--
BAC^2	237.5	1	20.2	< .0001	--
BAC^3	-661.0	1	12.1	<.0005	--

Cox-Snell R^2 = .057; likelihood ratio chi-square = 652.4; df = 36; p < .0001.

Type of occupation did not show a significant relationship with crash risk although education and employment status (unemployed, employed, retired, student, etc.) did prove to be significant predictors. The failure to find a type of occupation – crash risk relationship is believed to reflect inconsistencies in the classifications of drivers' occupations. Both sites encountered problems with the complex taxonomy that was adopted for the study.[13]

Covariates from the following domains were related to crash risk: marital status, education, ethnicity, employment status, vehicle type, gender, age and BAC. The Cox-Snell R^2 for the model is .057. Table 22 includes all variables significant at $p \leq .20$. Many were also significant at $p \leq .05$.

Compared to controls, crash-involved drivers were:

- More likely to be widowed.
- More likely to have less than 9 years and less than 12 years schooling.
- Less likely to be Latino and more likely to be classified as Caucasian or "other" (Native Americans, Pacific Islanders, multiethnic).
- More likely to be unemployed, a student or disabled.
- Less likely to be driving two-door coupes or convertibles and more likely to be driving pickups and vans.
- In Long Beach, more likely to be female.
- More likely to be under 21.

Inspection of the relative risk curve (to be presented later in Figure 12) indicates that crash drivers were more likely to have 0.05% and higher BACs. The inclusion of additional demographic covariates increased the predictive power of the model compared to a model with only age and gender, but the coefficients for BAC do not differ substantially from those of the age-gender adjusted equation (Table 20). They do differ from the equation that included alcohol consumption variables (Table 21).

All covariates. Two regressions were computed with the entire covariate pool, including alcohol consumption, sleep, trip origin, and demographic variables, self-reported drug and medicine use, and consequences of alcohol consumption (DUIs, driving after drinking). One equation included all variables significant at $p \leq .20$ (Table 23 which begins on the next page and carries over to the following page), and the other used all variables significant at $p \leq .05$ (see Appendix G, Table G-4).

---

[13] The main problem was that many subjects early in the study answered the question "What kind of work do you do?" with a response such as "I work for Company *X*." The interviewers recorded this rather than probing for a more definitive statement of occupation. The problem was corrected towards the middle of the study, but this left the occupational classification of many subjects somewhat uncertain.

Table 23. Logistic Regression Equation with All Covariates (Retention $p \leq .20$)

N = 11,373

Variable	Regression Coefficient	*Df*	Wald Chi-square	*P*	Adjusted odds ratio
Marital status	--	4	14.5	.006	--
Single (ref.)	--	--	--	--	1.00
Cohabiting	-.130	1	1.05	.31	0.88
Divorced or separate	.086	1	1.09	.30	1.09
Widowed	.474	1	8.32	.004	1.61
Married	-.067	1	1.46	.23	.94
Education	--	6	11.62	.07	--
0-8	.325	1	9.51	.002	1.38
9-11	.189	1	3.66	.056	1.21
12 (high school)	.103	1	2.01	.16	1.11
13-15	.069	1	0.98	.32	1.07
16 (BA) (ref.)	--	--	--	--	1.00
17-18	.087	1	0.85	.36	1.09
19+ (Doctorate)	.180	1	2.19	.139	1.20
Ethnicity	--	4	34.5	< .0001	--
Caucasian (ref.)	--	--	--	--	1.00
African American	-.097	1	1.91	.168	0.91
Asian	.012	1	.013	.91	1.01
Latino	-.314	1	19.45	< .0001	0.73
All others	.193	1	5.00	.025	1.21
Employment status	--	6	23.10	< .001	--
Full time (ref.)	--	--	--	--	1.00
Part-time	.071	1	0.82	.37	1.07
Unemployed	.397	1	14.1	.0002	1.49
Retired	.065	1	0.28	.59	1.07
Student	.213	1	4.32	.038	1.24
Homemaker	-.107	1	0.40	.53	0.90
Other/disabled	.371	1	6.76	.009	1.45
Trip origin	--	4	37.3	< .0001	--
Bar (ref.)	--	--	--	--	1.00
Own home	-.127	1	2.17	.14	0.88
Others home	.208	1	4.90	.027	1.23
Work/business	.086	1	.988	.32	1.09
Other	.257	1	7.50	.006	1.29
Vehicle type	--	6	26.0	.0002	--
4-door sedan (ref.)	--	--	--	--	1.00
2-door coupe	-.225	1	16.48	< .0001	.80
Sports	-.046	1	.116	.73	.96
Pick up	.081	1	1.07	.30	1.08
Convertible	-.393	1	3.43	.064	.68
Van/wagon/sub.	-.019	1	.058	.81	1.02
All other vehicles	-.004	1	.002	.97	1.00
Age group	--	5	35.4	< .0001	--
Under 21 (ref.)	--	--	--	--	1.00
21-24	-.230	1	4.42	.036	.80
25-34	-.425	1	17.10	< .0001	.65
35-44	-.542	1	24.80	< .0001	.58
45-54	-.611	1	26.26	< .0001	.54
55+	-.387	1	8.97	.003	.68

Table 23 continued

Variable	Regression Coefficient	*Df*	Wald Chi-square	*P*	Adjusted odds ratio
Gender (0=female)	-.253	1	14.05	.002	.78
Gender x site	.363	1	15.47	< .0001	1.44
Commercial vehicle (yes=1)	.759	1	7.09	.008	2.14
Average hours slept (past week)	--	3	46.7	< .0001	--
0-4	-.622	1	14.9	.0001	.54
5-7	-.327	1	36.1	< .0001	.72
8-9 (ref.)	--	--	--	--	1.00
10+	.098	1	.91	.34	1.10
Hours awake	-.00081	1	23.7	< .0001	--
Driving after drinking	-.031	1	6.0	.015	--
Drinking days per month	-.015	1	9.7	.002	.985
Drinks per setting	--	5	20.6	.001	--
0 (ref.)	--	--	--	--	1.00
1	-.038	1	.38	.54	.96
2	-.108	1	2.02	.16	.90
3-4	-.352	1	15.3	< .0001	.70
5-7	-.460	1	7.02	.008	.63
8+	-.324	1	.28	.60	.72
Deception (1=yes)	.505	1	14.9	.0001	1.66
Hours last slept	--	3	60.25	< .0001	--
0-4	-.577	1	25.08	< .0001	.56
5-7	-.361	1	40.58	< .0001	.70
8-9 (ref.)	--	--	--	--	1.00
10+	.096	1	1.35	.25	1.10
BAC	-8.99	1	4.66	.03	--
BAC^2	308.4	1	24.24	< .0001	--
BAC^3	-865.6	1	16.18	< .0001	--

Cox-Snell R^2 = .098; likelihood ratio chi-square = 1112.4; *df* =56; *p* < .001.

The equation retained 56 variables. Some nominal variables are non-significant (p > .20) but were retained because they contain both significant and non-significant categories. For example, the contrast between the referent group (single) and three marital status groups (married, cohabitation and divorced/separated) yielded p >.20. The contrast with the widowed group, however, was significant, and all were retained to preserve the structure of the indicator variable contrast coding. This requirement was relaxed for the $p\leq$ .05equation (Appendix G, Table G-4). With few exceptions, the demographic covariates that were significant in the previous analyses were also statistically significant in the all-covariates equation.

With respect to the alcohol consumption-related covariates, the previously noted counterintuitive finding regarding the relationship between alcohol consumption and crash risk continues to be evident in these results. In fact, all non-zero levels of number of drinks per setting are directionally associated with decreased crash risk, and the associations are statistically significant for drivers averaging 3-4 and 5-7 drinks per setting. It is again found that the number of drinking days per month is inversely associated with crash risk, i.e., crash drivers claim to drink less often than the controls. Crash drivers also reported less frequent driving after consuming alcohol, but they also were significantly more likely to give deceptive responses to alcohol questions.

Trip origins for crash drivers were more often reported as someone's home or other location rather than a bar. The driver's own home was the only trip origin that was associated with less crash risk than a bar.

Compared to crash drivers, control drivers reported sleeping fewer hours during the prior week and fewer total hours on the last sleep occasion. Zero to four hours was reported both as average sleep and as the hours on the last sleep occasion, and both were associated with the lowest crash risk. These two low sleep groups were .54 and .56 times, respectively, as likely to be crash involved as drivers who reported sleeping 8 to 9 hours. Also, more elapsed hours since awakening were associated with a decreased likelihood of being in the crash group.

The regression coefficients for the two sleep variables, simultaneously adjusted for their intercorrelation ($r = .57$) are remarkably similar. The relationship between hours last slept and crash risk represents that relationship after it was adjusted for the effects of average sleep hours and the other covariates. The relationship between average sleep hours and crash risk was similarly partialled for these intercorrelations. The BAC estimates continue to reflect a cubic polynomial structure, but the coefficients are notably larger for all three polynomial orders. The Cox-Snell R^2 for the model is .098.

The $p \leq .05$ criterion for covariate retention reduced the number of parameters in the model to 33 (Appendix G, Table G-4). The pattern of significant covariates and the BAC estimates, however, are very similar to the model with 56 (Table 23). The Cox-Snell R^2 is almost as high as in the larger model (.096 vs. .098).

<u>All covariates except those related to alcohol-consumption</u>. Table 24 displays the results when alcohol consumption variables are deleted from the all-covariate equation; the actual equation is not shown. Values and significance level for the retained covariates were not appreciably affected except those for the BAC parameters, which were increased by the inclusion of alcohol consumption variables. The net contribution of BAC to the predictive power of the models is relatively unchanged, however, as indicated by the score statistic chi-square (313.0 vs. 316.4).

<u>Analyses of interactions by crash subtype</u>. Table 25 summarizes the logistic regression analyses of the crash type interactions. Three analyses, one for each crash type, included the appropriate BAC x crash subtype product terms in the all-covariates model ($p \leq .20$). Severity x BAC and number of vehicles x BAC interactions are not significant by the Wald chi-square or the hierarchical stepwise statistic (likelihood ratio chi-square). The time x linear BAC interaction for the cubic model is significant, and the *p* value for the likelihood ratio test is suggestive. Figure 6 graphs the logit slope for this interaction (a constant of four has been added to the logit scale to eliminate negative values). Daytime and early night crashes exhibit the familiar cubic pattern, including the dip at low BACs. Late-night crashes exhibit a relatively linear increase throughout the BAC range and do not show a dip.

Table 24. All Alcohol Consumption Variables Deleted from All- Covariates Equation ($p \leq .20$) N = 11,373

	Wald chi-square and p value	
Variable	**All covariates**	**After deletion of alcohol**
Marital status	14.5 (p = .006)	13.9 (p < .008)
Education	11.6 (p = .07)	13.1 (p = .04)
Ethnicity	34.5 (p < .001)	36.0 (p < .001)
Employment status	23.1 (p < .001)	25.3 (p < .003)
Trip origin	37.3 (p < .001)	38.1 (p < .001)
Vehicle type	26.0 (p =. 002)	26.7 (p = .002)
Average hours slept per week	46.7 (p < .001)	51.7 (p < .001)
Mean hours last slept	60.25 (p < .001)	65.2 (p < .001)
Hours awake	23.7 (p < .001)	22.5 (p < .001)
Commercial plates	7.09 (p = .008)	6.76 (p =. 009)
Gender	14.05 (p = .002)	21.9 (p < .001)
Gender x site	15.5 (p < .001)	15.1 (p = .001)
BAC linear coefficients	-8.99	-6.30
BAC quadratic coefficient	308.4	251.3
BAC cubic coefficient	-865.6	-686.0
Score statistic for BAC terms	313, df = 3, p < .001	316.4 df = 3, p < .001

Table 25. Test of BAC by Crash Type Interaction, All Variables Model N = 11,373

	Component			
Interactions	Linear	Quadratic	Cubic	Likelihood ratio chi-square for interaction (df = 3)
Time x BAC	14.8 (p < .04)	-173.4 (p = .14)	464.8 (p = .27)	5.4*
Severity x BAC	-6.39 (p = .45)	112.8 (p = .42)	-326.3 (p = .45)	3.5**
Single vs. multiple vehicles x BAC	-4.19 (p = .73)	231.7 (p = .25)	-948.4 (p = .17)	.07**

* p ~ .15
** p > .25

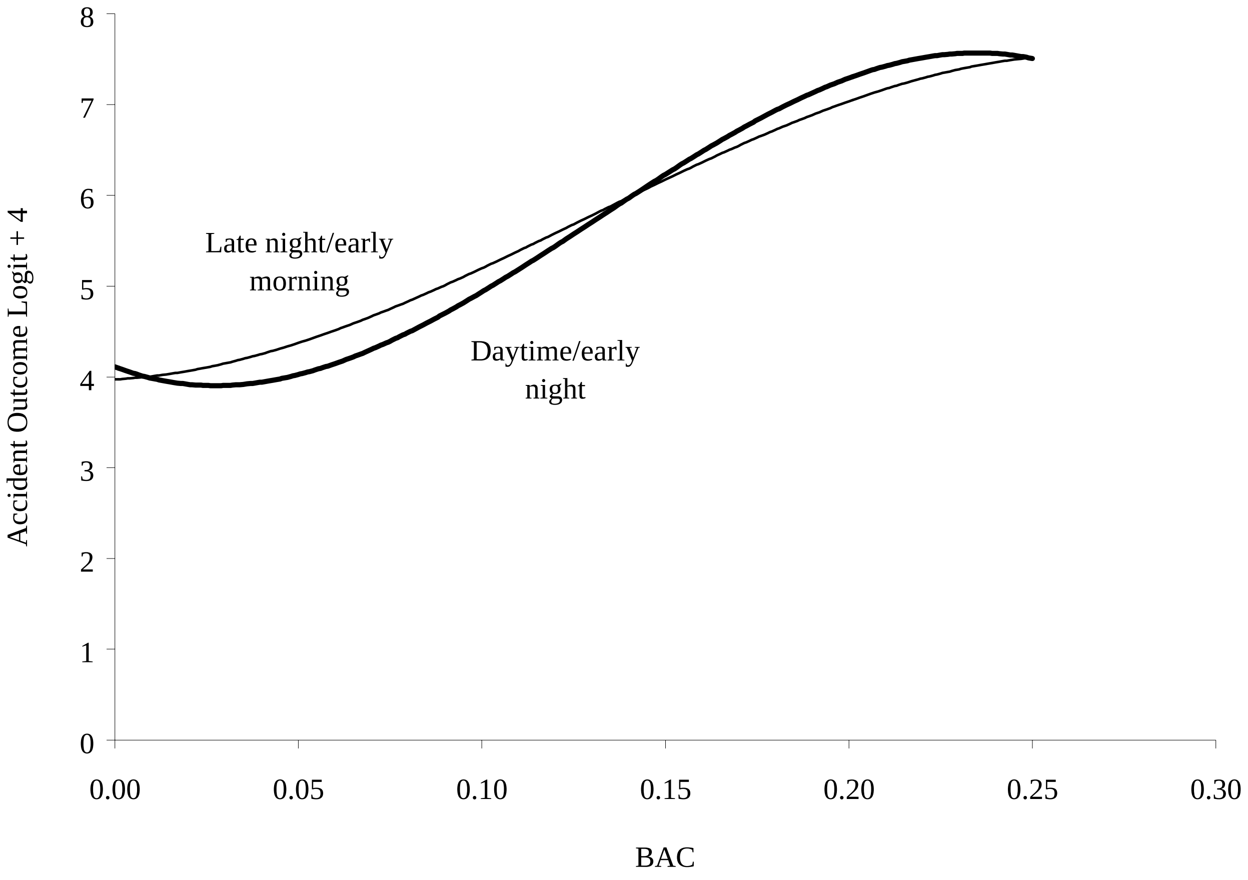

Figure 6. Crash time by BAC interaction

Small *Ns* for the less frequent types of crashes reduced the statistical power of the interaction tests. The proportion of crashes involving late-night, injury/fatalities and single vehicles were, respectively, 17%, 22% and 9%. Although the BAC x crash severity and BAC x single vs. multiple vehicle crashes interactions are not statistically significant, they have been graphed in Figures 7 and 8, and it can be seen that the slope differences are consistent with a priori expectations. Single vehicle and injury/fatal crashes exhibit steeper BAC risk gradients for 0.08 – 0.24% BACs.

Interactions with sleep and alcohol consumption covariates. Table 26 summarizes the evaluations of two-way interactions related to sleep and alcohol consumption. The evaluations used the general approach described above for analyses of crash types.

It is interesting to note that none of the equations show significant age and gender interactions. This result was expected based on the logistic regressions presented earlier. With the addition of alcohol consumption variables, however, a number of significant interactions emerge (BAC x 2 drinks per setting, BAC x 3 to 4 drinks per setting, BAC x 5 to 7 drinks per setting). The positive sign of the regression coefficients indicates that, compared to non-drinkers, the linear relationship between crash risk and BAC is significantly steeper among drivers who consume 2-7 drinks per setting than would be expected based on the main effect relationship between drinks per setting and crash risk. Recall that number of drinks per setting (3-4 and 5-7 drinks) was inversely associated with crash risk. These significant interactions indicate that the magnitude of that inverse relationship is not the same within the drinks per setting categories. Interpretation of these interactions is complicated by the fact that the association between BAC level and crash risk among “non drinkers” is attributable to drivers giving deceptive responses since, if all respondents were truthful, the BAC level of the 0-drinks group would have been 0.

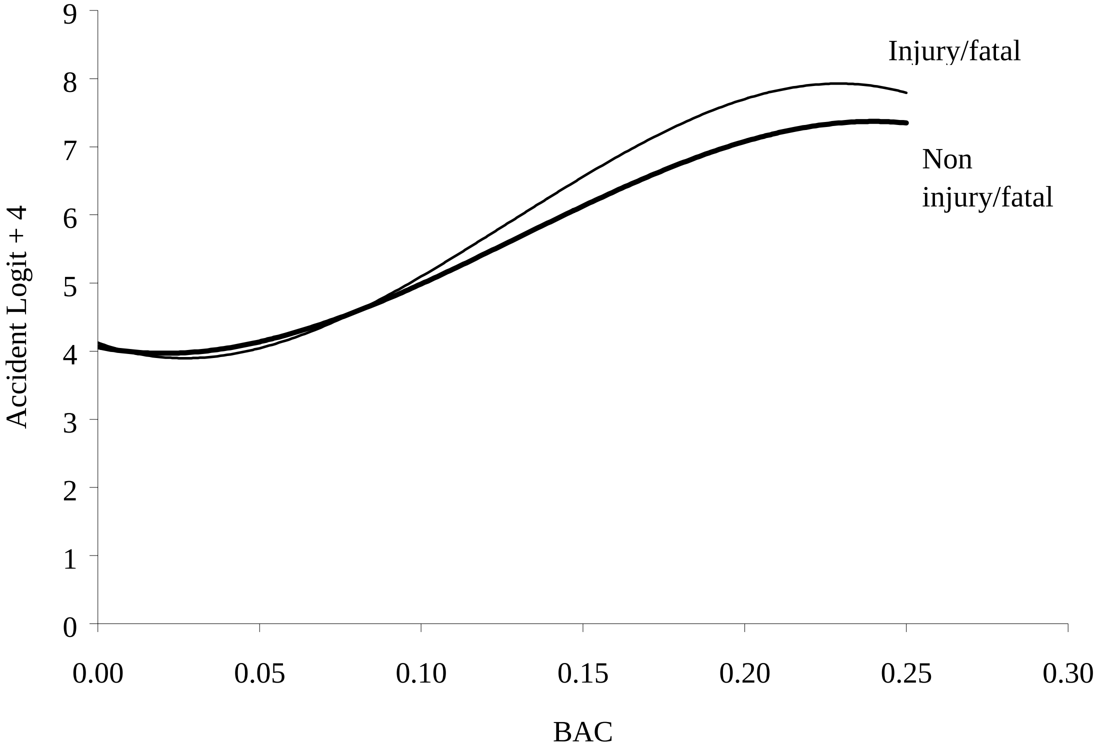

Figure 7. Crash severity by BAC interaction

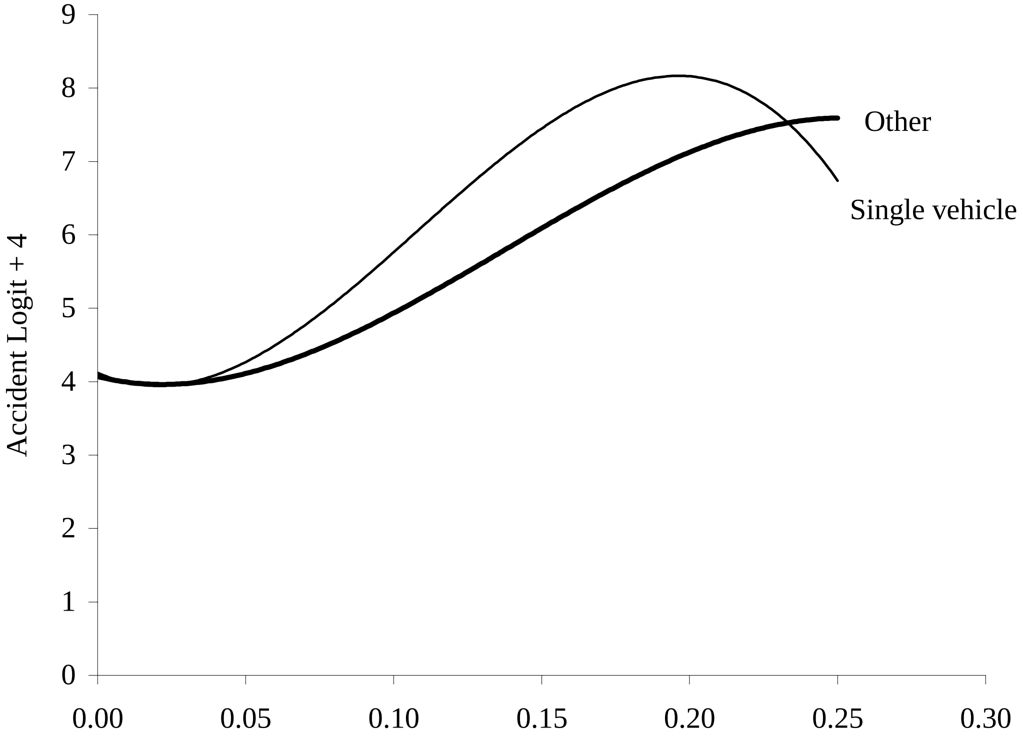

Figure 8. Crash type (single vehicle vs other) by BAC interaction

Table 26. Summary of Significant Interactions Involving Age, Gender, Sleep Hours and Alcohol Consumption Covariates (Logistic Regression Coefficients and *p* Values)

***N* = 11,373**

Covariate	**Equation** **Age and gender**	**Age, gender alcoholic consumption**	**All covariates ($p \leq .20$)**
BAC x gender	N.S. (p = .75)	N.S. (p = .63)	N.S. (p = .47)
BAC x age	N.S. (p = .76)	N.S. (p = .83)	N.S. (p = .71)
BAC x drink days per month	NA*	N.S. (p = .76)	N.S. (p = .69)
BAC x 1 drink per setting	NA	N.S. (p = .12)	N.S. (p < .07)
BAC x 2 drinks per setting	NA	9.19 (p = .02)	10.84 (p < .005)
BAC x 3-4 drinks per setting	NA	10.7 (p = .01)	12.15 (p < .002)
BAC x 5-7 drinks per setting	NA	3.35 (P = .07)	9.91 (p < .04)
BAC x 8 drinks per setting	NA	N.S.	28.9 (p < .20)
BAC x Deception	NA	N.S.	NA
BAC x 0-4 hours last slept	NA	NA	12.7 (p < .001)
BAC x 5-7 hours last slept	NA	NA	5.78 (p < .005)
BAC x 10+ hours last slept	NA	NA	N.S. (p = .72)
***p* value of likelihood ratio chi-square increase**	> .50	~ .25	< .001

*Note: NA Indicates that interaction was not applicable or was not evaluated. The *p* values are based on a comparison of likelihood ratios of models with and without (main effects only) interaction terms.

The stepwise likelihood ratio chi-square for the second equation in Table 26 requires clarification. The chi-square value (16.2, *df* = 13) is not significant (p ~ .25), which ordinarily would call into question the reality of the interactions. The set of interactions included the non-significant age x BAC and gender x BAC interactions, however, and the six degrees of freedom used to include these interactions reduced the sensitivity of the likelihood ratio chi-square test for all interactions considered as a set.

Figure 9 graphs the logistic regression main effect and interaction slopes.[14] The last equation in Table 26 includes both BAC x sleep and BAC x alcohol consumption interactions, and the overall likelihood ratio chi-square is highly significant ($p < .001$). The interactions with number of drinks per setting are significant or approach significance for drivers averaging 1 to 7 drinks per setting, and each interaction coefficient has a positive value. Figure 10 plots the interaction slopes (logits). Note the similarity of the slopes in Figures 9 and 10.

Although the maximum difference in slope is between non-drinkers and 8+ drinks per setting, the latter category is based on very small samples, and positive BACs for the 0-drink category reflect deception. The interaction slopes for 3 to 4 and 5 to 7 drinks per setting are statistically significant. Both groups have lower crash risks than the 0-drink group at 0.00% and very low BACs, but they have higher crash risks at $\geq$ 0.05% BACs. These interactions are difficult to interpret given the unknown reliability of self-reported alcohol consumption data and the bias introduced by missing data and non-participants.

The results reported in Table 26 also show highly significant interactions involving hours last slept. Recall that hours last slept had a significant main effect with crash risk, with drivers reporting 0-4 hours slept and 5-7 hours slept having lower crash risks than those reporting 8-9 hours last slept. The significant positive regression coefficient for the interaction parameters indicates that the relationship between BAC and crash risk is moderated by hours slept. These interactions are plotted in Figure 11. As expected, increasing levels of BAC are more deleterious in terms of increasing crash risk for drivers with less than average amounts of sleep, particularly 0-4 hours.

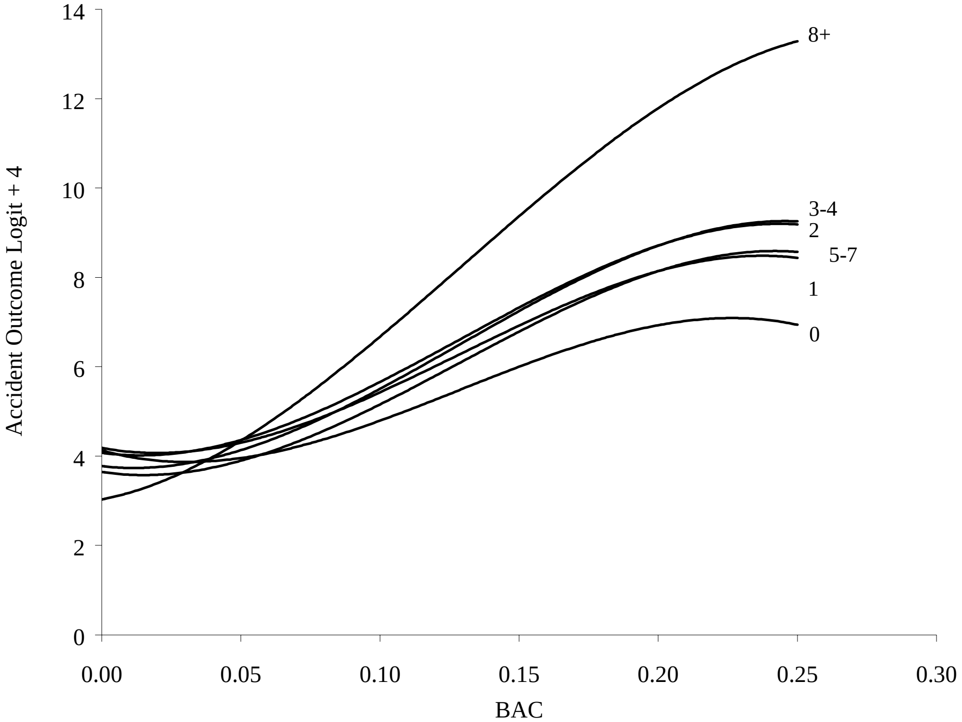

Figure 9. BAC linear by average number of drinks per setting interaction (BAC cubic equation)

---

[14] NOTE: When interactions are included in the model and plotted, the coefficients for the main effects are altered and the coefficients for the BAC terms are no longer identical to those shown in the "main effects" only model.

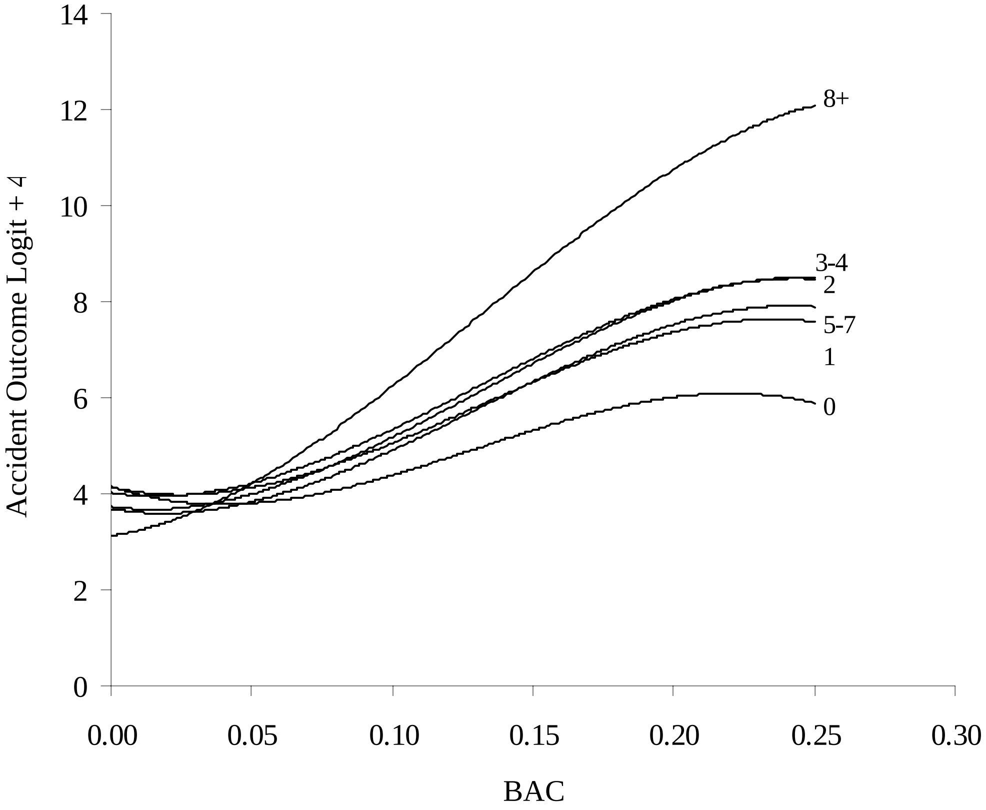

Figure 10. BAC linear by average number of drinks per setting interaction for model with all covariates (BAC cubic equation)

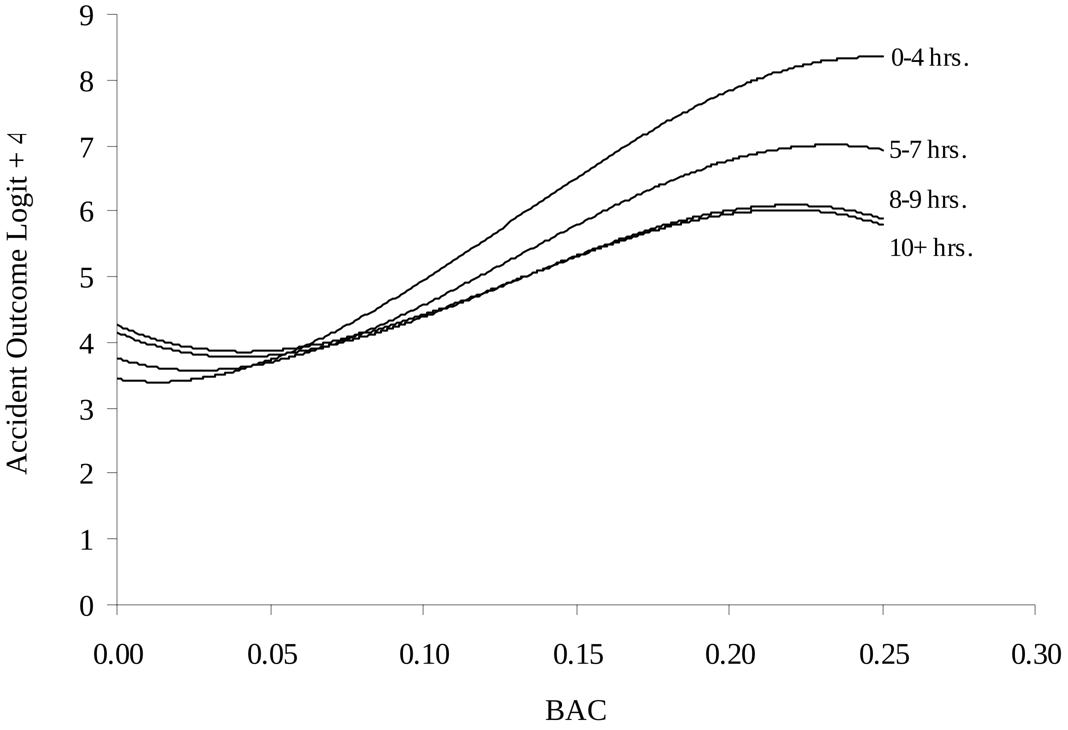

Figure 11. BAC linear by hours last slept interaction for model with parameters significant at p of .20 (BAC cubic equation)

Effects of covariates on BAC parameter estimates. Table 27 presents BAC parameter estimates from the various logistic regression equations. All models show a statistically significant cubic polynomial structure, but it is difficult to interpret the estimates for reasons noted previously. Age, gender and non-reactive demographic covariates reduce the size of the coefficients, but sleep and alcohol consumption increase the parameter values. The relative risk scores generated from each of the models (Table 28, Figure 12) provide a clearer understanding of the BAC-crash risk relationship.

Table 27. Summary of Parameter Estimates for BAC Polynomial Components

Equation	Linear	Quadratic	Cubic
BAC only 1 (*N* = 11,994)	-7.01	266.1	-730.4
BAC only 2 (11,373)	-7.65	266.1	-750.8
Age & gender 1 (*N* = 11,979)	-4.00	231.1	-627.3
Age & gender 2 (*N* = 11,373)	-4.72	234.3	-654.4
Age, gender & alcohol consumption indices (*N* = 11,373)	-8.72	303.3	-860.2
Age, gender & non-reactive demographic variables (*N* = 11,373)	-4.97	237.5	-661.0
All covariates – $p < .05$ (*N* = 11,373)	-9.29	312.9	-878.3
All covariates – $p \leq .20$ (*N* = 11,373)	-8.99	308.4	-865.6
All covariates excluding alcohol consumption variables – $p \leq .20$ (*N* = 11,373)	-6.30	251.3	-686.0

As covariates are added, the relative risk function becomes steeper at BACs above 0.10%. This is particularly true for the alcohol consumption variables, which also minimize the dip in risks at BACs below 0.05%. This can be seen in Table 28 but is hidden in Figure 12 by the graphic scale and the small absolute magnitude of the decline.

Table 28 does not contain risk values for an equation in which BAC is adjusted for age and gender only. A larger sample was used for the age and gender-adjusted values in Table 18, and the relative risks in Table 18 are fairly close to those of the non-reactive demographic equation until BACs exceed 0.22%, at which point they resemble those of the all covariates equation.

The study assessed the explanatory power of the regression models in general and of the predictive contribution of BAC in particular. The Cox-Snell R^2s increase as covariates are added (Table 29), but the most inclusive model (all covariates, $p \leq .20$) explains only 9.8% of the variance, which equates to a multiple R of .31 in ordinary least squares regression. The contribution of BAC is highly significant in all cases as indicated by the score statistic chi-square, and its incremental contribution is constant across all equations, varying from 3.0% - 3.2%.

Table 28. Relative Risks for BAC Levels in Equations After Deleting All Subjects With Missing Values on Covariates
***N* = 11,373**

BAC	No covariates	Non-reactive demographic covariates	All covariates except drinking consumption	All covariates $p \leq .20$	All covariates $p \leq .05$
0.00	1.00	1.00	1.00	1.00	1.00
.01	.91	.94	.92	.89	.89
.02	.87	.92	.89	.85	.84
.03	.87	.94	.91	.85	.85
.04	.92	1.00	.96	.90	.89
.05	1.00	1.10	1.05	1.00	.99
.06	1.13	1.25	1.19	1.14	1.14
.07	1.32	1.46	1.39	1.36	1.35
.08	1.57	1.74	1.66	1.68	1.66
.09	1.92	2.12	2.03	2.11	2.10
.10	2.37	2.62	2.53	2.70	2.70
.11	2.98	3.28	3.18	3.52	3.52
.12	3.77	4.14	4.14	4.62	4.63
.13	4.78	5.23	5.15	6.09	6.11
.14	6.05	6.60	6.57	8.02	8.06
.15	7.61	8.31	8.35	10.48	10.56
.16	9.48	10.35	10.52	13.55	13.67
.17	11.64	12.74	13.00	17.22	17.40
.18	14.00	15.43	16.06	21.40	21.65
.19	16.45	18.31	19.29	25.88	26.20
.20	18.78	21.20	22.60	30.29	30.68
.21	20.74	23.85	25.73	34.15	34.58
.22	22.07	25.99	28.37	36.88	37.30
.23	22.51	27.30	30.15	37.95	38.33
.24	21.92	27.55	30.75	37.03	37.31
.25+	20.29	26.60	29.99	34.08	34.22

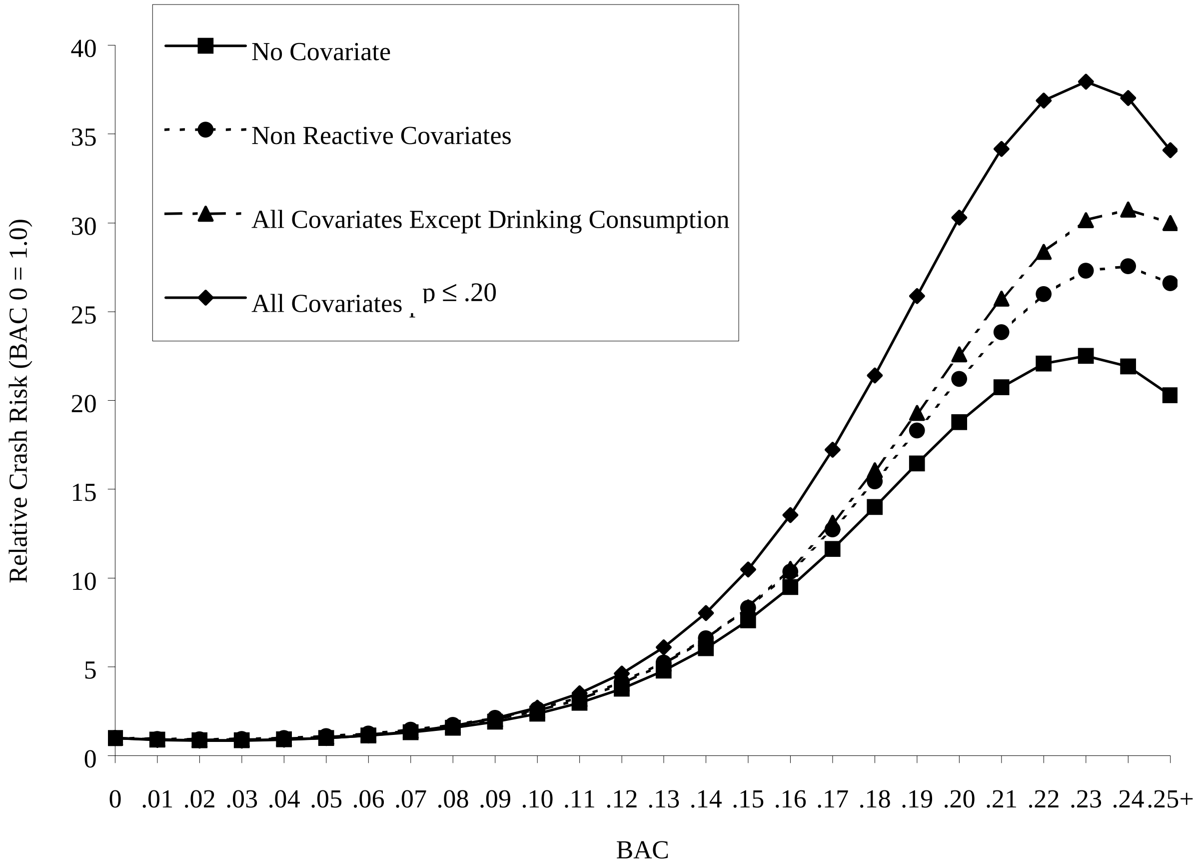

Figure 12. Relative risk by BAC as a function of covariate pool (all missing value subjects deleted)

Table 29. Explanatory Power of Logistic Regression Equations With and Without BAC Terms
N = 11,373

	Cox-Snell R^2 x 100			
Covariates	**BAC Excluded**	**BAC Included**	**Difference**	**Score test chi-square for increase***
Age & gender	1.2%	4.4%	+3.2%	363.6
Age, gender & alcohol consumption	2.3%	5.5%	+3.2%	351.3
All non-reactive	2.6%	5.7%	+3.1%	350.6
All covariates except alcohol consumption related	5.9%	8.9%	+3.0%	316.4
All covariates $p \leq .05$	6.5%	9.6%	+3.1%	317.1
All covariates $p \leq .20$	6.7%	9.8%	+3.1%	313.0

* $df = 3$; all p's ≤ .0001.

4.6.6 Adjusting Relative Risks for Bias

Three sources of bias were anticipated at the outset of the study and substantiated by the data analyses:

- Differential non-participation rates between the crash and control groups.

- Missing covariate data as a consequence of differential non-participation rates and the failure of subjects to complete the interview.

- Hit-and-run driver attrition from the crash group.

In all three instances, the effect of the bias was to underestimate the crash risk associated with elevated BACs. Alcohol positive crash drivers more often refused to participate or to complete the interview, and the hit-and-run drivers were more likely to be alcohol-impaired. It was therefore decided to adjust the computed relative risk values to compensate for the magnitude of these three biases as calculated in the analyses.

Data reported earlier showed that drivers who refused to participate either partially or completely had significantly higher BACs, as estimated by the PAS. To compensate for missing alcohol data, the BACs of drivers who refused to provide a breath specimen were estimated through an ordinary least squares regression procedure with PAS estimates. Valid BACs and PAS estimates were obtained from 12,584 subjects. The BACs were regressed against the PAS scores separately for the sample of crash and control drivers within each site, using a quadratic equation to account for the non-linearity of the relationship. Table 30 presents the multiple R^2s for the four resulting equations. The relationship between BACs and PAS scores is substantial for all four groups, but the relationship is stronger for the crash groups at both sites.

Table 30. R^2 for Regressions of BAC on PAS Scores by Site and Group
N = 12,584

Site and group	R^2
Long Beach controls	.59
Long Beach crashes	.74
Fort Lauderdale controls	.61
Fort Lauderdale crashes	.72
Total combined	.65

Note: Regressions based on a quadratic model.

The regression coefficients (linear and quadratic) for the PAS variable were used to estimate BACs for 232 drivers who refused to provide breath specimens but for whom PAS estimates were obtained. These BACs were added to the sample of 12,584 drivers with measured BACs to produce an enhanced sample of 12,816 subjects with measured or imputed BACs. A logistic regression was then computed on this sample to re-estimate the BAC – crash risk relationship using a cubic polynomial model. The relative risk estimates appear in the third column of Table 31, and the second column shows the relative risks prior to adding the imputed values. Figure 13 graphs the two curves.

Table 31. A Comparison of Relative Risks Adjusted for Non-Participation Bias (Refusal) with Unadjusted Relative Risks

BAC	Unadjusted (*N* = 12,584)	PAS adjusted (*N* = 12,816)	Net difference	Percent difference
0.00	1.00	1.00	--	.0
.01	.89	.90	+.01	1.1
.02	.84	.86	+.02	2.4
.03	.83	.85	+.02	2.4
.04	.85	.89	+.04	4.7
.05	.91	.96	+.05	5.5
.06	1.01	1.07	+.06	5.9
.07	1.15	1.24	+.09	7.8
.08	1.34	1.47	+.13	9.7
.09	1.60	1.77	+.17	10.6
.10	1.95	2.18	+.23	11.8
.11	2.41	2.71	+.30	12.4
.12	3.00	3.41	+.41	13.7
.13	3.76	4.29	+.53	14.1
.14	4.72	5.40	+.68	14.4
.15	5.90	6.76	+.86	14.6
.16	7.32	8.38	+1.06	14.5
.17	9.00	10.24	+1.24	13.8
.18	10.88	12.28	+1.40	12.9
.19	12.92	14.39	+1.47	11.4
.20	14.97	16.41	+1.44	9.6
.21	16.88	18.12	+1.24	7.3
.22	18.44	19.29	+.85	4.6
.23	19.43	19.71	+.28	1.4
.24	19.68	19.25	-.43	-2.2
.25+	19.07	17.89	-1.18	-6.2

The inclusion of imputed BACs increases the relative risks with the maximum difference occurring at 0.15% BAC. At this level, the relative risk was 14.6% higher after the addition of imputed values. The differences decline at very high BACs, and at 0.24% or 0.25% BAC the relative risks are lower than they were without the imputed values. This is believed to reflect the instability of quadratic regression estimates due to small *N*s at BAC extremes and the 0.12% BAC measurement ceiling of the PAS. Based on this finding, a correction factor of 14.6% for all BACs greater than 0.15% was therefore employed since the non-monotonicity in the curve was likely an artifact.

The relative risk values required yet another adjustment. Because there were PAS data for only 232 of the drivers who refused to provide breath specimens, it was necessary to increase each adjustment weight in Table 31 by a factor of 2.8 (656/232). For example, at 0.15% BAC the adjustment was 14.6% x 2.8 = 40.9%.

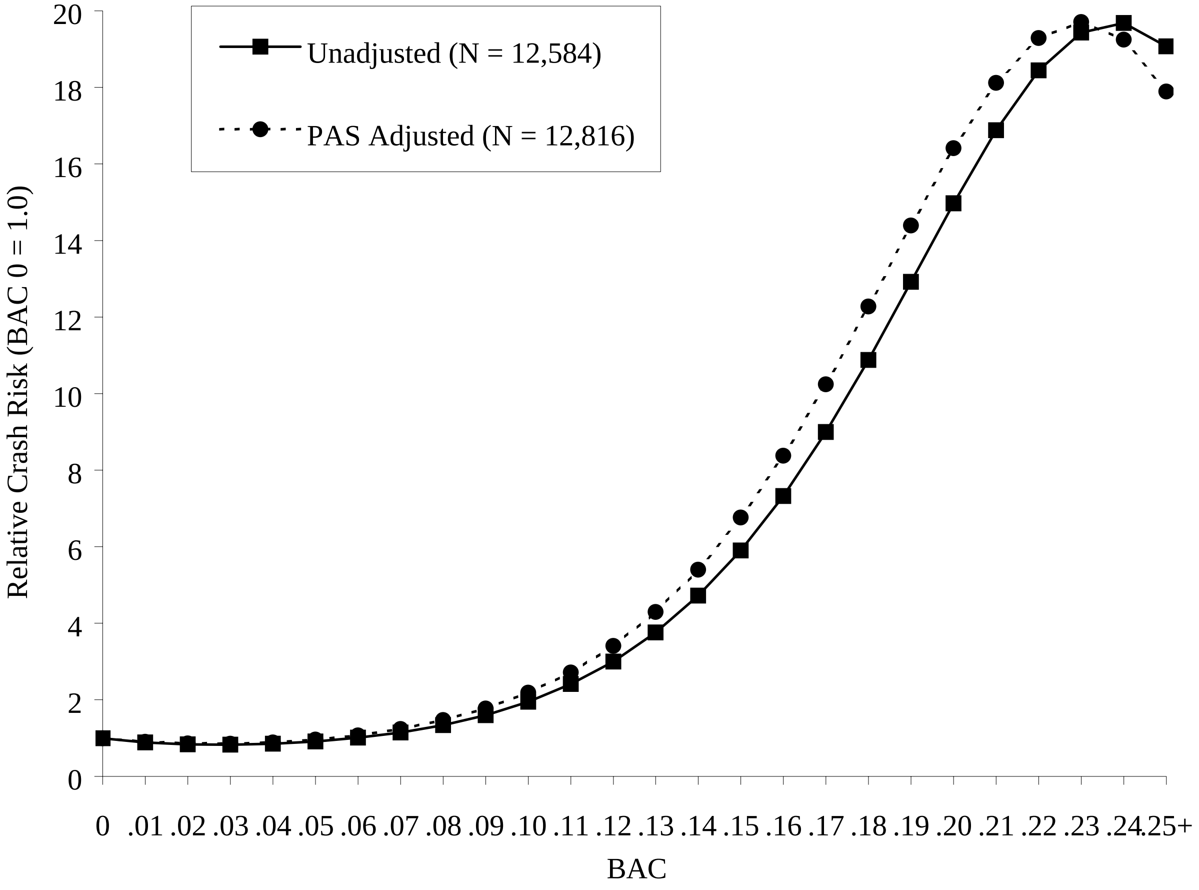

Figure 13. A comparison of PAS-adjusted relative risks with unadjusted relative risks

The above analyses substantiate the conclusion that the higher refusal rate of the crash group introduced a bias into the relative risk estimates. These adjustments, however, do not account for bias due to missing covariate data. Covariate data were missing for more crash drivers than control drivers, and crash drivers were more likely to have positive BACs. It is possible to estimate the net bias produced by the deletion of cases with missing data by comparing relative risk curves prior to and subsequent to the deletion. The relative risk values prior to deletion are based on the equation without covariates (Table 14). This comparison is shown in Table 32 and Figure 14.

Note that the relative risks are virtually identical through 0.09% BAC and then begin to diverge gradually. At 0.16% BAC, the differences are pronounced. At 0.24% BAC, the relative risk estimate is attenuated 32.3% by the deletions. These percent differences were computed for each BAC to serve as a missing data correction factor.

Table 32. Effects on Relative Risk of Deleting Subjects with Missing Values on Covariates (Hit-and-Runs Included)

BAC	Before deletion*	After deletion**
0.00	1.00	1.00
.01	.91	.91
.02	.86	.87
.03	.87	.87
.04	.91	.92
.05	.99	1.00
.06	1.12	1.13
.07	1.31	1.32
.08	1.57	1.57
.09	1.93	1.92
.10	2.41	2.37
.11	3.06	2.98
.12	3.92	3.77
.13	5.04	4.78
.14	6.49	6.05
.15	8.32	7.61
.16	10.59	9.48
.17	13.31	11.64
.18	16.45	14.00
.19	19.91	16.45
.20	23.49	18.78
.21	26.90	20.74
.22	29.76	22.07
.23	31.69	22.51
.24	32.31	21.92
.25+	31.42	20.29
N	11,994	11,373

* Includes 94 hit-and-run drivers.

** Includes 72 hit-and-run drivers.

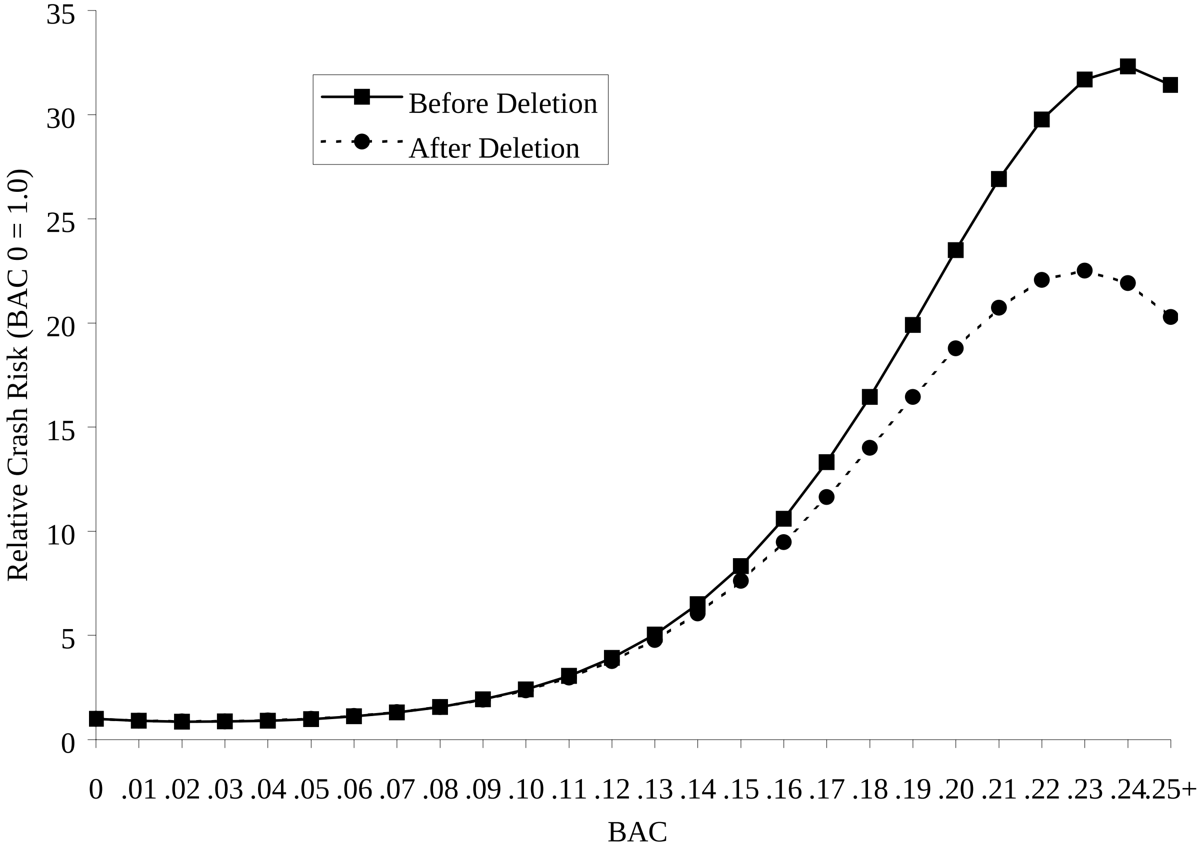

Figure 14. Effects on relative risks of deleting subjects with missing values on covariates

A final adjustment involved the bias introduced by hit-and-run drivers. A large percentage of the hit-and-run drivers who were apprehended and breath tested (15.6%) had high BACs (Table 15). This source of bias is largely independent of those introduced by refusal to participate and incomplete covariate data, and it is possible to apply a final weighting factor to adjust the relative risks for all three sources of bias. To adjust for the hit-and-run bias, the odds ratio differences were increased by a factor of 5.4 for the BACs of the 1:2 conditional model (Table 15). This correction adjusts for the fact that measured BACs are not available for 84.4% of the hit-and-run drivers (509 of 603).

The net effects of the series of adjustments summarized in Table 33 and Figure 15 can be illustrated with a specific example. Using the estimates from the non-reactive covariate equation (Table 28), the relative risk for 0.10% BAC prior to adjustment is 2.62. Adjustments are as follows:

- Hit-and-run: .17 x 5.4 = .92. 2.62 + .92 = 3.54
  This is an estimate of what the relative risk would have been at 0.10% BAC
  - if all hit-and-runs had been recovered, and
  - if the distribution of BACs for non-recovered and recovered hit-and-run drivers were the same.

- Imputation of BAC from the PAS data (Table 31)

$$\frac{2.18-1.95}{1.95}\ x\ 2.8 = .118\ x\ 2.8 = .3304 = 33.04\%$$

  - 3.54 x 1.3304 = 4.71
  - Multiplication of the relative risk adjusted for hit-and-run drivers produces an estimate of what the relative risk would have been if BACs had been obtained for all subjects including those for whom PAS scores were not available.

- Missing data adjustment (Table 32)

$$\frac{2.41}{2.37} = 1.02$$

  - 4.71 x 1.02 = 4.80
  - This multiplication of the relative risk adjusted for hit-and-run drivers and imputed BACs yields an estimate of relative risk if the analysis had not deleted subjects with valid BACs who had missing data on one or more of the covariates.
  - The estimate of 4.80 is essentially identical to the value of 4.79 shown for 0.10% BAC in the last column of Table 33.

The adjusted relative risks are dramatically different from the unadjusted risks, particularly at high BACs (Table 33, Figure 15). For example, at 0.25% BAC and higher, the relative risk is 153.68, compared to 26.60 before adjustment. In addition, the dip at low BACs disappears and is replaced by a very small elevation in risk beginning at 0.01% BAC. Note that the adjusted risks at 0.04% BAC and higher are significantly greater than 1.00 since a value of 1.00 is not contained within their lower and upper confidence bounds (see below).

The validity of the preceding adjustments is based on a number of assumptions, which are addressed in Section 5. Note, too, that there is a small, non-orthogonal (overlapping) component in the adjustments arising because the recovered hit-and-runs were more likely than others to have missing covariate data. The most likely effect of the non-orthogonality is a slight inflation of the bias correction factor.

4.6.7 Classification Accuracy of the Logistic Regression Models

Although the Cox-Snell R^2 values indicate that none of the logistic regression models accurately differentiate between control and crash drivers based on the covariate information, a more concrete illustration of predictive accuracy can be achieved by examining the classification matrix from a logistic regression equation. The logistic routine in SAS® produces 2 x 2 tables for various probability cut-points, but the PHREG procedure does not include this classification option. The SAS® logistic routine was used, therefore, to compute an unconditional logistic regression from the same variables that were significant at $p \leq .20$ in the all-covariates 1:2 conditional equation. The elimination of the matching requirement increased the available sample size from 11,373 to 12,418. Table 34 gives the results for one of the 2 x 2 classification tables.

Table 33. Relative Risks Adjusted for Bias Due to Refusal, Missing Data and Non-Recorded Hit-and-Runs

BAC	Before hit-and-runs adjusted	After hit-and-runs adjusted*	PAS imputation and missing data multiplier**	Final adjusted estimate
0.00	1.00	1.00	--	1.00
.01	.94	0.99	1.03	1.03
.02	.92	0.97	1.05	1.03
.03	.94	0.99	1.07	1.06
.04	1.00	1.05	1.12	1.18
.05	1.10	1.21	1.14	1.38
.06	1.25	1.41	1.16	1.63
.07	1.46	1.73	1.21	2.09
.08	1.74	2.12	1.27	2.69
.09	2.12	2.71	1.30	3.54
.10	2.62	3.54	1.35	4.79
.11	3.28	4.63	1.38	6.41
.12	4.14	6.19	1.44	8.90
.13	5.23	8.20	1.47	12.06
.14	6.60	10.87	1.51	16.36
.15	8.31	14.36	1.54	22.10
.16	10.35	18.77	1.57	29.48
.17	12.74	24.30	1.61	39.05
.18	15.43	30.87	1.65	50.99
.19	18.31	38.40	1.70	65.32
.20	21.20	46.53	1.76	81.79
.21	23.85	54.74	1.82	99.78
.22	25.99	62.12	1.90	117.72
.23	27.30	67.85	1.98	134.26
.24	27.55	70.91	2.07	146.90
.25+	26.60	70.61	2.18	153.68

* Odds ratio differentials in Table 14 multiplied by 5.4.

** Product of the PAS missing data correction multipliers.

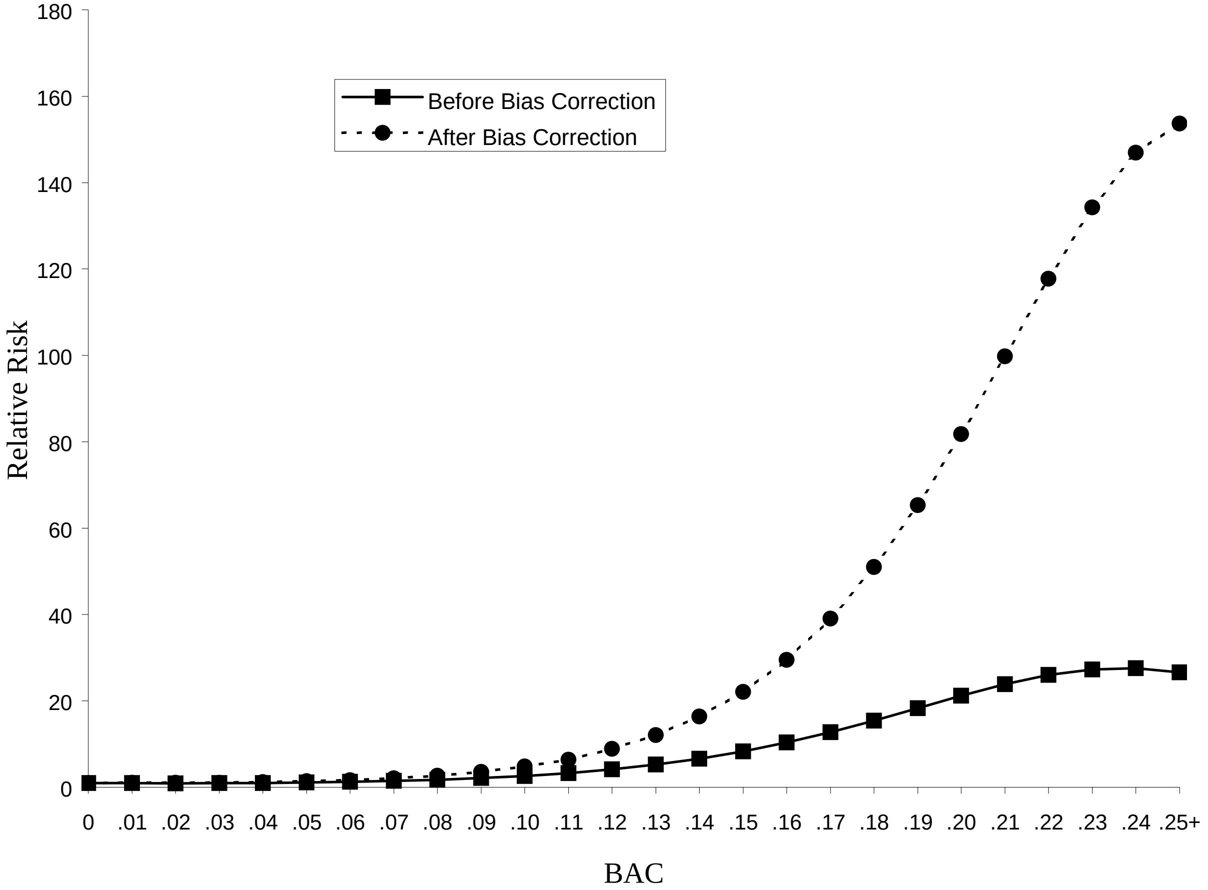

Figure 15. A Comparison of relative risk after adjustment for subject deletion bias

Table 34. Classification Matrix for Unconditional Logistic Regression Equation, All Covariates p ≤ .20
***N* = 12,418**

	Actual			
Predicted	**Control**	**Crash**	**Total**	**Percentage correct**
Control	6589	1994	8583	76.7%
Crash	2038	1797	3835	46.8%
Total	8627	3791	12418	--
Percentage correct	76.4%	47.4%		

phi = .24

The probability cut-point for formulating Table 34 was selected to equalize approximately the marginal distribution. That is, since 30.5% of the sample was crash involved, a logistic probability cut-point was used that predicted (approximately) 30.5% of the sample to be crash involved. This strategy tends to be optimal in the sense that the phi coefficient can only attain a value of 1.0 if the marginals are equal. Another consequence is that the number of false negative and false positive predictions will be the same when the marginal distributions are equal. Hence, the weight or disutility given to the two types of prediction error are, by implication, the same.

The selected probability cut-point came very close to producing equal marginals. In terms of classification accuracy, 76.4% of the control drivers were correctly classified compared to 47.4% of the crash drivers. Thus, the specificity of the classifications was much better than the sensitivity. The phi coefficient for the classification was only .24, which means that only 5.8% ($.24^2$) of the variance in individual crash risk can be explained by the classification function. This correlation is not much different from the various Cox-Snell R^2s reported in preceding sections.

A change in the cut-point alters the tradeoff between sensitivity and specificity, but the two are inversely related. If the probability threshold were lowered to .20, the sensitivity of the model would increase to 87%, but its specificity would decrease to 29.2%. Almost 90% of the crash-involved drivers would be correctly identified but at the price of erroneously identifying 70% of the control drivers as crash involved.

The classification matrix output of the SAS® Logistic procedure provides a number of association measures on the relationship between the predicted probability scores from the logistic regression equation and observed status (control vs. crash). One such measure, the C statistic, represents the area under the complete receiver operating characteristic (ROC) curve. The value of C can range from .5 to 1.0. A value of .5 indicates that the accuracy of classifying each subject does not exceed chance. A value of 1.0 indicates that the model always assigns higher probabilities to crashes than to controls. For these data, C= .685, indicating that the crash driver was assigned the higher logistic probability score in 68.5% of all possible crash-control pairings.

This low degree of predictive accuracy initially appears inconsistent with the steep relative risk function for the BAC parameter. Traffic crashes, however, are subject to large stochastic influences that limit their correlation with individual differences among drivers (Peck, McBride and Coppin, 1971; Gebers, 1998; Cobb, 1940). Although alcohol greatly increases crash risk, the BAC of 81.7% of the crash-involved drivers was 0.00%. The increase in relative risk conditional on a given BAC did not exceed .50 for most analyses until about 0.08% BAC. Furthermore, only 12% of the control drivers' positive BACs exceeded 0.10%. The very small proportion of subjects with high BACs suppresses the magnitude of the correlation coefficient even when relative risks are extreme. It is for this reason that most epidemiologists prefer relative risks and odds ratios to correlation coefficients as measures of disease propensity. Also, the correlation between logistic probabilities and crash risk for the data summarized in Table 34 has been substantially attenuated by the biases from non-participation and hit-and-run attrition. The classification function could not be adjusted for these biases as it was for the relative risks shown in Table 33.

4.6.8 Relative Risk Confidence Intervals

The SAS® PHREG conditional logistic regression procedure does not provide confidence intervals for the relative risk estimates. It was possible, however, to output the

PHREG variance-covariance matrix of the non-reactive covariate equation and to insert the appropriate matrix elements into an Excel® program. That program produced confidence intervals for the logit of each positive BAC as estimated from the regression coefficients for the BAC parameters (linear, quadratic and cubic). The lower and upper bounds for the relative risks were computed by exponentiating the logit differences:

$$\exp[B \pm 1.96\ x\ \mathrm{var}(d)]$$

where B represents the logit difference from the 0.0 BAC logit and var(d) represents the variance for each linear combination of the BAC coefficients.

The formula for computing d can be found in Hosmer and Lemeshow (2000, pp. 18-20 and 40-41).

Table 35 tabulates and Figure 16 graphs the lower and upper confidence bounds. The confidence intervals are narrow at low BACs but become very large at high BACs. At 0.25%+ BAC, the 95% confidence interval for the relative risk estimate is 10.25 - 69.00. These confidence intervals are based on the logistic regression equation and relative risk curve prior to the bias adjustments used to produce the final estimates shown in the second column of Table 33.

It is not possible to produce a mathematically rigorous set of confidence intervals for the final adjusted estimates because of the unknown statistical properties (accuracy and precision) of the adjustment weights. Confidence interval estimates for the final adjusted estimates can be calculated, however, if the adjustment weights are considered as fixed error-free constants. In this instance, the computed intervals shown in Table 35 are applied to the adjusted relative risk point estimates in Table 33. Table 36 shows results of this superimposition for selected BACs. The lower bound for all relative risk estimates is above 1.00 for all positive values above 0.03% BAC. By 0.15% BAC, there is a dramatic increase in the lower and upper bound risk relativities and at 0.25%+ BAC, the 95% confidence interval ranges from 124.28 to 183.10. Because of the unknown validity of the bias adjustments and the non-testable assumptions implicit in the derivation and application of the adjustment weights, these estimates should be used with caution.

4.6.9 Relative Risks by Site

Site x BAC interactions from the preliminary logistic regression model are not statistically significant. Based on that result, relative risk estimates were computed for the combined sites. Even under ideal circumstances, however, the statistical power is much lower for detecting interactions than for main effects (Breslow, 1980; McClelland and Judd, 1993). This limitation is exacerbated by the large standard errors implied by the width of the confidence intervals (Table 35). Thus, it was deemed appropriate to examine the relative risk function for each site separately.

Relative risk functions were generated for Long Beach and Fort Lauderdale by using the variables that were significant at $p \leq .20$ in the non-reactive covariate equation. New logistic regression models were generated for each site without adjustments for hit-and-runs, missing data, or refusals. The BAC parameters for the separate equations appear in Table 37.

Table 35. 95 Percent Confidence Intervals for Relative Risk Estimates from the Equation Containing BAC and All Non-Reactive Covariates

N = 11,373

BAC	Lower bound	Upper Bound
0.00	1.00	1.00
.01	.927	.943
.02	.864	.972
.03	.854	1.03
.04	.888	1.12
.05	.965	1.25
.06	1.09	1.43
.07	1.26	1.69
.08	1.49	2.04
.09	1.78	2.53
.10	2.15	3.21
.11	2.61	4.13
.12	3.19	5.36
.13	3.91	6.98
.14	4.80	9.08
.15	5.89	11.70
.16	7.19	14.89
.17	8.71	18.65
.18	10.39	22.92
.19	12.13	27.64
.20	13.71	32.78
.21	14.81	38.43
.22	15.10	44.74
.23	14.36	51.89
.24	12.65	64.39
.25+	10.25	69.00

*Indexed to crash odds of the 0.0 BAC group.

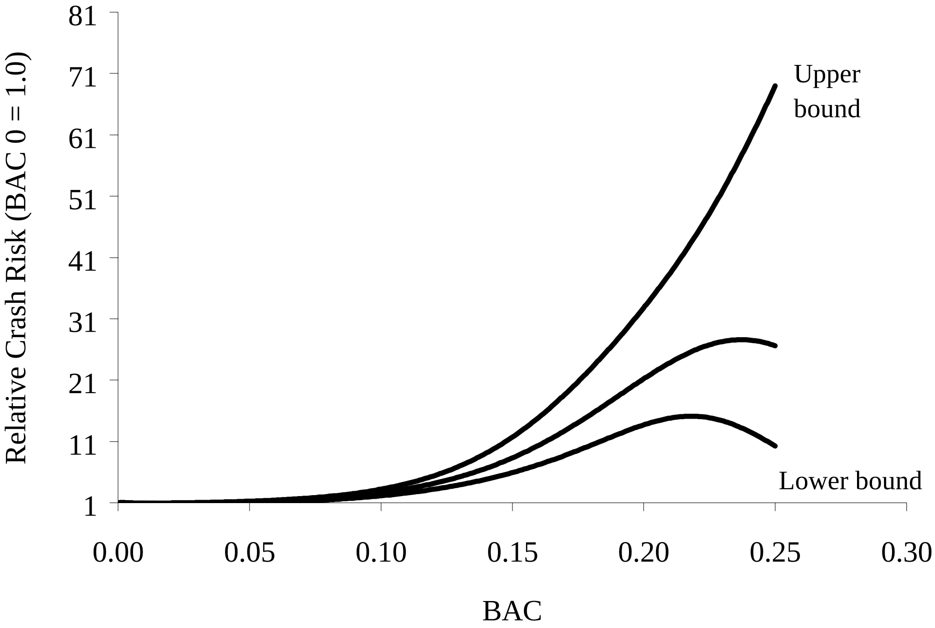

Figure 16. 95 percent confidence interval for odds ratio relative to 0.0 BAC for model including BAC and non-reactive demographic covariates

Table 36. 95 Percent Confidence Intervals of Relative Risk Estimates for Selected BAC Values Adjusted for Test Refusal and Other Subject Attrition Biases

BAC	Lower bound	Upper bound
0.00	1.00	1.00
.02	.980	1.08
.03	.970	1.15
.04	1.06	1.30
.06	1.46	1.80
.08	2.42	2.96
.10	4.26	5.32
.15	19.20	25.00
.20	72.24	91.34
.25+	124.28	183.10

Table 37. BAC Parameter Estimates and Cox-Snell R^2 for Separate Long Beach and Fort Lauderdale Equations

Parameter	Long Beach (*N* = 5,667)	Fort Lauderdale (*N* = 5,706)
BAC	-7.98 ($p = .15$)	-3.73 ($p = .34$)
BAC^2	277.3 ($p = .002$)	229.3 ($p = .001$)
BAC^3	-727.9 ($p = .02$)	-693.0 ($p < .001$)
Cox-Snell R^2	.073	.054

Both equations show a significant cubic structure, but the linear component is much larger for Long Beach. That equation also has more explanatory power than the Fort Lauderdale equation (R^2.073 vs. R^2.054). Figure 17 plots the relative risk curves, and Table 38 tabulates the values for selected BACs. Although the values for Fort Lauderdale are somewhat larger than for Long Beach until 0.12% BAC, the curves are quite similar until 0.15% BAC. At 0.15% BAC the relative risk value for Long Beach is noticeably larger than for Fort Lauderdale, and there is an extreme divergence at 0.20% BAC (Fort Lauderdale 16.08; Long Beach 31.14). The difference is due to more crash drivers and fewer control drivers at very high BACs in Long Beach.

Although the differences between the two curves cannot be explained by postulating one or two unusual events, the sensitivity of polynomial regression to small changes in the number of data points (drivers) at extreme values, as well as to the size of the confidence intervals at high BACs, must be noted. The larger number of hit-and-run drivers in Long Beach provides an alternate explanation of site differences. Seventy-four of 94 recovered hit-and-run drivers with valid BACs were from Long Beach. Given the very high BACs of recovered hit-and-runs, their under-representation in Fort Lauderdale attenuates the relative risk function at high BACs at that site (Table 38, Figure 17).

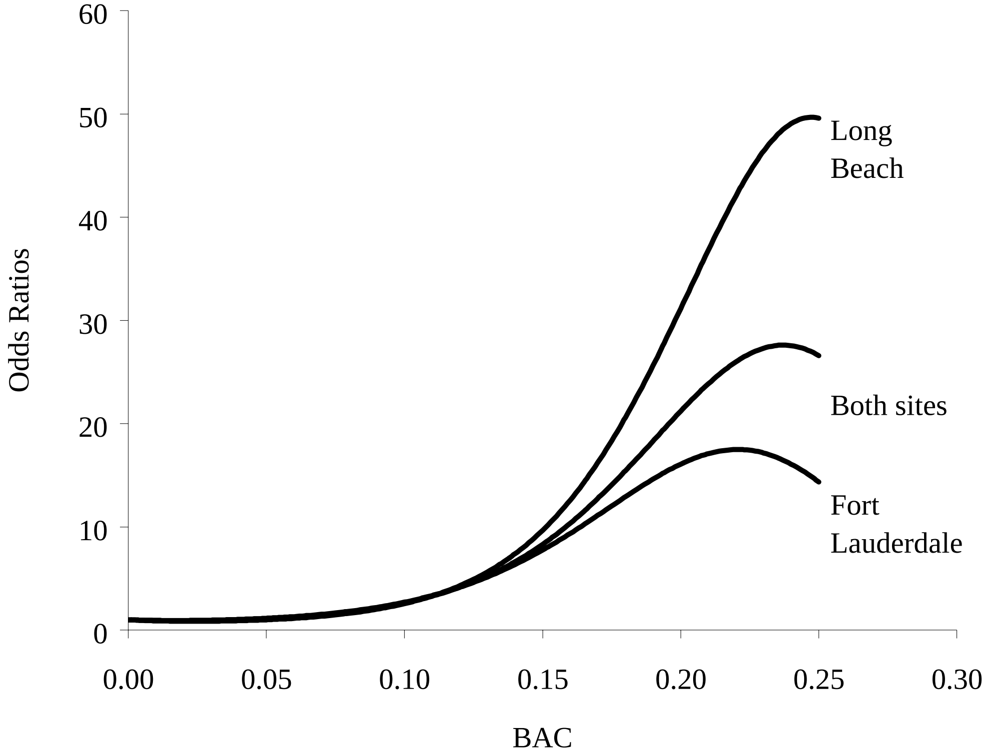

Figure 17. Relative crash risk (odds) by BAC level for each site (BAC 0 = 1.0)

Table 38. Relative Crash Risk for Selected BAC Values from Separate Equations for Long Beach and Fort Lauderdale

BAC	Long Beach	Fort Lauderdale
0.00	1.00	1.00
.02	.86	.94
.04	.92	1.04
.06	1.00	1.31
.08	1.63	1.82
.10	2.57	2.71
.15	9.67	7.76
.20	31.14	16.08
.25+	49.58	14.34

4.6.10 Grand Rapids Dip and Simpson's Paradox

Section 2 and Appendix A of this report review several studies, including the landmark Grand Rapids Study (Borkenstein et al., 1964) and report on re-analyses of the Grand Rapids data by Allsop (1966) and Hurst et al. (1994). The latter papers focused much attention on the apparent reduction in crash risk at low BACs (0.01% – 0.04%) and attributed this dip to a statistical phenomenon known as Simpson's Paradox. In the most extreme form of the phenomenon, marginal distributions (means, rates, odds ratio) of a covariate are opposite in trend to those of its sub-groupings. Allsop (1966) noted that several age groups showed an increasing crash risk at all BACs, but the marginal BAC distribution showed a reduction in crash risk at low BACs. Hurst et al. (1994) reported a more dramatic example for subgroups defined by drinking frequency. The multiple logistic regressions in the present study revealed no evidence of Simpson's Paradox, although the groupings and methodology are not directly comparable to those used by Allsop (1966) and Hurst et al. (1994).

Before presenting relative risk tabulations by age and drinking frequency comparable to those in the Allsop (1966) and Hurst et al. (1994) studies, it is instructive to examine the relationship among crash probability, age and BAC as estimated from the logistic regression model containing all non-reactive covariates. These estimates are shown graphically in Figure 18.

drivers aged 55+. The substantially higher risk ratio (.742) for the latter group suggests that alcohol is a less important factor in older drivers' crashes.

Table 39. Relative Risk (Odds) of 0.00% BAC Drivers' Crash Involvement by Age Group
***N* = 13,909**

Age	Relative Risk*
Under 21	.540
21-24	.531
25-34	.583
35-44	.594
45-54	.593
55+	.742
All ages	.564

$$* RR = \frac{Odds\ of\ 0-BAC\ being\ crash\ involved}{Odds\ of\ positive\ being\ crash\ involved}$$

Relative risks as a function of BAC within age categories are shown in Table 40 and Figure 19. The BAC distributions have been capped at .12+% BAC due to sample size limitations and to facilitate comparisons with Allsop (1966).[15] Contrary to the findings of Allsop (1966), the current analysis finds the dip at low BACs in all age groups. Only at 0.06% - 0.079% BACs are the relative risks above 1.0 for all age groups. The general shapes of the relative risk gradients are remarkably similar across age when one takes into account the instability introduced by small *N*s in some cells. The very large increase in risk for drivers under age 21 at $\geq$ 0.12% BAC is an exception.

The study data provide no evidence of Simpson's Paradox involving age. The most dramatic evidence of a Simpson's Paradox explanation of the Grand Rapids dip, however, comes from drinking frequency (Hurst et al., 1994). Using categories that closely approximate those in Hurst et al. (1994), crash probability as a function of BAC and number of drinking days per month were calculated. Table 41 summarizes the crash probabilities, and relative risks indexed to the crash odds of the 0.0 BAC groups are shown in Table 42. Dips at low BAC levels are evident, which is contrary to the findings from the Hurst et al. (1994) analysis of the Grand Rapids data.

The relative risk curves within number of drinking days (Table 42) are plotted graphically in Figure 20. A similar graphical display using number of drinks per setting as the subgroup dimension is also included (Figure 21). This covariate was a somewhat stronger predictor of crash risk and BAC level than was number of drinking days in the logistic regression analyses. It also shows no evidence of a Simpson's paradox artifact.

---

[15] The primary motivation in generating Table 40 is to determine whether the previously noted and widely discussed dip in risk at low BACs vanishes when the data are examined within age group.

Table 40. Relative Crash Risk as a Function of BAC Within Age Group
N = 13,760

BAC	Under 21	21-24	25-34	35-44	45-54	55+	N
0.00	1.00	1.00	1.00	1.00	1.00	1.00	11,978
.01 - .019	1.89*	1.09	1.07	.44	.72	.42	394
.02 - .039	.59	1.17	.68	.64	.79	.95	362
.04 - .059	1.30	.66	.81	1.12	.87	.61	253
.06 - .079	1.04	1.02	1.75	1.54	1.06	1.57	201
.08 - .099	(2.02)**	2.31	1.99	5.10	4.75	4.91	101
.10 - .119	(2.02)**	4.62	4.19	1.13	1.28	.73	83
12+	21.53	8.96	8.11	8.25	10.01	9.62	388
N	1,216	1,476	4,140	3,368	1,998	1,562	13,760

*Relative risk values are indexed to the crash odds of the 0-BAC levels within age.

**Based on the collapsed interval, .08 - .119, due to 0-entries in one or more of the cells of the non-collapsed intervals.

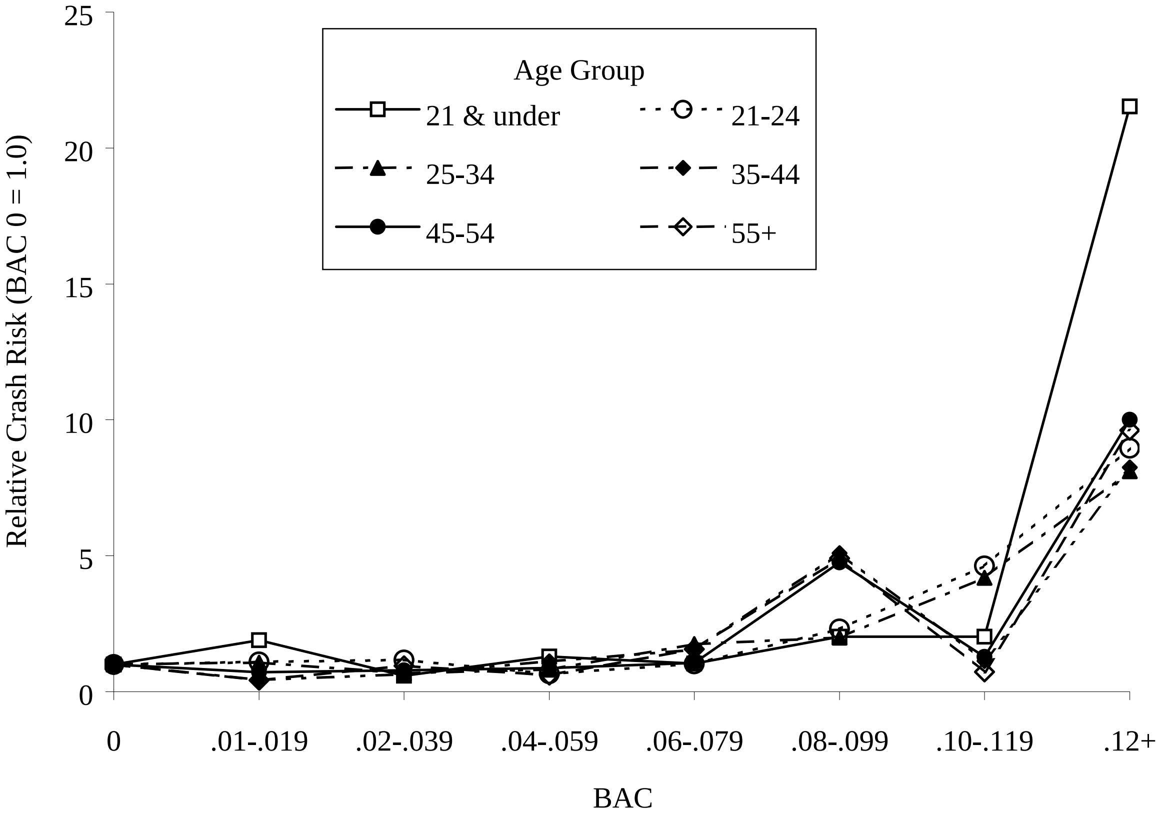

Figure 19. Relative crash risk as a function of BAC within age group

Table 41. Crash Probability as a Joint Function of BAC Level and Number of Drinking Days

BAC	Number of drinking days				
	0	1-3	4-11	12-27	28
0.00	.308*	.269	.215	.181	.224
.001 - .009	.241	.118	.326	.143	.250
.01 - .049	.250	.310	.209	.166	.102
.05 - .079	.222	.415	.292	.292	.257
.08 - .109	.667	.500	.519	.313	.370
.11+	.750	.817	.688	.701	.682
All BAC levels	.309	.293	.251	.227	.284

*Entry is proportion involved in crashes within each BAC x drinking days category.

Table 42. Relationship Between BAC and Relative Risk Within Number of Drinking Days Per Month (Past 28 Days)

BAC	Number of drinking days					
	0	1-3	4-11	12-27	28	N
0.00	1.00	1.00	1.00	1.00	1.00	11,984
.001 - .009	.72*	.37	1.77	.76	1.16	142
.01 - .049	.75	1.22	.97	.90	.40	751
.05 - .079	.65	1.93	1.51	1.87	1.20	320
.08 - .109	4.31	2.70	3.93	2.06	2.04	136
.11+	6.76	12.17	8.07	10.67	7.46	419
N	6,515	2,998	2,698	985	556	13,752

*Entries are relative risk (odds ratio) indexed to the crash odds of the 0-BAC group.

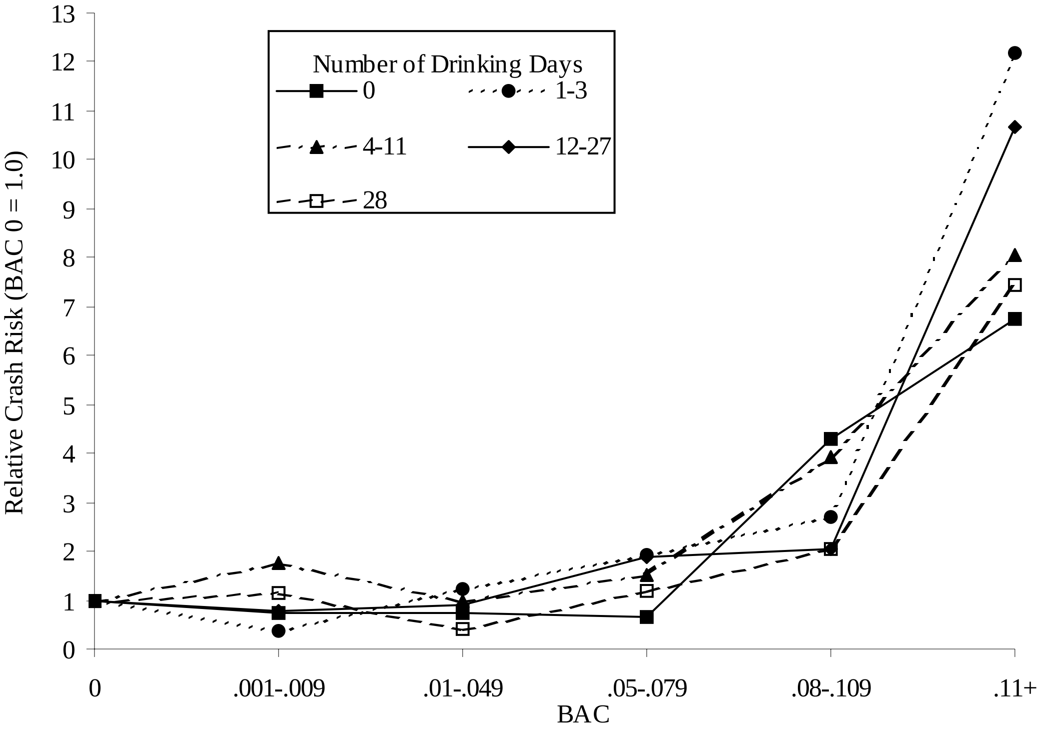

Figure 20. **Relationship between BAC level and relative risk within number of drinking days per month**

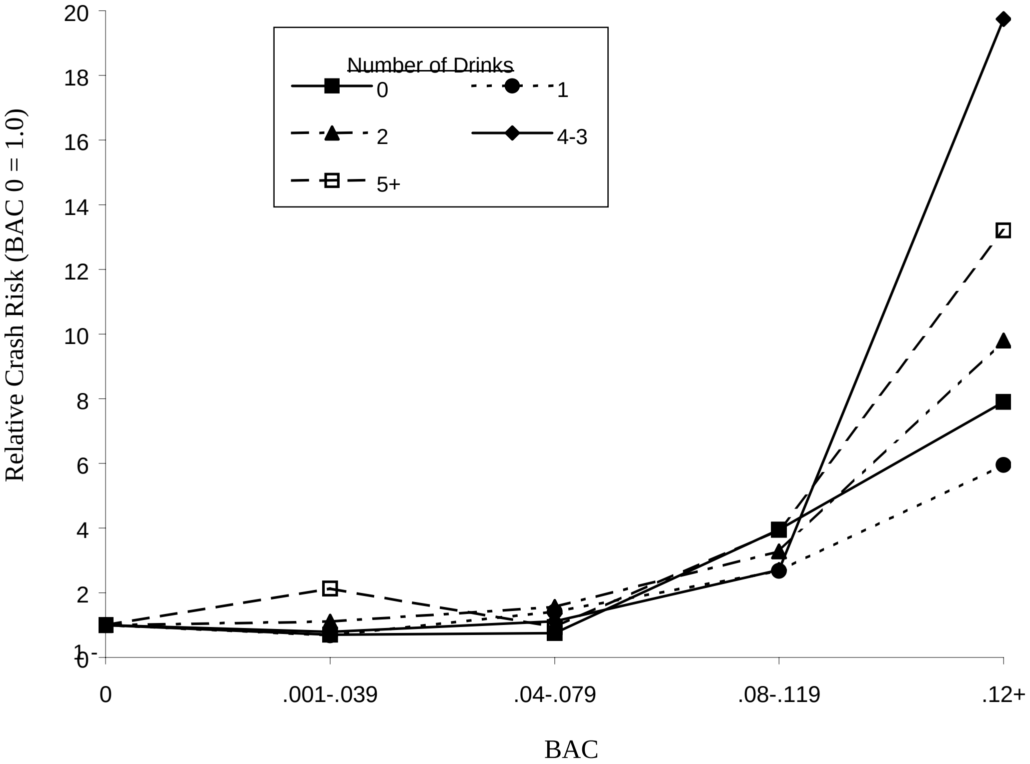

Figure 21. **Relative crash risk as a function of BAC level within number of drinks per setting**

0.12%, but the risk curves were nevertheless adjusted with the predicted BAC, and logistic regressions were re-computed.

A further adjustment for non-participation was necessary, because PAS estimates were not available for all non-participants. The adjustment used a simple weighting procedure, based on the assumption that the distribution of the 232 available PAS scores could be extrapolated to 424 (64.4%) non-participants for whom the data were not available. This is a reasonable assumption since the reasons PAS readings could not be obtained were random (environmental conditions; inability to get the instrument close to the driver).

The higher proportion of crash subjects than controls who provided a breath specimen but did not complete the interview introduced the third source of non-participation bias. This bias does not affect the BAC risk curve without covariates and has only a negligible effect on the BAC risk curve adjusted for age and gender because age and gender were available on almost 100% of the participants. Since drivers who failed to complete the survey were more likely to be in the crash group and to have higher BAC values, their deletion attenuated the BAC-crash risk curves that were adjusted for multiple covariates. The effect of this bias was removed by a weighting procedure based on the change in relative risks when subjects with missing survey covariate data were deleted from the equation containing no covariates. This procedure adjusts for the fact that subjects with missing data were more likely to have positive and higher BAC values and were more likely to be crash-involved, but it does not take into account the possibility that the score pattern on the covariates (e.g., drinking pattern, education, etc.) for the missing data subjects might be different than that of complete data subjects.

It is also possible that the counterintuitive direction of the relationship between some covariates and crash risk is attributable to the bias stemming from the non-participants and non-recovered hit-and-runs. For example, a significantly *decreased* crash risk was found to be associated with heavy and frequent drinking, lack of sleep, and Hispanics. The adjustment of the BAC relative risk curve for bias did not involve imputation of missing covariate data. The selective inaccuracy of self-report may also have contributed to these unexpected findings.

Despite the above qualifications, the consistent direction of the biases and the rationale underlying the adjustments strongly supports the conclusion that the adjusted relative risks shown in Table 33 provide a much more accurate portrayal of how BAC affects crash risk, particularly at high BAC levels than do the unadjusted versions.

5.2 Curve Fitting

Polynomial power terms were used to represent the non-linear form of the logits, and a cubic model provided an adequate fit to almost every equation. It should be noted, however, that limitations apply to the polynomial regression models.

First and foremost, the polynomial parameter estimates (regression coefficients) can be highly sensitive to the range and frequency distribution of the exposure variable (BAC) and to specific points where “bends” in the relative risk function occur (local maxima and minima). Because BACs higher than 0.25% were sparse, the analyses were capped at that level to reduce instability. As a consequence, the calculation of crash risk at a higher BAC requires that the BAC be converted to 0.25% to avoid extrapolating beyond the range of the modeled data. Other than the fact that the fitted models cannot differentiate crash risk for 25% BAC vs. 0.35% BAC, for example, the limitation has little practical importance. The proportion of drivers at such extremely high BACs is small, and differences in relative risk beyond 0.20% BAC have no policy implications.

Another limitation of the use of polynomial (power) terms to represent non-linearity relates to multicollinearity. The linear, quadratic and cubic power terms are very highly correlated, which increases the standard errors of the respective regression coefficients and can affect numerical accuracy in the parameter estimates. The data were centered by subtracting the grand mean BAC from each individual BAC before fitting the models. This standard procedure eliminated the correlation between the linear and quadratic terms, but a substantial correlation between the linear and cubic term remained.

A similar problem occurs with the introduction of interaction terms, which are highly correlated with main effect terms. Although the resultant models are unbiased, the multicollinearity among the components increases the standard errors of the parameter estimates when all terms are included in the model.

5.3 Confidence Intervals

The confidence intervals calculated for the final relative risk curve are very narrow for 0.00% and low BACs and very large for high BACs. They reflect the standard errors of the regression coefficients for the three BAC polynomial terms considered as a set in the context of the final covariate-adjusted equation. These confidence intervals were calculated prior to the bias adjustments and cannot take into account variance or bias associated with the adjustments. They do not reflect all of the uncertainty and variability associated with the relative risk curve. Nonetheless, they provide compelling evidence, even at a 0.04% BAC, of a strong dose-response relationship between alcohol (BAC) and crash risk.

5.4 Covariate Selection

A large number of logistic regressions were calculated with different subsets of covariates. A major emphasis was the conditional logistic regression with non-reactive demographic covariates (age, gender, education, ethnicity, etc.). Equations also included alcohol consumption and sleep patterns as covariates. These adjust the BAC-crash risk for variations in drinking habits and sleep patterns and merit some caution in interpretation. A proper adjustment for a covariate requires that it be relatively free of measurement error, especially systematic or non-random error. For alcohol-consumption measurement errors to be random, they would have to be uncorrelated with a driver’s true drinking frequency and group membership (crash or control). If the magnitude of error (under- or over-estimate) is influenced by how much a person actually drinks or by the occurrence of a crash, the covariate adjustment will be biased. If the non-random error is sufficiently large, it is possible for the covariate-adjusted relative risks to be more biased than it would have been without adjustment.

It is unlikely that measurement error for some of the covariates is random or unaffected by group membership. Problem drinkers tend to underestimate their alcohol consumption, and it appears that the tendency is influenced by crash involvement.

It is conceivable, too, that persons in a crash were more likely to deny sleep deprivation or traveling from a bar because of their perception that these factors are considered by many to be undesirable from a public safety perspective and might have contributed to the crash.

Multiple logistic regression and analysis of covariance are not well suited for modeling certain complex causal relationships among independent variables and a dependent variable. For an association to be viewed as completely causal, it is necessary that the association between the independent and dependent variables be adjusted for any concomitant association with exogenous variables. Care is necessary, however, when covariates are functionally related to both the independent and dependent variable, and this raises a question about valid covariates. For example, quantity and frequency of alcohol consumption has an obvious connection with BAC, and it is likely that the extent of impairment and crash propensity is causally influenced by how much and how often a driver drinks. In that case, an adjustment of the BAC-crash risk relationship for quantity and frequency of alcohol consumption would remove a causal component or mediator of the relationship.

Under what circumstances might adjustment for drinking practices result in a more valid assessment of the BAC-crash risk relationship? The Grand Rapids Study found that non-drinkers had substantially higher crash risks than all other drinking categories (Borkenstein et al., 1964, 1974). A similar but less pronounced relationship appears in the present study’s data. If it were determined that non-drinkers have traits not related to alcohol use that predispose them to higher crash rates, that would be an appropriate bias to control. An alternate explanation, however, is related to measurement error. If crash-involved drinkers are more likely to claim to be non-drinkers, this would produce the observed relationship, and it would be inappropriate to adjust for a bias that is an artifact of differential measurement error. For this reason, the logistic regression results that exclude alcohol consumption and sleep covariates are emphasized here. These issues actually are of little practical concern, because the BAC risk curves prior to non-participant-bias adjustment were quite similar across all modes of analyses and covariate sets.

It was hypothesized that teenage drivers would exhibit elevated crash risks at lower BACs than older drivers, but the BAC x age interaction, though suggestive, was not statistically significant. Instead, the data suggest that drivers under age 20 years are more affected than older drivers at high BACs (0.11%+). Tests of interaction have much lower power than tests of main effects, and the sample size was not adequate to isolate a 16-17 year old age stratum.

The BAC – crash risk relationship did not differ significantly between Long Beach and Fort Lauderdale, and the data were collapsed across sites for most of the logistic regressions. An inspection of the within-site risk curves, however, revealed a much steeper risk gradient for the California site at BACs of 0.15% and higher. If BAC actually was more strongly associated with crash risk in Long Beach than in Fort Lauderdale, that presumes a BAC x site interaction. Given the very large standard errors of the crash risk estimates for high BACs, it is possible that failure to detect a BAC x site interaction was a Type II error.

Another possibility is that the difference in the risk curves may be explained by the greater success in Long Beach in recovering and testing hit-and-run cases. Of the 94 recovered cases, only 20 were from Fort Lauderdale. This between-site difference in the proportion of recovered hit-and-run cases alters the steepness of the two curves in the direction and manner observed.

5.5 Comparisons with the Grand Rapids Study

The “Grand Rapids Dip”, which has been a controversial finding in the impaired driving literature, sometimes has been interpreted as evidence that small amounts of alcohol decrease the risk of causing a crash. Note, however, that the analysis of the Grand Rapids data was limited to univariate comparisons of crash and control groups. When the data were subjected to bivariate analysis, the dip disappeared (Allsop, 1966). Allsop (1966) and Hurst, Harte and Frith (1994) attributed it to a statistical phenomenon known as Simpson’s Paradox.

The univariate marginal risk distributions from the present study follow a pattern that is similar to the Grand Rapid Study result and display an apparent dip in the relative crash risk at low BACs. Again, however, additional analyses (the bias corrections presented in the previous section) eliminated that seeming decrease in risk at low BACs. More specifically, the small reduction in crash probability at BAC’s below 0.05% disappears, and a statistically significant increase in crash risk begins at 0.04% BAC. There are small but not statistically significant increases in relative risk at 0.01% – 0.03% BACs, and the direction of changes are consistent with a monotonically increasing crash risk beginning at BACs close to zero.

Although the present study supports the conclusion that the Grand Rapids Dip is largely artifactual, it provides no evidence of Simpson’s Paradox in the BAC – crash risk curves even prior to the adjustments for bias. The most likely explanation for absence of the paradox in the present study is its use of a matched case-control design and multiple covariates, which eliminated the confounding that was the likely cause of the dip in the Grand Rapids data.

Table 43 compares the results for three of the models produced by the present study with relative risk estimates reported by Allsop (1966) for the Grand Rapids data. The table illustrates the modest difference in risk produced by adding the covariates and the dramatic effects of adjusting for bias. Note the similarity between this study’s unadjusted relative risks and those obtained in the Grand Rapids Study.

Figure 22 plots the final adjusted estimate from Table 43. This bias-adjusted curve shows a significant increase in relative risk at about 0.04% BAC with an exponential acceleration of risk at 0.10% BAC and above. Note, for example, that drivers at a 0.15% BAC are 8.2 times more likely to be crash involved than drivers at a 0.08% BAC (22.10 ÷ 2.69 = 8.2). At 0.25% BAC, relative risk compared to 0.08% BAC increases by a factor of 57.1(153.68 ÷ 2.69 = 57.1).

5.6 Future Data Analysis and Research

The data obtained in this study can support additional examinations. Furthermore, many questions about alcohol involvement in motor vehicle crashes await further research.

Table 43. Comparison of Calculated Relative Risks with Grand Rapids Results

BAC	No covariates*	Non-reactive demographic covariates **	Final adjusted estimate***	Grand Rapids****
0.00	1.00	1.00	1.00	1.00
.01	.91	.94	1.03	.92
.02	.87	.92	1.03	.96
.03	.87	.94	1.06	.80
.04	.92	1.00	1.18	1.08
.05	1.00	1.10	1.38	1.21
.06	1.13	1.25	1.63	1.41
.07	1.32	1.46	2.09	1.52
.08	1.57	1.74	2.69	1.88
.09	1.92	2.12	3.54	1.95
.10	2.37	2.62	4.79	5.93
.11	2.98	3.28	6.41	
.12	3.77	4.14	8.90	4.94
.13	4.78	5.23	12.06	
.14	6.05	6.60	16.36	10.44
.15	7.61	8.31	22.10	
.16	9.48	10.35	29.48	21.38
.17	11.64	12.74	39.05	
.18	14.00	15.43	50.99	
.19	16.45	18.31	65.32	
.20	18.78	21.20	81.79	
.21	20.74	23.85	99.78	
.22	22.07	25.99	117.72	
.23	22.51	27.30	134.26	
.24	21.92	27.55	146.90	
.25+	20.29	26.60	153.68	

*From Table 28, column 2.
**From Table 28, column 3 or Table 33, column 2.
***From Table 33, column 5.
****From Grand Rapids Study data reported in Table 25 (a) of Allsop (1966).

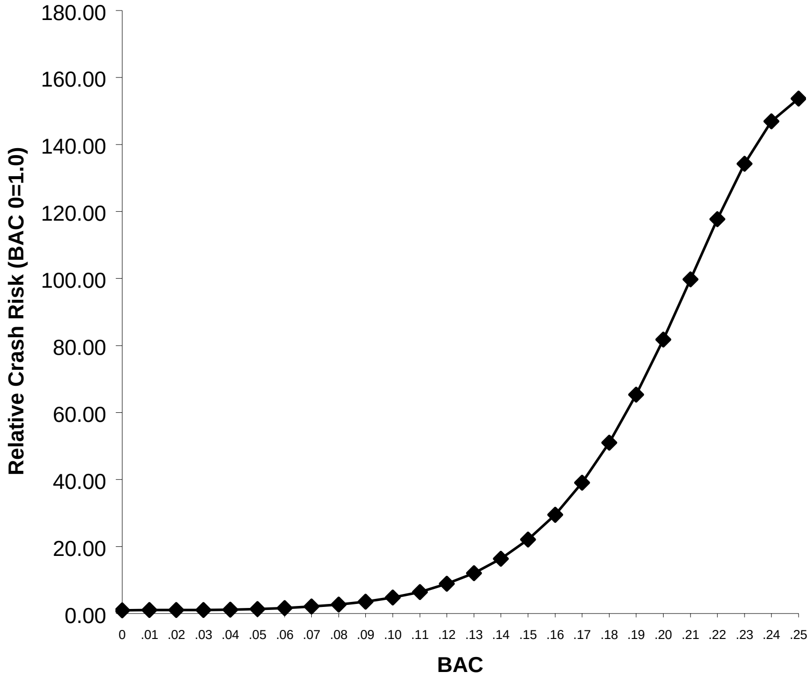

Figure 22. Final relative risk estimate

5.6.1 Additional Analyses with Respect to the Relative Risk of Alcohol

Exploration of alternative curve-fitting methods. Improved fit of response surfaces can sometimes be obtained through use of piece-wise regression models known as splines and polynomial splines (Darlington, 1990; Stone and Koo, 1985). These techniques have both advantages and disadvantages over polynomial regression but often result in slightly improved fits on sampled data.

The use of additive as opposed to multiplicative models is another alternative that might be explored. Logistic regression transforms the crash probabilities from the original additive scale to logarithms of relative risk (logits). It creates a model in which the BAC and covariate effects on crash risk are expressed as additive sums of multiplicative effects. Although the logistic transform is appropriate and is the method of choice for binary data, models based on the original additive scale offer advantages in some cases (Breslow and Storer, 1985; Coughlin et al., 1991).

Each of the relative risk curves produced by the cubic polynomial models showed a downward trend in crash risk above 0.25% BAC. In most cases, the maximum crash risk was observed at 0.23 – 0.24% BAC. It is conceivable that very high BAC drivers might be less of a crash risk, perhaps because their impairment is apparent to other drivers. The cubic structure of the obtained models, however, is based on a very small number of control observations with extremely high BACs. A more parsimonious (quadratic) model with significantly smaller standard errors might result if the distribution were truncated at 0.20% or 0.22%.

Application of more sophisticated missing data imputation procedures. Values for missing data can be imputed with multiple imputation procedures, and the E-M algorithm can be used to supply maximum likelihood estimates for the missing data elements based on the pattern of covariate scores in the complete data vector (Rubin, 1987; Rao, 1996). Analysis of the study data applied simple weighting procedures to the covariate-adjusted marginal BAC-crash-risk association. This approach avoided the necessity of supplying estimates for each missing covariate value but at the price of uncertainty as to the accuracy and precision (standard errors) of the adjustments. In view of the diverse nature of the missing data and the small number (*N*=72) of hit-and-run drivers with valid data on all covariates, it is not certain that a more mathematically formal imputation would be successful. Nevertheless, it warrants consideration in future analyses of the study data.

More sophisticated versions of the weighting adjustments could be used. For example, instead of using the raw differences in risk relativities at each BAC, it would be possible to derive bias adjustment curves from these changes in relative risk and alter the uncorrected curves through numerical integration. The resultant adjustments would be slightly smoothed, but this would come at the expense of greater analytical complexity. A similar method could eliminate potential non-orthogonality in the adjustment weights for drivers who did not complete the interview.

The imputation strategies used in this study were confined to the total sample, and no attempt was made to adjust the subgroup analyses (e.g., age, drinking frequency, etc.) for non-participation bias. The biases likely are not the same for all subgroups. Although imputation for subgroups would further reduce sample size, the additional analyses are warranted.

<u>Additional analyses of the drinking consumption indices.</u> Alcohol consumption questions were based on the conceptual structure and statistical model developed by Gruenewald and Nephew (1994), but that model’s parameters, which involve the application of a log-logistic model and mathematically complex considerations, was not used for the analysis. Instead, standard quantity-frequency (Q-F) scales were derived from drivers’ responses.

The length of time (past 28 days) covered by the alcohol consumption questions creates a problem for data analysis. If a subject reported zero drinks in the past 28 days, he/she was categorized as a non-drinker. The questions did not determine how often and how much alcohol a driver might have consumed during periods other than the past 28 days. Gruenewald and Nephew (1994) asked subjects who reported no drinking in the past 28 days about their alcohol consumption over the prior 12-month period. In the present study, interview time was limited and questions about the longer time could not be asked. Nonetheless, the Gruenewald and Nephew (1994) approach could be used for additional analysis.

<u>Development of causal models.</u> This study’s case-control design and logistic regression model were intended to remove, or greatly reduce, confounds from the BAC-crash risk association. This objective appears to have been met. More complex statistical techniques could assess further the causal structure and hierarchical relationship among covariates that affect the BAC-crash risk relationship. Structural equation models can construct and evaluate causal models, including those that adjust for measurement errors (Jöreskog and Sörboing, 1993; Bollen and Long, 1993). These models normally are based on a priori theory and assumptions about the hierarchical pathways by which covariates exert indirect, direct and mediating effects on the dependent variable. Although the present study was limited to the role of alcohol and BAC, a majority of the sampled crashes did not involve

alcohol. The collected data are relevant to an examination of the general issue of crash correlates and general crash proneness.

Population attributable risk (PAR) associated with alcohol impaired driving. An estimate of PAR makes it possible to examine the potential payoff of countermeasures and treatments for a given risk factor. To estimate PAR for alcohol-involved crashes, it would be necessary to determine the proportion of the driving population who are at given BACs. The present study did not provide that data, and the control group data do not suffice for this purpose. Control drivers were selected from the same time and location as crashes, and crashes were selected from hours of the day when drivers who have been drinking alcohol are more likely to be on the roadway. The drivers sampled in Long Beach and Fort Lauderdale, therefore, are not representative of the population at risk. The data cannot be used to estimate the prevalence of alcohol in the population or the proportion or number of crashes that could be prevented if all driving at specific BACs were eliminated. This constraint was exacerbated by the operationally necessary decision to exclude freeway crashes from the present study. Possibly, however, prevalence estimates from other studies could be applied to the risk relativities obtained here to estimate PAR fractions.

5.6.2 Additional Research

Relative risk research. Although alcohol is known to influence both absolute speed and speed differential, the study data do not include the vehicle speeds or differences from the prevailing speed at the time and place of the crash or control sampling. The speed of control vehicles could have been measured with radar, but it is not feasible at the present to obtain accurate speed measures for crash vehicles. The speeds reported by independent witness are highly inaccurate, and those reported by involved drivers are likely to be both inaccurate and significantly biased. If data recorders become standard equipment in vehicles, it then will be possible to obtain speed data for crash vehicle. With that development, a study of relative risk will be able to examine speed as an important covariate.

Population at risk. The control group of more than 10,000 drivers provides data for non-crash involved drivers who were exposed to the same types of traffic conditions as the study crash drivers. It would be of interest to know whether that large control group is representative of the total at-risk population. A comparison of the control group data to roadside survey data from other research could shed light on this question. The roadside surveys need not be contemporaneous or co-located with this study. For example, data collected by Voas et al. (1998) could be compared with the study control data.

This Page Intentionally Left Blank

6. REFERENCES

Allsop, R. E. (1966). Alcohol and Road Accidents. (Road Research Laboratory Report No. 6). Harmondsworth, England: Road Research Laboratory, Ministry of Transport.

Appleton, D. R., French, J. M., and Vanderpump, M. P. J. (1996, November). Ignoring a covariate: An example of Simpson's paradox. The American Statistician, 50(4), 340-341.

Biecheler, M. B. (1983, Avril). Alcohol, driving and safety: Summary and analysis of research done by the French National Road Safety Agency. I.N.R.E.T.S. [National Institute for Transportation Safety], Arcueil, France.

Biecheler, M. B., Rambach, M. C., Filou, C., Goffetter, D. and Monseur, J. L. (1970). Étude alcool conduite et accidents de la route. I.R.T.-O.N.S.E.R., Arcueil, France.

Biecheler, M. B., Lefort, E., Rambach, M. C., Filou, C., Goffette, D., and Monseur, J. L. (1971). Étude alcool conduite et accidents de la route. I.R.T.-O.N.S.E.R., Arcueil, France.

Biecheler, M. B., Remond, M. C., Berlioz, C., Filou, C. and Monseur, J. L. (1974, Mai). Alcoolémie des conducteurs et accidents de la route. Cahiers D'Études, O.N.S.E.R., nº 32.

Blomberg, R.D. and Preusser, D.F. (1974, Vol. 6). Narcotic use and driving behavior. Accident Analysis and Prevention, 6, 23-32.

Bollen, K. and Long, J. (editors)(1993) Testing structural equation models. Newbury Park, Ca: Sage Publications Inc.

Borkenstein, R. F. (1985, July). Historical perspective: North American traditional and experimental response. Journal of Studies on Alcohol, Supplement No.10.

Borkenstein, R. F., Crowther, R. F., Shumate, R. P., Zeil, W. W., and Zylman, R. (1964). The role of the drinking driver in traffic accidents. Bloomington, IN: Department of Police Administration, Indiana University.

Borkenstein, R. F., Crowther, R. F., Shumate, R. P., Zeil, W. W., and Zylman, R. (1974). The role of the drinking driver in traffic accidents, 2nd ed. Blutalkohol: Alcohol, Drugs and Behavior,11 (Supplement1).

Breslow, N.E. (1966). Statistics in epidemiology: the case-control study. Journal of American Statistical Association, 91(433), 14-28.

Breslow, N.E. (2000). Statistics. Epidemiological Reviews, 22, 126-130.

Breslow N.E. and Storer B.E. (1985). General relative risk functions for case-control studies. American Journal of Epidemiology, 122, 149-62.

Carlson, W.L. (1979). Estimation of nonrespondent BAC using a priori judgment. Accident, Analysis and Prevention, 11, 35-41.

Cobb, P.W. (1940). The limit of usefulness of accident rate as a measure of accident proneness. Journal of Applied Psychology, 24, 154-159.

Cohen, J. and Cohen, P. (1983). Applied Regression/Correlation Analyses for the Behavioral Sciense. Hillsdale, NJ: Lawrence Erlbaum Assoc.

Committee on Trauma Research, Commission on Life Sciences, National Research Council and the Institute of Medicine. (1985). Injury in America: A Continuing Public Health Problem. Washington, D.C.: National Academy Press.

Coughlin, S., Nass, C., Pickle, L., Trock, B. and Brunin, C. (1991) Regression methods for estimating attributable risk in population-based case-control studies:a comparison of additive and multiplicative models. American Journal of Epidemiology, 133(3), p305-311.

D'Agostino, Jr., R.B. and Rubin, D.B. (2000). Estimating and using propensity scores with partially missing data. Journal of the American Statistical Association, 95(491).

Darlington, R.B. (1990). Regression and Linear Models. New York, NY: McGraw-Hill, Inc.

(Editorial) (1904). Drinking and motor wagon accidents, inebriate retreats and reformatories on England. Quarterly Journal of Inebriation, 26, 308-309.

Evans, L. and Frick, M. C. (1993). Alcohol's effect on fatality risk from a physical insult. Journal of Studies on Alcohol, 54, 441-449.

Ferguson, J. K. W. (1955). Keynote address. Alcohol and Road Traffic, Proceedings of the Second International Conference on Alcohol and Road Traffic (1953). Toronto: Canada: Garden City Press Co-operative, 4-8.

Gebers, M.A. (1998). Exploratory multivariable analyses of California driver record accident rates. Transportation Research Board Record, 1635, 72-80.

Graham, J.W. and Donaldson, S.I. (1993). Evaluating interventions with differential attrition: the importance of nonresponse mechanisms and use of follow-up data. Journal of Applied Psychology, 78(1), 119-128.

Greenland, S. (2000). Causal analysis in the health sciences. JASA, 95(449), 286-289.

Gruenewald, P. and Nephew, P. (1994). Drinking in California: theoretical and empirical analyses of alcohol consumption patterns. Addiction, 89, 707-723.

Haddon, Jr., W. (1964, June). Personal correspondence to Robert Borkenstein. Library of the Insurance Institute for Highway Safety.

Haddon, Jr., W., Valien, P., McCarroll, J. R., and Umberger, C. J. (1961). A controlled investigation of the characteristics of adult pedestrians fatally injured by motor vehicles in Manhattan. Journal of Chronic Diseases, 14(6), 655-678.

Harrington, D.P. (2000). The randomized clinical trial. Journal of the American Statistical Association, 95(449), 312-315.

Heise, H. A. and Halporn, B. (1932). Medicolegal aspects of drunkenness. Pennsylvania Medical Journal, 36, 190-195.

Holcomb, R. L. (1938). Alcohol in relation to traffic accidents. Journal of the American Medical Association, III, 1076-1086.

Hosmer, D.W. and Lemeshow, S. (2000). Applied Logistic Regression. New York, NY: John Wiley & Sons.

Hurst, P. M. (1973). Epidemiological aspects of alcohol in driver crashes and citations. Journal of Safety Research, 5(3), 130-148.

Hurst, P.M., Harte, D. and Frith, W.J. (1994). The Grand Rapids dip revisited. Accident Analysis and Prevention, 26(5), 647-654.

Hurst, P. M. and Darwin, J. H. (1977). Estimation of alcohol for nonrespondents in roadside breath surveys. Accident, Analysis and Prevention, 9, 119-123.

Jöreskog, K and Sörboing, D (1993) Structural equation modeling with the SIMPLIS Command language. Hillsdale, N J:LEA publishers.

Klein, D. and Waller, J. A. (1970, July). Causation, culpability and deterrence in highway crashes. Washington, D.C: U.S. Department of Transportation.

Lucas, G. H., Kalow, W., McColl, J. D., Griffith, B. A. and Smith, H. W. (1955). Quantitative studies of the relationship between alcohol levels and motor vehicle accidents. Alcohol and Road Traffic, Proceedings of the Second International Conference on Alcohol and Road Traffic (1953). Toronto, Canada: Garden City Press Co-operative,139-142.

Lundsgaard, E. (1951). Proceedings of the First International Conference on Alcohol and Traffic(1950). Stockholm, Sweden: Kugelbergs Boktryckeri, 261-262.

McCarroll, J. R., and Haddon, W., Jr. (1962). A controlled study of fatal automobile accidents in New York City. Journal of Chronic Diseases, 15, 811-826.

McClelland, G.H. and Judd, C.M. (1993). Statistical difficulties of detecting interactions and moderator effects. Psychological Bulletin,114(2), 376-390.

McLean, A. J., Holubowycz, O. T. and Sandow, B. L. (1980). Alcohol and crashes: Identification of relevant factors in this association (Report No. CR 11). Adelaide, Australia: Department of Transport, Office of Road Safety, Road Accident Research Unit, University of Adelaide.

Mittal, Y. (1991). Homogeneity of subpopulations and Simpson's paradox. JASA, 86(413), 167-172.

NHTSA. (2004). Traffic Safety Facts 2002: A compilation of motor vehicle crash data from The Fatality Analysis Reporting System and The General Estimates System (DOT HS 809 620). Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration, National Center for Statistics and Analysis.

Nielson, R.A. (1965, September). Alcohol involvement in fatal motor vehicle accidents. San Francisco,CA: California Traffic Safety Foundation.

Nielson, R.A. (1967, April). A survey of post-mortem blood-alcohols from 41 California counties in 1966 (10pp plus tables). San Francisco, CA: California Traffic Safety Foundation.

Peck, R.C., McBride, R.S. and Coppin, R.S. (1971). The distributions and prediction of driver accident frequencies. Accident Analysis and Prevention, 2, 243-299.

Perrine, M. W. (1970a). Drinking and driving: A comparison of deceased drivers and roadblock control motorists. Alkohol und Verkehrssicherheit, Konferenzbericht der 5. Internationalen Konferenz über Alkohol und Verkehrssicherheit Freiburg im Breisgau, 22-27. 9. 1969. Hans Ferdinand Schulz Verlag, Freiburg im Breisgau, 1970.

Perrine, M.W. (1970b). Identification of personality, attitudinal, and biographical characteristics of drinking drivers. Behav. Res. Highway Safety, 2, 207-226.

Perrine, M. W. (1970c). The spectrum of drinking drivers. Alcohol and Highway Safety, North Carolina Symposium on Highway Safety: Vol. 3. Chapel Hill, NC.

Perrine, M. W., Waller, J. A., and Harris, L. S. (1971). Alcohol and highway safety: Behavioral and medical aspects (Final Report on Contracts No. FH-11-6609 and FH-11-6899). Washington, D.C: U.S. Department of Transportation.

Rao, JNK. (1996). On variance estimation with imputed survey data. JASA, 91(434), 499-506.

Rosenbaum, P.R. and Rubin, D.B. (1985). Constructing a control group using multivariate sampling methods that incorporate the propensity of score. American Statistician, 39(1), 33-38.

Rothman, K.J. and Greenland, S. (1998). Modern Epidemiology. Philadelphia, PA: Lippincott-Raven.

Rubin, D.W. (1987) Multiple Imputation for Nonresponse in Surveys. New York, NY: John Wiley and Sons.

Samuels, M.L. (1993, March). Simpson's paradox and related phenomena. Journal of the American Statistical Association, 88(424), 81-88.

SAS. (1995). Logistic regression using the SAS system, version 6, first edition, Cary, NC: SAS Institute, Inc.

SAS. (1996). SAS/STAT Software: Changes and Enhancements, release 6.11, Cary, NC: SAS Institute, Inc.

Smith, H. W. and Popham, R. E. (1951). Accident contribution of drivers as a function of their blood alcohol levels. Proceedings of the First International Conference on Alcohol and Traffic. Stockholm, Sweden: Kugelbergs Boktryckeri.

Solnick, S. J. and Hemenway, D. (1994). Hit the bottle and run: The role of alcohol in hit–and-run pedestrian fatalities. Journal of Studies on Alcohol, 55, 679-684.

SPSS. (1997). SPSS Professional Statistics 7.5. Chicago, IL: SPSS, Inc.

State of California. (1999). 1998 annual report of fatal and injury motor vehicle traffic collisions, statewide integrated traffic records system (SWITRS), ix.

Stone, C. and Koo, C. (1985). Additive splines in statistics. Proceedings of the Statistical Computing Section, American Statistical Association, 45-48.

Tabachnick, B. and Fidell, L. (2000). Multivariate Statistics, 4th edition. Boston, MA: Allyn and Bacon.

U.S. Department of Transportation, National Highway Traffic Safety Administration. (1999, October). Traffic safety facts 1998: A compilation of motor vehicle crash data from the fatality analysis reporting system and the general estimates system,.Washington, D.C.

Vamosi, M. (1960). Determination of the amount of alcohol in the blood of motorists. Traffic Safety Res., Rev., 4(3), 8-11.

Vamosi, M. (1963). Experiences with non-alcoholic road traffic in Czechoslovakia. Proceedings of the Third International Conference on Alcohol and Road Traffic. London, England: Tavistock Square, British Medical Association House.

Voas, R. B., Wells, J., Lestina, D., Williams, A., and Greene, M. (1998). Drinking and driving in the Unites States: The 1996 national roadside survey. Accident, Analysis and Prevention, 30(2), 267-275.

Zador, P.L., Krawchuk, S.A. and Voas, R.B. (2000). Relative risk of fatal crash involvement by BAC, age and gender. Rockville, MD: Westal, Inc., (DOT HS809050).

Zobeck, T. S., Stinson, F. S., and Bertolucci, D. (1992, November). Surveillance report #22: Trends in alcohol-related fatal traffic crashes, United States: 1979-1990. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism, Division of Biometry and Epidemiology.

Zylman, R. (1971). Analysis of studies comparing collision-involved drivers and non-collision involved driver. Journal of Safety Research, 3(3), 116-128.

This Page Intentionally Left Blank

Appendix A

Additional Background Relevant to Study Design

This Page Intentionally Left Blank

1. ADDITIONAL BACKGROUND

This Appendix contains additional background material from the literature relevant to the design of the study procedures and analyses. References cited in this section are included in the reference list at the end of the body of the main report.

1.1 Key Methodological Issues

A review of the literature, together with an initial assessment of the task at hand, identified critical methodological issues. The design of a sound study of crash risk required a resolution of these issues

1.1.1 Controlled Experiment Versus Epidemiological Approach

Many variables affect alcohol-and-vehicle crashes, and an epidemiological study of the relationship, by definition, is difficult. Identification and control of all possible variables is seldom possible. That limitation and many other challenges confront investigators who wish to collect real world data from similar populations of drivers who have or have not crashed.

The alternate approach, a controlled on-the-road experiment, presents even greater challenges. Ethical and technological constraints preclude the use of controlled experimental designs that require alcohol administration to subjects and subsequent observation of their driving performance**.** Researchers and institutions are unwilling to assume the financial and moral responsibilities of allowing alcohol-dosed subjects to drive vehicles on the open road. Because drivers do not chose to drink or to drive at random, it can also be argued that a controlled experimental design, even if feasible, would have less ecological validity than a well done retrospective statistical analysis of epidemiological data.

It is therefore necessary to rely on data obtained from crashes as naturally occurring events. With such an approach, a vehicle crash serves as the trigger for sampling, and drivers are the unit of analysis. Non-crash drivers, or controls, are selected to be closely representative of the pool of drivers from which the crash-involved drivers came. Potential members of either the crash or control group, however, can and do refuse to participate. Since the missing data that results from such refusals can bias the results, non-participation is a serious problem for epidemiological studies.

1.1.2 Dealing with Covariates—The Case-Control Study

In a true, randomized experimental study, two groups of subjects would be equally likely to be involved in crashes except for the influence of an administered alcohol dose. In an epidemiological study, however, variables other than alcohol affect the probability of being involved in a crash, and they may be differentially present in crash-involved and non-crash-involved drivers. The number of such variables is large. It includes, but is not limited to, crash location, direction of travel, time of day, roadway geometry, weather, vehicle type and age, traffic density, driver experience, age, gender, duration of vehicle ownership, physical/emotional status, licit and/or illicit drug use, and miles driven per year. A valid comparison of the role of alcohol for crash-involved drivers and control drivers requires matching or compensating for these covariates.

Although it is not possible to know all of the relevant characteristics of drivers prior to their inclusion in a control sample, they must be as comparable as possible to the membership of the crash groups. A case-control study is one method for accomplishing the comparability. A case-control design increases the efficiency of the matching procedures by selecting control drivers who are matched on specific characteristics to members of the crash group. In the present study, control drivers were selected by random procedures exactly one week after a crash. They were drawn from a stream of traffic going in the same direction, on the same day of the week, at the same time of day and at the same location as the crash-involved drivers. This method reduces sampling error and increases the similarity of the crash and control groups.

Whether a study uses representative sampling or case-control methods to obtain the control group, the frequency of many variables important to crash causation will nonetheless differ between the two groups. The differences can bias the assessment of the influence of alcohol, the variable of primary interest. The residual biases that are due to differences between the crash and control samples in the frequencies of variables of interest can, however, be controlled through multiple regression techniques.

1.1.3 Measures of Relative Risk

All of the relevant studies discussed in the literature review (Section 2) used relative risk as a measure of crash probability. Allsop (1966) defined this measure as “the ratio of the number of crash drivers to the number of control drivers in class i [a BAC classification], divided by the corresponding ratio for the 0-9 mg/100 ml class.” The latter was the lowest BAC class in the Grand Rapids Study (Borkenstein et al., 1964). Thus, relative risk is defined as:

$$\frac{\frac{A_i}{C_i}}{\frac{A_0}{C_0}} \quad \text{which is equal to} \quad \frac{A_i \, x \, C_0}{C_i \, x \, A_0}$$

where: i defines a BAC level class
0 defines the lowest BAC level class (BAC < 0.001 in the Grand Rapids Study)
A defines the number of drivers in a given crash BAC level class
C defines the number of drivers in a given control BAC level class

For example, using the Grand Rapids Study data (Borkenstein et al., 1964) the following relative risk can be calculated for BACs greater than or equal to 0.10 and less than 0.12.

BAC (%)	Number of Crash Group Drivers	Number of Control Group Drivers	Relative Risk
0.00 – 0.009	4,992	6,756	1.0
0.10 – 0.119	92	21	5.93

In this example, the relative risk for drivers at a BAC of less than 0.001% (the “zero” group in the Grand Rapids Study) is set at 1.0 by convention. This is a “relative” risk measure, because it is relative to the base comparison drivers (the lowest BAC group). Based on the above formula, the relative risk for drivers with a BAC between 0.10% and 0.119% is calculated as:

$$\frac{92\,x\,6{,}756}{21\,x\,4{,}992} = 5.93$$

Note that the relative risk is actually an odds ratio, which, in case-control studies, provides the appropriate epidemiological index of relative risk.[1] A crash involvement curve can be generated by successively calculating the relative risk for all "i" BAC classes.

In evaluating effects on (or differences among) variables measured in terms of odds ratios, the ratios are usually transformed to logarithms "logits." In essence, this converts an additive scale (proportions) to a multiplicative scale ("logits"), which better conforms to the assumption of linearity in the parameter estimates of a logistic regression model (Hosmer & Lemeshaw, 2000). The logits, however, are less meaningful intuitively and are frequently transformed back to relative risks (odds ratios) through sample exponentiation. Both metrics are presented in the data analysis (Section 4) of this report.

1.1.4 Matching Control Drivers to Crash Drivers

Relative risk calculations assume that the members of the four sets of drivers in each calculation (crash and control; zero and elevated BAC) are the same with regard to crash propensity, except for their BACs. If the groups differ in other important ways, it is not possible to make a proper logical deduction about the risk-inducing effects of BAC. To illustrate, suppose the zero BAC crash group (A_0) contains many drivers under age 18 who have a very high crash rate without alcohol. Further suppose there were few or none of these young drivers in the crash groups at positive BACs (A_i). The calculation of relative risk then would underestimate the effect of alcohol, because young drivers, who have a high likelihood of a crash without alcohol, are not among the positive BACs. The overrepresentation of these young drivers in the base comparison group would make it difficult to discern the influence of alcohol since, in effect, the analysis would compare the crash rate of older drivers with positive BACs to the crash rate of young drivers with zero BACs. Thus, matching of control to crash drivers or adjusting for covariates is required for logically consistent comparisons of crash rate at different BACs.

1.1.5 Representative Samples vs. Refusals to Participate

It is essential to ensure that both crash and control samples are representative of crash and control populations. In a review of methodological issues for a study of highway crashes, Klein and Waller (1970) noted that police and other officials reported only 36% to 81% of the crashes that respondents acknowledged during research interviews. Also, it is generally accepted that many highway crashes, both property damage and injury, go unreported. Crashes that are not witnessed, especially single vehicle crashes, nighttime crashes, crashes involving animals, and crashes with minor damage are the most likely to not be reported. Single vehicle crashes that occur at night are highly likely to involve alcohol. Estimates of the BACs of drivers involved in unreported crashes, however, are speculative.

---

[1] For reasons noted later, the computation of relative risk is often based on the logarithm of the odds ratio.

Even when a crash is reported, and police and researchers respond, not all crash drivers agree to participate in a research study. There is evidence, discussed in Section 4 of this report, that the drivers who refuse to provide breath specimens or fail to complete questionnaires are more likely to have positive BACs. Hit-and-run drivers present an even more disturbing problem. In California in 1998, hit-and-run drivers were estimated to be about 10% of total injury and fatal crashes (State of California, 1999). The inclusion of less severe property damage only crashes would likely raise the hit-and-run percentage. From an examination of the Fatality Analysis Reporting System (FARS) data, Solnick and Hemenway (1994) reported that 19% of all pedestrian fatalities for 1989 and 1990 were hit-and-run. Based on the characteristics of the drivers who eventually were caught and crash times, they concluded that alcohol was a factor in many of these crashes.

To prevent an underestimation of the prevalence of alcohol, researchers also must estimate the BACs of potential control group drivers who do not cooperate. Although it is easier to sample drivers for a control group than to sample crash drivers, willingness to participate remains an issue. Voas, et al. (1998) found breath sampling refusal rates of 13.7%, 6.3% and 4.3% in national roadside surveys conducted in 1973, 1986 and 1996. In the latter two surveys, data obtained with passive alcohol sensors demonstrated that the drivers who refused were more likely to have positive BACs than those who cooperated.

To note that samples must be representative does not mean that the entire populations of drivers or of crashes must be sampled. In studies of the influence of alcohol, little would be gained by sampling drivers during the daytime hours when the incidence of alcohol is low. To be representative, the characteristics of drivers who enter the study must be known, and they must be representative for the times and places that the samples were drawn.

1.1.6 Measurement of BAC

Accurate, quantitative assessments of crash participants' BACs are essential to studies of the role of alcohol in crash causation. In early studies, the valid and reliable assessment of alcohol in drivers was problematic. Prior to the development of improved sobriety tests and portable breath testers (PBTs), the information was obtained from police officers' classifications of crashes and their assessments of the drivers. Studies indicate, however, that police officers tend to underestimate alcohol involvement in crashes. Borkenstein (1985) noted that the Pennsylvania Deputy Commissioner for Motor Vehicles declared in 1931 that less than 1% of more than 2,000 drivers killed that year were drunk. Heise responded by performing urine alcohol tests on 15 consecutively-hospitalized drivers who had been injured in motor vehicle crashes. He found that nine had consumed considerable alcohol (Heise and Halporn, 1932).

Police training academies, as well as special seminars and conferences, now offer training for officers in the recognition of the signs and symptoms of alcohol. Also, many officers now routinely use standardized sobriety tests and PBTs when they suspect drivers of being under the influence of alcohol.

1.1.7 Effects of Crash Severity

The probability that alcohol is involved in a crash differs by crash severity (fatal, injury, property damage only). For example, 1998 data (NHTSA, 1999) show that alcohol was present in 3% of 7,560,000 drivers in property-damage-only crashes. It was present, however, in 5% of 3,751,000 drivers in injury crashes and 23% of 56,543 drivers in fatal crashes. The BAC distribution by crash in Table 46 of the Grand Rapids Study (Borkenstein, et al., 1974) also shows that the severity of injury increases as the probability of alcohol presence increases.

It is possible to compute the relative risk of crash involvement as a function of alcohol presence and crash severity for the Grand Rapids data. The computation for alcohol presence yields 1.3 for no-injury crashes, 1.3 for pain complaint but no evident injury, 1.7 for visible injury and 1.9 for serious injury. Thus, both U.S. Federal agency data and the Grand Rapids Study data demonstrate that the relationship between crash involvement and BAC differs by severity of injury.

Why are there differences in relative risk as a function of crash severity? It has been shown that the crash characteristics and also the body’s responses to trauma are affected by alcohol. A report by the Committee on Trauma Research, the National Research Council and the Institute of Medicine (1985) reported that the more severe the event, the higher the percentage in which alcohol plays a role in both highway and non-highway events. The studies found that alcohol is important in injury causation and increases the severity of the injury. Furthermore, its chronic use interferes with normal body repair processes. A decreased probability of safety belt use and the higher velocity of involved vehicles also underlie the increased severity of injuries in alcohol related crashes.

1.2 Methodology of the Grand Rapids Study

The Borkenstein et al. study (1964, 1974) was a large controlled study of alcohol-involved collisions. It was not, however, a matched case-control study. The control group (8,008 drivers) was created by sampling four drivers at 2,000 crash sites selected at random from a pool of 27,000 crashes that had occurred during the three preceding years. Four control drivers were sampled at all sites at the time of day and day of the week of the selected prior crash. The direction of traffic for sampling control drivers was randomly chosen. The BACs of 5,985 collision drivers and 7,590 control drivers were measured. Some drivers in both groups refused to cooperate, some crash drivers fled, and some potential control drivers avoided contact. During peak crash periods, it was not possible to go to all crashes. The investigators did not compute the relative risk of *involvement* in a crash as a function of BAC. Rather, they created a BAC distribution of crash-causing drivers and generated a figure purported to indicate the relative risk of *causing* a crash as a function of BAC.

There were important differences between the crash and control groups. Four control drivers were sampled for each of the 2,000 control crash sites regardless of the number of drivers in the crash, and this over-represented control drivers for sites of single vehicle crashes (see Table 44 in Borkenstein et al., 1974). The overrepresentation is noteworthy, because single-vehicle crash drivers differ from multiple-vehicle crash drivers on various characteristics, including more frequent positive BACs.

Also, factors known to be relevant determinants of crashes such as time of day, day of week and month of year differed between the crash and control groups and added error variance. Controls were matched on time of day and day of week to crashes sampled from all crashes during the previous three-year period not the crashes that entered the study sample. The failure to match the control group to the crash group was ameliorated to some extent by the very large number (2000) of crash sites used to define the control sample.

An inappropriate assumption underlies the creation of the study's Chart XV that presents a relative risk curve. Hurst's 1973 analysis showed that the BAC distributions of not-at-fault crash-involved drivers and control drivers were not the same. There were more positive BACs among the not-at-fault than among the control drivers. Additional evidence comes from a study of 361 drivers killed in collisions for which police attributed responsibility to the other driver. Twenty-one percent of the fatally injured drivers were alcohol positive, and the BACs of 12% were greater than 0.10% (Nielson, 1965, 1967). It appears possible that alcohol prevents a driver from making maneuvers to avoid crashes from actions initiated by others.

The creation of the crash group from the distribution of at-fault drivers only resulted in a larger proportion of drivers with alcohol than contained in the total group of crash-involved drivers. A review of the Grand Rapids Study, reported by Allsop (1966), includes a relative risk calculation for crash involvement as a function of BAC. The relative risk curve of drivers "causing" a crash (Borkenstein, et al., 1974, Chart XV) rises at roughly double the rate for the complete, original sampled group (Allsop, 1966, Table 1). The Grand Rapids Study estimates of the reductions in total crashes that would be achieved by preventing driving at various BACs are, therefore, roughly triple the estimates by Allsop. Compare Borkenstein et al., 1974, Chart XV with Allsop, 1966, Figure 3.

The Borkenstein causation curve shows the relative risk at 0.01% - 0.04% BACs to be less than the risk at zero BAC. This Grand Rapids dip, as it is known, has generated controversy and further analyses, and several explanations are possible. The dip could be an artifact of the crash – control univariate comparison, which did not control for various crash-determining covariates. In several analyses with bivariate techniques, the dip has been eliminated. For example, in an analysis of BAC, drinking frequency and relative crash involvement risk (Allsop, 1966), the dip disappeared. The following excerpts from Allsop (1966) are informative:

> *… for each drinking frequency class the lowest accident rate may well occur in the 0-9 mg/100 ml alcohol level interval, i.e. even the frequent drinker should keep his drinking and driving separate in order to minimize his accident risk.*
>
> *Borkenstein's report does not give details of two-factors analysis by drinking frequency and the other seven variables, but it does state that age-groups having the best and worst accident experience are respectively over and under represented among the frequent drinkers compared with the infrequent drinkers, and that the same is true when age is replaced by any other major variable except alcohol level. It thus appears that part of the difference between the accident risks of frequent and infrequent drinkers may be attributable to the effect of other variables.*

Hurst (1973) also independently calculated a bivariate analysis of relative risk of crash involvement for BAC by drinking frequency for all BACs from the Grand Rapids data (Figure 1), and again the dip disappears. As Allsop had previously pointed out, there is a unimodal increase in crash rate for any increase in BAC for all drinking frequencies. These and the following analyses are not without their own problems, however, as noted below.

Hurst, Harte & Frith (1994) used a general linear model to adjust the BAC – crash risk relationship for drinking frequency and produced a steeper risk function that eliminated the dip at low BAC levels. There are, however, a number of problems with their analysis. First, it assumed a model in which the relationship between BAC and crash risk does not vary as a function of drinking frequency and in which the self-reported drinking levels are relatively free of systematic error from deception or mis-recollection. Both of these assumptions are questionable. A visual inspection of Figure A-1 is sufficient to raise questions about the validity of the self-reported frequencies. Note, for example, there is a perfect inverse ordinal association between drinking frequency and crash risk. That is, the more often drivers reported drinking, the less likely they were to be crash involved. Daily drinkers have the lowest crash rates and yearly or non-drinkers have the highest rates. This relationship would be expected if crash-involved drivers were more likely to deny or conceal their actual drinking frequency. Even if the self-reported drinking frequency levels were free of error, it is not clear why one should adjust the BAC-crash risk curve for drinking frequency if, for example, frequent drinkers acquire some degree of tolerance (i.e., are able to perform better at given BAC levels than drivers who seldom drink).

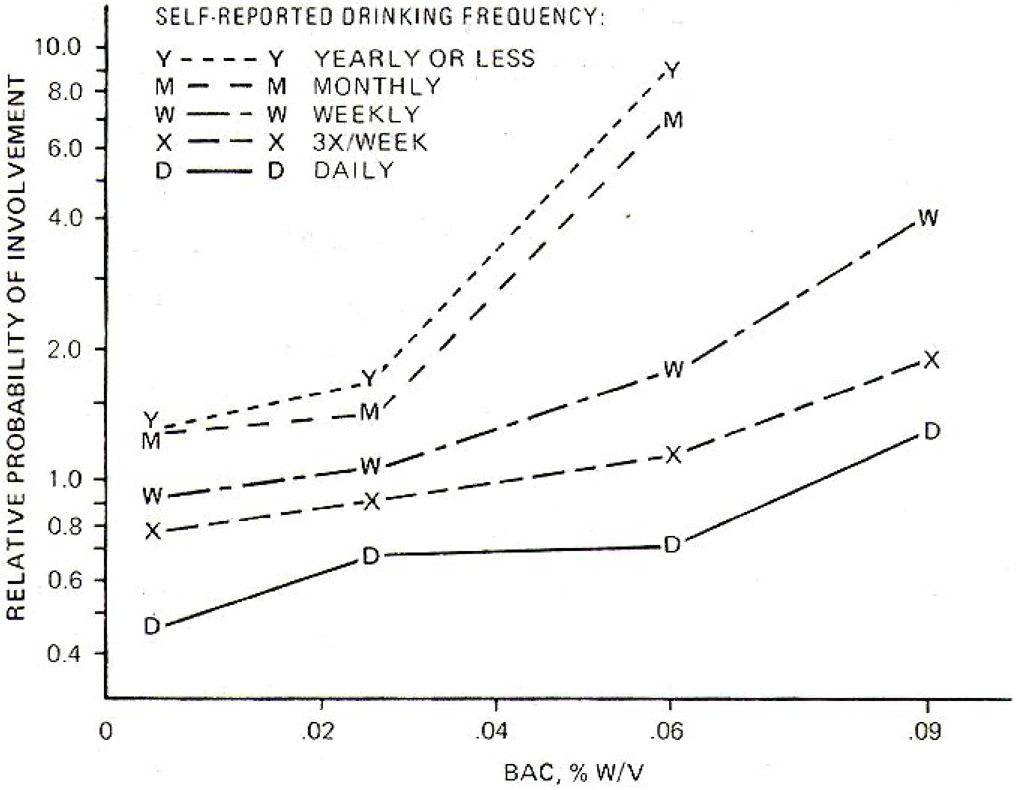

Figure A-1. Relative probability of crash involvement by self-reported drinking frequency (Figure 3 from Hurst, 1973)

The Allsop monograph (1966) contains an algebraic explanation of the statistical paradox illustrated by the Grand Rapids Study. More extensive discussion of the statistical reason for the phenomenon, which is known as “Simpson’s Paradox,” is beyond the scope of this discussion but can be found in the literature (cf. Appleton, et al., 1996; Samuels, 1993). This topic is also revisited with respect to the present study’s data in Section 4.6.10 of this report.

Although there is no evidence that the Grand Rapids dip actually exists, the decline at low BACs could be the result, at least in part, of compensatory behavior by drivers. It is possible that after ingesting a small amount of alcohol drivers exercise extra caution and thereby compensate for the performance decrements caused by very low levels of alcohol. This phenomenon has been documented with respect to driving by heroin addicts (Blomberg and Preusser, 1974) but not for drinking drivers.

No statistical adjustment was made to control for covariates so the unbiased relationship between BAC and crash involvement is unknown. Even without detailed analysis, however, the patterns of the covariates suggest that adjustments are necessary. For example, it can be seen in Figure A-1 that the various drinking frequency groups varied in crash probability by more than 300% at zero BAC. This is further compounded by the fact that it is highly likely that covariates such as age varied within the drinking frequency groups. Analysis of the total set of Grand Rapids data could have unraveled the intercorrelations that determine crash frequency, but the data are not available for reanalysis.

Two other lessons from the Grand Rapids Study are relevant to the present effort. One is the need to account for differential refusal rates between the crash and control groups in providing breath specimens or in completing the questionnaire. In the Grand Rapids Study, 4.7% of the crash drivers and 2.2% of the control group were in the refusal group, and the probability of refusal appears to have been greatest for drivers in the high drinking frequency group.

A second issue, hit-and-run drivers, is mentioned in the Grand Rapids Study report, but no quantitative data were reported (Borkenstein, et al., 1974, p. 28), This affects the validity of the crash group’s BAC distribution. In the crash group 16.6% was positive for alcohol, but the report did not indicate what portion of the crashes were hit-and-run, nor the BACs of the hit-and-run drivers who were apprehended a short time after the crash. The victims of the hit-and-run crash were included in the study. From a review of 1989 – 1990 FARS data, Solnick and Hemenway (1994) reported that hit-and-run drivers were involved in 19% of pedestrian fatalities. Roughly 50% of the drivers were eventually apprehended, and it was found likely that a high proportion had a positive BAC at the time of the crash. If hit-and-runs in Grand Rapids in 1962-3 approached the current U.S. average of 12% of all crashes, and if half had alcohol present, that would equate to 25% of the entire positive BAC crash population. That would have massively influenced the relative risk curve.

Appendix B

English Language Questionnaire

This Page Intentionally Left Blank

Interviewer I.D.________

Driver No. 2._ _ _ _._ _._

B. Code	N	S	O

ENGLISH QUESTIONNAIRE

(OMB Approval # 2127-0589)

1. In a typical week, how many miles do you drive?	

2. Where did you begin this trip?

1	Your home	
2	Other's home	
3	Liquor store	
4	Other business	
5	Bar / Tavern	
6	Restaurant	
7	Work	
8	Church / Temple	
9	Movie, game, other recreation	
10	Airport, bus or train depot	
11	Hotel or motel	
12	Other (specify) ____________ ____________	

3. What is your birthdate?

___/___/___
Month Day Year

4. What is your marital status?

1	Single	
2	Married	
3	Living Together	
4	Divorced	
5	Separated	
6	Widowed	

5. How many full years of regular school have you completed?

6. Are you currently employed full time, part time, unemployed and looking for work, retired, going to school, homemaker, disabled, or something else?

1	Employed Full Time	
2	Employed Part Time	
3	Unemployed and Looking for Work	
4	Retired	
5	Going to School	
6	Homemaker	
7	Disabled	
8	Other (specify) ______________________ ______________________	

7. (If employed or going to school) What are your work hours?

Start work @ ____

End work @ ____

8. (If employed) What kind of work do you do?
[Record drivers exact responses]

____

____

Occupation Code

Interviewer I.D.________ Driver No. 2._ _ _ _._ _._

9. **Please** look at this card and tell me the number that best describes your ethnic group.

1	African-American	
2	Asian	
3	Caucasian	
4	Hispanic / Latino/a	
5	Native American	
6	Pacific Islander	
7	Multi-ethnic	
8	Other (specify) ____________ ____________	

10. During the last week, how many hours did you sleep on average each night?

________ no. hours

11. The last time that you slept, how many hours did you sleep?

________ no. hours

12. What time did you wake up?

________ hours

13. Have you had any alcoholic beverages in the last 2 hours?

1	Yes	
2	No	

14. Have you had any alcoholic beverages in the last 30 minutes?

1	Yes	
2	No	

15. Which type of alcoholic beverage do you usually drink, if any?

1	Beer	
2	Liquor	
3	Wine	
4	Don't drink	

ASK ALL OF THE REMAINING QUESTIONS IN ALL CASES

A drink is defined as: 1 Can of Beer (12 Oz, 355mL)
1 Mixed Drink
1 Glass of Wine

16. Think about the alcohol that you drank during the past 28 days.
How many **days** did you have....

		# Days
1	..one or more drinks?	
2	..more than one drink?	
3	..three or more drinks?	
4	..six or more drinks?	
5	..nine or more drinks?	

17. Have you <u>ever</u> been arrested for driving under the influence of alcohol/DUI?

1	Yes	
2	No	

18. In the last 28 days (4 weeks), how many times have you driven a motor vehicle within
two hours of drinking an alcoholic beverage?

Interviewer I.D.________ Driver No. 2._ _ _ _._ _._

19. Are you using any medicines or drugs?	1	Yes	☐	What medicines or drugs? ________ ________ ________ ________ ________
	2	No	☐	

"Just one more thing to do, and then we'll be finished."
REQUEST BREATH SPECIMEN
GIVE INSTRUCTIONS FOR PRELIMINARY BREATH TESTER

20. Breath specimen time and reading	________ Time
	0.____%

If first breath specimen does not give a valid BAC, ask driver to blow into PBT again.

21. Second breath specimen time and reading	________ Time
	0.____%

C. Code	N	M	S

This Page Intentionally Left Blank

Appendix C

Crash Site Data Collection Forms

This Page Intentionally Left Blank

Crash Site Observation Form   Interviewer I.D.________   Case No. 2._ _ _ _.

1. Date (MM/DD/YY)	

2. Time Notified of Crash (2400HrTime)	

3. Time Arrived at Scene(2400 Hr Time)	

4. Weather

(Check 1 to 2 Items)

A	Clear	
B	Cloudy	
C	Raining	
D	Snowing	
E	Fog/Visibility FT.	
F	Other (describe) ___	
G	Wind	

5. Lighting

A	Daylight	
B	Dusk - Dawn	
C	Dark - Street Lights	
D	Dark - No Street Lights	
E	Dark - Street lights not Functioning	

6. Roadway Surface

A	Dry	
B	Wet	
C	Snowy - Icy	
D	Slippery (Muddy, Oily, etc.)	

7. Roadway Conditions

Mark 1 to 2 Items

A	Holes, Deep Ruts	
B	Loose Material on Rdwy	
C	Obstruction on Rdwy	
D	Construction - Repair Zone	
E	Reduced Rdwy Width	
F	Flooded	
G	Other	
H	No Unusual Conditions	

8. Type of Roadway and Number of Lanes

	TYPE OF ROADWAY	√	#
A	Freeway		
B	Interstate Highway		
C	City Surface		
D	Rural Road		
E	Other Type (describe)__ ____________		
F	Intersection (describe)__ ____________		

9. What Can Be Seen Within One Block of Crash Location

(Check All That Apply)

A	Alcohol Outlet (On Site - Bar, Tavern, Restaurant)	
B	Alcohol Outlet (Off Site - Liquor Store, Market)	
C	Homes	
D	Apartment Buildings	
E	Professional Buildings	
F	Retail Stores / Small Businesses	
G	Warehouses, Industry, Manufacturing	
H	Other. Specify:_	

10. Type of Crash

PDO		A	Head-on	
Injury		B	Sideswipe	
Fatality		C	Rear end	
Hit-Run		D	Broadside	
		E	Hit object	
		F	Overturned	
		G	Vehicle/Pedestrian	
		H	Vehicle/Train	
		I	Vehicle/Bicycle	
		J	Vehicle/Motorcycle	
		K	Vehicle/Animal	
		H	Other ____________ ____________	

11. Time of Crash (Estimated)	

Crash Site Observation Form

12. Number of Motor Vehicles Involved ☐

13. Number of Pedestrians ☐

14. Number of Bicycles ☐

Comments: ____________________

15. Sketch Scene.

Record Street Names, Numbers and Directions.

Label Vehicles by Number

✧ Indicate North

16. Vehicle/driver Information

Vehicle Number	Vehicle Type	Commercial Markings? (Y/N)	Gender of Driver (M/F)	Age of Driver	Ethnic Group	No. Of Passengers	Agree to Interview (Y/N/X)
#1							
#2							
#3							
#4							
#5							
	<u>Sample</u> 2=2-door 4=4-door T=Convertible D=Sports Car W=Station Wagon V=Van/Minivan P=Pickup O=Other		<u>Do NOT sample</u> M=Motorcycles S=Mopeds B=Buses/School buses K=Trucks C=Construction Equipment		1=African-American 2=Asian 3=Caucasian 4=Hispanic 5=Native American 6=Pacific Islander 7=Multi-ethnic 8=Other		

17. Suggested Control Location

✧ Indicate North

Fort Lauderdale
Police Crash Observation Form

Team Officer I.D. __________ **Case No. (5 digits)** 2 __ __ __ __

Report Number ______________________

Driver # as on Report	PAS Reading # of Bars	PAS Confidence (High, Medium, Low)	Agree to Interview? (Y/N/X)	Comments and Reason for refusal
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

Fort Lauderdale
Non-Project BAC Form*

Interviewer I.D.________

Driver No. 1._ _ _ _._ _.

1. Type of BAC reading?

1	Police Evidential	
2	Police PBT	
3	Hospital Blood	
4	Coroner	
5	Urine	
6	Other_______________________________	

2. What was the reading?

____0._______%

3. Time Sample was drawn? (2400 hr time)

____
Time

* To be used only if BAC is taken by non-team member with non-team equipment.

FLPD

Appendix D

Certificate of Confidentiality

This Page Intentionally Left Blank

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

National Institute on Alcohol
Abuse and Alcoholism
6000 Executive Boulevard
Rockville, MD 20892-7003

CONFIDENTIALITY CERTIFICATE

issued to:

Individuals Responsible for the Research Entitled

"Crash Risk of Alcohol Involved Driving"

NHTSA Contract Number: DTNH22-94-C-05001

Conducted by:

Dunlap and Associates, Inc.
Stamford, Connecticut 06905-5503

In accordance with the provisions of section 301(d) of the Public Health Service Act (42 U.S.C. § 241 (d)), this Certificate is issued in response to the request of Mr. Richard D. Blomberg, Principal Investigator, to protect the privacy of research subjects by withholding their identities from all persons not connected with the research. Mr. Blomberg is primarily responsible for the conduct of this research.

Under authority vested in the Secretary of Health and Human Services by Section 301(d) of the Public Health Service Act (42 U.S.C.), all persons who:

1. are employed by Dunlap and Associates, Inc., or its contractors and cooperating agencies; and

2. have, during the course of that employment, access to information which would identify individuals who are subjects of the research on alcohol and drug abuse in the study entitled: "Crash Risk of Alcohol Involved Driving"

are hereby authorized to protect the privacy of the individuals who are the subjects of that research by withholding their names and other identifying characteristics from all persons not connected with the conduct of that research, with the exceptions and limitations set forth below.

The primary goals of the proposed project are to determine: 1) the relative crash risk of drivers at various Blood Alcohol Concentrations (BACs) compared to drivers at zero BAC, while controlling for other factors (e.g., age, gender, drinking patterns) that may be related to crash risk; and 2) the relative risk of groups of drivers (e.g., young males, the elderly, heavy drinkers) at various BACs compared to similar groups at zero BAC.

As provided in Section 301(d) of the Public Health Service Act (42 U.S.C. § 241 (d)):

> "Persons so authorized to protect the privacy of such individuals may not be compelled in any Federal, State or local, civil, criminal, administrative, legislative, or other proceedings to identify such individuals."

The following conditions apply to the protection provided under this Certificate:

(1) This Certificate does not authorize Dunlap and Associates, Inc., or its contractors and cooperating agencies to refuse to reveal information concerning research subjects if any of the following conditions exist:

  (a) The subject (or, if he or she is legally incompetent, his or her guardian) consents in writing to disclosure of identifying information.

  (b) Authorized personnel of the United States Department of Health and Human Services request such information for audit or program evaluation of the research project, or for investigation of Dunlap and Associates, Inc., or its contractors or employees in carrying out the research project.

(2) This Certificate requires that there be no disclosures of identifying characteristics of research subjects in any Federal, State, or local civil, criminal, administrative, legislative, or other proceedings to compel disclosure of the identifying characteristics of research subjects, except as provided for in paragraph (1), above.

(3) The Confidentiality Certificate does not govern the voluntary disclosure of identifying characteristics of research subjects.

(4) This Certificate does not represent an endorsement of the research project by the Department of Health and Human Services.

(5) All research subjects in the project will be given a fair, clear explanation of the protection this certificate affords, and of the limitations and exceptions to the protection and will be given a copy of this Certificate, except when to do so would put them at risk.

(6) This Certificate is effective upon issuance and will expire on **June 1, 2002**, or sooner if the holder is notified of cancellation in accordance with the procedures set out in 42 C.F.R. § 2a.8. The protection afforded by this Confidentiality Certificate is permanent with respect to

subjects who participate in the research during any time the Certificate is in effect.

(7) This research project may also be subject to special rules for the confidentiality of alcohol and drug abuse patient records, under sections 544 and 548 of the Public Health Service Act (42 U.S.C. §§ 290dd-3 and 290ee-3), and implementing regulations at 42 C.F.R. Part 2, which restrict voluntary disclosures of information from covered patient records.

Enoch Gordis

for Mary Dufour, M.D., M.P.H.
Deputy Director
National Institute on
Alcohol Abuse and Alcoholism

MAY 23 1997
Date

This Page Intentionally Left Blank

Appendix E

Control Site Data Collection Form

This Page Intentionally Left Blank

Control Site Observation Form Interviewer I.D.________ Case No. 2._ _ _ _.

1. Date (MM/DD/YY) (Crash+1 Week) ________

2. Scheduled Start time(2400 Hr Time) ________

3. Weather

(Check 1 to 2 Items)

A	Clear	
B	Cloudy	
C	Raining	
D	Snowing	
E	Fog/Visibility FT.	
F	Other (describe) ___	
G	Wind	

4. Lighting

A	Daylight	
B	Dusk - Dawn	
C	Dark - Street Lights	
D	Dark - No Street Lights	
E	Dark - Street lights not Functioning	

5. Roadway Surface

A	Dry	
B	Wet	
C	Snowy - Icy	
D	Slippery (Muddy, Oily, etc.)	

6. Roadway Conditions

Mark 1 to 2 Items

A	Holes, Deep Ruts	
B	Loose Material on Rdwy	
C	Obstruction on Rdwy	
D	Construction - Repair Zone	
E	Reduced Rdwy Width	
F	Flooded	
G	Other	
H	No Unusual Conditions	

7. Type of Roadway and Number of Lanes

	TYPE OF ROADWAY	√	#
A	Freeway		
B	Interstate Highway		
C	City Surface		
D	Rural Road		
E	Other Type (describe)_ ____________		
F	Intersection (describe)_ ____________		

8. What Can Be Seen Within One Block of Crash Location

(Check All That Apply)

A	Alcohol Outlet (On Site - Bar, Tavern, Restaurant)	
B	Alcohol Outlet (Off Site - Liquor Store, Market)	
C	Homes	
D	Apartment Buildings	
E	Professional Buildings	
F	Retail Stores / Small Businesses	
G	Warehouses, Industry, Manufacturing	
H	Other. Specify:_	

Comments: ______________________________________

____

____

____

____

____

____

____

____

____

FLPD

Control Site Observation Form

9. Control Stops

Drivers Stopped	Vehicle Type	Commercial Markings (Y/N)	Gender (M/F)	Age	Ethnic Group	No. Passengers	PAS Reading (# of Bars)	PAS Confidence **H**igh, **M**edium, **L**ow	Behavioral Alcohol Assessment **N**o signs, **S**uspicion, **O**bvious signs	Agree to Interview (Y/N)	Comments and Reasons for Refusal	Crash Driver #	1 Digit Control #
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
	Sample 2=2-door 4=4-door T=Convertible D=Sports Car W=Station Wagon V=Van/Minivan P=Pickup O=Other		Do NOT sample Motorcycles Mopeds Buses/School buses Trucks Construction Equipment		1=African-American 2=Asian 3=Caucasian 4=Hispanic 5=Native American 6=Pacific Islander 7=Multi-ethnic 8=Other								

10. Actual sampling END time

FLPD

Appendix F

Sample Description

List of Tables in Appendix F

1. Descriptive Results

Different research teams obtained data a year apart in time at two sites that are separated by a continent. Nonetheless, the sites produced similar distributions of most of the study variables. This appendix describes the database, presents variable tabulations, and points out directional trends for data collected from 2,871 crashes and 14, 985 drivers in Long Beach, CA and Fort Lauderdale, FL. Statistical tests of between-site differences are not shown, however, because the large sample sizes would make small, operationally meaningless differences statistically significant. Also, univariate differences between sites are irrelevant to the determination of relative risk controlled for the effects of differences in the covariates. As can be seen in Section 4, the multivariate relative risk analyses do take into account the main effect of site, site x covariate interactions, and the main effect of important covariates.

1.1 Number of Crashes and Crash-Involved Drivers

The Long Beach site required a three-month extension of the data collection period to reach the target number of 1,300 crashes. During 15 months, 1419 crashes were sampled. The Fort Lauderdale site sampled 1,452 crashes in 12 months. Table F-1 presents the number of crashes and drivers by crash severity. The crash and driver distributions by severity are similar, as would be expected.

Property damage only (PDO) crashes account for 59.1% of the total sample, but the incidence differed by site. PDO crashes were 68.5% of the Fort Lauderdale and 49.5% of the Long Beach sample. At the opposite extreme of severity, 11 of 19 fatalities occurred in Fort Lauderdale.

A higher percentage of injury and hit-and-run crashes occurred in Long Beach than in Fort Lauderdale, a finding that may reflect differences in team responsibilities. Long Beach teams were first responders and were dispatched preferentially to the most severe crashes. Fort Lauderdale teams were dispatched to the closest crash without regard for severity.

Table F-1 includes Hit-and-Run (severity unknown) and Missing Severity cases to provide a complete accounting of sampled crashes. Crash severity sometimes was apparent even though one driver fled the scene, but the tabled number in the Hit-and-Run category includes cases for which severity could not be determined. The Missing Severity category (6 in Long Beach; 1 in Fort Lauderdale) consists of data collection anomalies.

The total and mean numbers of drivers per crash also appear in Table F-1. Across categories, the mean is the same for both sites, 1.7 drivers per crash. Consistent with previous research, the fatality category had the smallest mean number of drivers per crash.

Table F-1. Crashes and Drivers by Site and Severity

Crash Severity			Long Beach			Fort Lauderdale			Total		
			# Crashes	# Crash Drivers*	Drivers per Crash	# Crashes	# Crash Drivers*	Drivers per Crash	# Crashes	# Crash Drivers*	Drivers per Crash
	Property Damage Only	Count	702	1250	1.8	994	1760	1.8	1696	3010	1.8
		% of Crashes at Site	49.5%	51.6%		68.5%	70.5%		59.1%	61.2%	
		% of Total	24.5%	25.4%		34.6%	35.8%				
	Injury	Count	355	604	1.7	248	402	1.6	603	1006	1.7
		% of Crashes at Site	25.0%	24.9%		17.1%	16.1%		21.0%	20.5%	
		% of Total	12.4%	12.3%		8.6%	8.2%				
	Fatality	Count	8	12	1.5	11	12	1.1	19	24	1.3
		% of Crashes at Site	0.6%	0.5%		0.8%	0.5%		0.7%	0.5%	
		% of Total	0.3%	0.2%		0.3%	0.2%				
	Hit-and-Run (Severity Unknown)	Count	348	544	1.6	198	322	1.6	546	866	1.6
		% of Crashes at Site	24.5%	22.5%		13.6%	12.9%		19.0%	17.6%	
		% of Total	12.1%	11.1%		6.9%	6.5%				
	Missing Severity	Count	6	12	2.0	1	1	1.0	7	13	1.9
		% of Crashes at Site	0.4%	0.5%		0.1%	0.0%		0.2%	0.3%	
		% of Total	0.2%	0.2%		0.0%	0.0%				
	Total	Count	1419	2422	1.7	1452	2497	1.7	2871	4919	1.7
		% of Total	49.4%	49.2%		50.2%	50.8%		100.0%	100%	

*Classification of drivers is by the most serious outcome to anyone involved, and not the severity to the individual drivers.

1.2 Participation Rates

Table F-2 summarizes the numbers of drivers who participated and refused to participate in the study. Of the 10,066 drivers who were asked to participate as control subjects, 9,853 (97.9%) agreed to the interview and 9,821 (97.6%) provided breath specimens. The rates were similar for both sites.

For the total sample, 3,986 of the crash-involved drivers available for interview participated, and 279 refused to participate. Thus, 93.5% of the 4,265 drivers approached at the scene agreed to an interview. The total number of drivers actually eligible for the study (4,919), however, included an additional 654 as described below:

- 499 hit-run drivers not apprehended
- 104 hit-and-run apprehended drivers who participated (no breath specimen for 10)
- 51 drivers who were unable to participate (did not understand questionnaire languages).

Table F-2. Crash and Control Drivers' Participation by Site

		Long Beach			**Fort Lauderdale**			**Total**		
		Crash	**Control**	**Total**	**Crash**	**Control**	**Total**	**Crash**	**Control**	**Total**
Agreed	**Count**	1954	4861	6815	2032	4992	7024	3986	9853	13839
	% within Crash/Control Status	80.7	97.1	91.7	81.4	98.7	92.9	81.0	97.9	92.4
	% of Total	26.3	65.4	91.7	26.9	66.1	92.9	26.6	65.8	92.4
Did not Agree	**Count**	100	145	245	179	68	247	279	213	492
	% within Crash/Control Status	4.1	2.9	3.3	7.2	1.3	3.3	5.7	2.1	3.3
	% of Total	1.3	2.0	3.3	2.4	.9	3.3	1.9	1.4	3.3
Hit-and-Run Recovered	**Count**	77		77	27		27	104		104
	% within Crash/Control Status	3.2		1.0	1.1		.4	2.1		.7
	% of Total	1.0		1.0	.4		.4	.7		.7
Hit-and-run not Recovered	**Count**	291		291	208		208	499		499
	% within Crash/Control Status	12.0		3.9	8.3		2.8	10.1		3.3
	% of Total	3.9		3.9	2.8		2.8	3.3		3.3
Unable to Participate	**Count**				51		51	51		51
	% within Crash/Control Status				2.0		.7	1.0		.3
	% of Total				.7		.7	.3		.3
Total	**Count**	2422	5006	7428	2497	5060	7473	4919	10066	14985
	% of Total	32.6%	67.4%	100%	33.0%	67.0%	100%	32.8%	67.2%	100%

In total, 4,090, or 83.1% of 4,919 eligible drivers participated (3,986 at the crashes plus 104 recovered hit-and-run drivers). As will be discussed later, 81.3% provided breath specimens.

1.3 Time Variables

Table F-3 presents the numbers of sampled crash and control drivers by month. Control drivers account for 67.2% of the total; i.e., two controls for each crash-involved driver. Bearing in mind that data acquisition in Long Beach extended over two years for the months of June, July and August, the sampling across months was relatively constant. Somewhat higher totals occurred in Fort Lauderdale in April after the corps of field interviewers was augmented by three thereby making it possible to staff more shifts.

The numbers of sampled crash and control drivers by day of the week, as shown in Table F-4, reflect the sampling schedules (Tables 2 and 3 in the body of the report). The numbers increased on the days when three teams, instead of two, were in the field. An extra team in Long Beach increased the number of crashes and controls sampled, although not by 50%. An extra team in Fort Lauderdale increased the number of drivers sampled on Saturday by 150% compared to a typical weekday, produced only a marginal increase on Fridays, and produced no change on Wednesdays.

Table F-5 presents the numbers of crashes sampled by hour of the day. Sampling began at 1600 and ended at 0200 in Long Beach. The Fort Lauderdale teams began and ended an hour later. In Long Beach, 52.4% of the crashes occurred before 1900 whereas 55.4% of Fort Lauderdale crashes occurred after 2000. The congestion of the evening commute may have contributed to the early crashes in Long Beach. It is also possible that the team's duties as first-responders delayed their responses to additional crashes

1.4 Vehicle Type

As can be seen in Table F-6[1], most of the crash and control vehicles at both sites were 2-door and 4-door sedans. Fort Lauderdale vehicles included more sports cars, convertibles and other types, whereas Long Beach vehicles included more pickup trucks, vans, minivans and station wagons.

1.5 Passengers

Between 55% and 60% of vehicles at both sites were occupied by the driver only (Table F-7). About 25% had a single passenger, and about 8% had two passengers.

1.6 Weekly Mileage

Drivers were asked to estimate their typical weekly mileage (Table F-8). Just over 60% of all drivers at both sites estimated their weekly mileage as less than 200 miles per week. Two hundred miles per week was the mean, and 100 miles per week was the median. More than 20% of the respondents, however, reported that they drove 300 or more miles in a typical week.

---

[1] All tables presented to this point contain the full sample of 14,985 drivers or 4,919 crashes. No data are missing since they are based on team observations or information that could be accurately acquired for all cases. In Table 9 and most of the following tables, some data is missing because it was unavailable or the subject driver refused to provide it. Since the purpose of these tables is to describe the known sample, the tabulation of missing drivers is not shown. The reader can readily assess the extent of missing data by comparison with the site and study marginals in any of the preceding tables.

Table F-3. Crash and Control Sampling by Site and Month

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Month	**January**	**Count**	139	283	422	208	418	626	347	701	1048
		% within Crash/Control Status	5.7%	5.7%	5.7%	8.3%	8.3%	8.3%	7.1%	7.0%	7.0%
		% of Total	1.9%	3.8%	5.7%	2.8%	5.5%	8.3%	2.3%	4.7%	7.0%
	February	**Count**	159	322	481	162	328	490	321	650	971
		% within Crash/Control Status	6.6%	6.4%	6.5%	6.5%	6.5%	6.5%	6.5%	6.5%	6.5%
		% of Total	2.1%	4.3%	6.5%	2.1%	4.3%	6.5%	2.1%	4.3%	6.5%
	March	**Count**	206	426	632	218	440	658	424	866	1290
		% within Crash/Control Status	8.5%	8.5%	8.5%	8.7%	8.7%	8.7%	8.6%	8.6%	8.6%
		% of Total	2.8%	5.7%	8.5%	2.9%	5.8%	8.7%	2.8%	5.8%	8.6%
	April	**Count**	199	408	607	253	515	768	452	923	1375
		% within Crash/Control Status	8.2%	8.2%	8.2%	10.1%	10.2%	10.2%	9.2%	9.2%	9.2%
		% of Total	2.7%	5.5%	8.2%	3.3%	6.8%	10.2%	3.0%	6.2%	9.2%
	May	**Count**	168	357	525	203	411	614	371	768	1139
		% within Crash/Control Status	6.9%	7.1%	7.1%	8.1%	8.1%	8.1%	7.5%	7.6%	7.6%
		% of Total	2.3%	4.8%	7.1%	2.7%	5.4%	8.1%	2.5%	5.1%	7.6%
	June	**Count**	241	510	751	221	449	670	462	959	1421
		% within Crash/Control Status	10.0%	10.2%	10.1%	8.9%	8.9%	8.9%	9.4%	9.5%	9.5%
		% of Total	3.2%	6.9%	10.1%	2.9%	5.9%	8.9%	3.1%	6.4%	9.5%
	July	**Count**	344	718	1062	179	367	546	523	1085	1608
		% within Crash/Control Status	14.2%	14.3%	14.3%	7.2%	7.3%	7.2%	10.6%	10.8%	10.7%
		% of Total	4.6%	9.7%	14.3%	2.4%	4.9%	7.2%	3.5%	7.2%	10.7%
	August	**Count**	335	693	1028	179	363	542	514	1056	1570
		% within Crash/Control Status	13.8%	13.8%	13.8%	7.2%	7.2%	7.2%	10.4%	10.5%	10.5%
		% of Total	4.5%	9.3%	13.8%	2.4%	4.8%	7.2%	3.4%	7.0%	10.5%
	September	**Count**	153	319	472	194	393	587	347	712	1059
		% within Crash/Control Status	6.3%	6.4%	6.4%	7.8%	7.8%	7.8%	7.1%	7.1%	7.1%
		% of Total	2.1%	4.3%	6.4%	2.6%	5.2%	7.8%	2.3%	4.8%	7.1%
	October	**Count**	151	309	460	238	477	715	389	786	1175
		% within Crash/Control Status	6.2%	6.2%	6.2%	9.5%	9.4%	9.5%	7.9%	7.8%	7.8%
		% of Total	2.0%	4.2%	6.2%	3.1%	6.3%	9.5%	2.6%	5.2%	7.8%
	November	**Count**	166	333	499	223	454	677	389	787	1176
		% within Crash/Control Status	6.9%	6.7%	6.7%	8.9%	9.0%	9.0%	7.9%	7.8%	7.8%
		% of Total	2.2%	4.5%	6.7%	3.0%	6.0%	9.0%	2.6%	5.3%	7.8%
	December	**Count**	161	328	489	219	445	664	380	773	1153
		% within Crash/Control Status	6.6%	6.6%	6.6%	8.8%	8.8%	8.8%	7.7%	7.7%	7.7%
		% of Total	2.2%	4.4%	6.6%	2.9%	5.9%	8.8%	2.5%	5.2%	7.7%
	Total	**Count**	2422	5006	7428	2497	5060	7557	4919	10066	14985
		% of Site/Total	32.6%	70.6%	100%	33.0%	67.0%	100%	32.8%	67.2%	100%

Table F-4. Crash and Control Sampling by Site and Day of Week

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Day of the Week	**Sunday**	**Count**	315	642	957	356	719	1075	671	1361	2032
		% within Crash/Control Status	13.0%	12.8%	12.9%	14.3%	14.2%	14.2%	13.6%	13.5%	13.6%
		% of Total	4.2%	8.6%	12.9%	4.7%	9.5%	14.2%	4.5%	9.1%	13.6%
	Monday	**Count**	271	563	834	275	556	831	546	1119	1665
		% within Crash/Control Status	11.2%	11.2%	11.2%	11.0%	11.0%	11.0%	11.1%	11.1%	11.1%
		% of Total	3.6%	7.6%	11.2%	3.6%	7.4%	11.0%	3.6%	7.5%	11.1%
	Tuesday	**Count**	261	549	810	325	652	977	586	1201	1787
		% within Crash/Control Status	10.8%	11.0%	10.9%	13.0%	12.9%	12.9%	11.9%	11.9%	11.9%
		% of Total	3.5%	7.4%	10.9%	4.3%	8.6%	12.9%	3.9%	8.0%	11.9%
	Wednesday	**Count**	299	624	923	327	659	986	626	1283	1909
		% within Crash/Control Status	12.3%	12.5%	12.4%	13.1%	13.0%	13.0%	12.7%	12.7%	12.7%
		% of Total	4.0%	8.4%	12.4%	4.3%	8.7%	13.0%	4.2%	8.6%	12.7%
	Thursday	**Count**	393	806	1199	326	664	990	719	1470	2189
		% within Crash/Control Status	16.2%	16.1%	16.1%	13.1%	13.1%	13.1%	14.6%	14.6%	14.6%
		% of Total	5.3%	10.9%	16.1%	4.3%	8.8%	13.1%	4.8%	9.8%	14.6%
	Friday	**Count**	481	996	1477	388	794	1182	869	1790	2659
		% within Crash/Control Status	19.9%	19.9%	19.9%	15.5%	15.7%	15.6%	17.7%	17.8%	17.7%
		% of Total	6.5%	13.4%	19.9%	5.1%	10.5%	15.6%	5.8%	11.9%	17.7%
	Saturday	**Count**	402	826	1228	500	1016	1516	902	1842	2744
		% within Crash/Control Status	16.6%	16.5%	16.5%	20.0%	20.1%	20.1%	18.3%	18.3%	18.3%
		% of Total	5.4%	11.1%	16.5%	6.6%	13.4%	20.1%	6.0%	12.3%	18.3%
	Total	**Count**	2422	5006	7428	2497	5060	7557	4919	10066	14985
		% of Total	32.6%	67.4%	100%	33.0%	67.0%	100%	32.8%	67.2%	100%

Table F-5. Police Reported Times of Crashes by Site

			Long Beach	Fort Lauderdale	Total
Estimated Time of Crash	1500 - 1559	Count	2		2
		% within Location	.1%		0.0%
	1600 - 1659	Count	379		379
		% within Location	15.6%		7.7%
	1700 - 1759	Count	534	341	875
		% within Location	22.0%	13.7%	17.8%
	1800 - 1859	Count	423	435	858
		% within Location	17.5%	17.4%	17.4%
	1900 - 1959	Count	245	353	598
		% within Location	10.1%	14.1%	12.2%
	2000 - 2059	Count	247	236	483
		% within Location	10.2%	9.5%	9.8%
	2100 - 2159	Count	242	281	523
		% within Location	10.0%	11.3%	10.6%
	2200 - 2259	Count	142	272	414
		% within Location	5.9%	10.9%	8.4%
	2300 - 2359	Count	111	265	376
		% within Location	4.6%	10.6%	7.6%
	0000 - 0059	Count	67	183	250
		% within Location	2.8%	7.3%	5.1%
	0100 - 0159	Count	30	110	140
		% within Location	1.2%	4.4%	2.8%
	0200 - 0259	Count		21	21
		% within Location		.8%	.4%
	Total	Count	2422	2497	4919

1.7 Trip Origin

Drivers were asked where their trips had originated (Table F-9). Over 40% said they were coming from their own or someone else's home, and relatively few said they were coming from an alcohol outlet. Fort Lauderdale drivers said they were coming from a restaurant almost twice as often as Long Beach drivers. Trip origin did not differ between crash and control drivers.

1.8 Age, Marital Status, Gender and Ethnicity

Respondents' ages and marital status, as determined by responses to questionnaire items, are in Tables F-10 and F-11. Gender (Table F-12) was observed and recorded by the research team. More crash than control drivers were under age 25 and over age 65. This overrepresentation in the crash group of youthful and older drivers is consistent with other studies of drivers' ages.

Table F-6. Vehicle Types Driven by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Vehicle Type	**2-door**	Count	704	1792	2496	519	1178	1697	1223	2970	4193
		% within Crash/Control Status	30.6%	35.8%	34.2%	21.5%	23.3%	22.7%	26.0%	29.5%	28.4%
		% of Total	9.6%	24.6%	34.2%	6.9%	15.8%	22.7%	8.3%	20.1%	28.4%
	4-door	Count	864	1779	2643	1109	2273	3382	1973	4052	6025
		% within Crash/Control Status	37.6%	35.6%	36.2%	45.9%	44.9%	45.3%	41.9%	40.3%	40.8%
		% of Total	11.8%	24.4%	36.2%	14.8%	30.4%	45.3%	13.4%	27.4%	40.8%
	Sports Car	Count	44	75	119	100	223	323	144	298	442
		% within Crash/Control Status	1.9%	1.5%	1.6%	4.1%	4.4%	4.3%	3.1%	3.0%	3.0%
		% of Total	.6%	1.0%	1.6%	1.3%	3.0%	4.3%	1.0%	2.0%	3.0%
	Other	Count	74	142	216	203	415	618	277	557	834
		% within Crash/Control Status	3.2%	2.8%	3.0%	8.4%	8.2%	8.3%	5.9%	5.5%	5.6%
		% of Total	1.0%	1.9%	3.0%	2.7%	5.6%	8.3%	1.9%	3.8%	5.6%
	Pickup	Count	320	591	911	238	458	696	558	1049	1607
		% within Crash/Control Status	13.9%	11.8%	12.5%	9.9%	9.1%	9.3%	11.8%	10.4%	10.9%
		% of Total	4.4%	8.1%	12.5%	3.2%	6.1%	9.3%	3.8%	7.1%	10.9%
	Convertible	Count	4	13	17	42	127	169	46	140	186
		% within Crash/Control Status	.2%	.3%	.2%	1.7%	2.5%	2.3%	1.0%	1.4%	1.3%
		% of Total	.1%	.2%	.2%	.6%	1.7%	2.3%	.3%	.9%	1.3%
	Van/Minivan	Count	216	399	615	130	262	392	346	661	1007
		% within Crash/Control Status	9.4%	8.0%	8.4%	5.4%	5.2%	5.2%	7.3%	6.6%	6.8%
		% of Total	3.0%	5.5%	8.4%	1.7%	3.5%	5.2%	2.3%	4.5%	6.8%
	Station Wagon	Count	72	210	282	73	123	196	145	333	478
		% within Crash/Control Status	3.1%	4.2%	3.9%	3.0%	2.4%	2.6%	3.1%	3.3%	3.2%
		% of Total	1.0%	2.9%	3.9%	3.5%	1.6%	2.6%	1.0%	2.3%	3.2%
	Total	Count	2298	5001	7299	2414	5059	7473	4712	10060	14772
		% of Total	31.5%	68.5%	100%	32.3%	67.7%	100%	31.9%	68.1%	100%

Table F-7. Number of Passengers with Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	**Control**	**Total**	**Crash**	**Control**	**Total**	**Crash**	**Control**	**Total**
Number of Passengers	**0**	**Count**	1338	3001	4339	1337	2823	4160	2675	5824	8499
		% within Crash /Control Status	60.5%	60.0%	60.2%	57.7%	55.8%	56.4%	59.1%	57.9%	58.3%
		% of Total	18.6%	41.6%	60.2%	18.1%	38.3%	56.4%	18.3%	39.9%	58.3%
	1	**Count**	521	1315	1836	645	1565	2210	1166	2880	4046
		% within Crash /Control Status	23.6%	26.3%	25.5%	27.8%	30.9%	30.0%	25.7%	28.6%	27.7%
		% of Total	7.2%	18.2%	25.5%	8.7%	21.2%	30.0%	8.0%	19.7%	27.7%
	2	**Count**	194	342	536	190	396	586	384	738	1122
		% within Crash /Control Status	8.8%	6.8%	7.4%	8.2%	7.8%	7.9%	8.5%	7.3%	7.7%
		% of Total	2.7%	4.7%	7.4%	2.6%	5.4%	7.9%	2.6%	5.1%	7.7%
	3	**Count**	87	232	319	103	183	286	190	415	605
		% within Crash /Control Status	3.9%	4.6%	4.4%	4.4%	3.6%	3.9%	4.2%	4.1%	4.1%
		% of Total	1.2%	3.2%	4.4%	1.4%	2.5%	3.9%	1.3%	2.8%	4.1%
	4	**Count**	45	67	112	24	64	88	69	131	200
		% within Crash /Control Status	2.0%	1.3%	1.6%	1.0%	1.3%	1.2%	1.5%	1.3%	1.4%
		% of Total	.6%	.9%	1.6%	.3%	.9%	1.2%	.5%	.9%	1.4%
	5	**Count**	12	28	40	13	18	31	25	46	71
		% within Crash /Control Status	.5%	.6%	.6%	.6%	.4%	.4%	.6%	.5%	.5%
		% of Total	.2%	.4%	.6%	.2%	.2%	.4%	.2%	.3%	.5%
	6	**Count**	9	8	17	6	7	13	15	15	30
		% within Crash /Control Status	.4%	.2%	.2%	.3%	.1%	.2%	.3%	.3%	.2%
		% of Total	.1%	.1%	.2%	.1%	.1%	.2%	.1%	.1%	.2%
	7	**Count**	1	2	3	1	1	2	2	3	5
		% within Crash /Control Status	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.0%
		% of Total	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.0%	.0%
	8	**Count**		4	4		1	1		5	5
		% within Crash /Control Status		.1%	.1%		.0%	.0%		.0%	.0%
		% of Total		.1%	.1%		.0%	.0%		.0%	.0%
	9	**Count**	3	1	4				3	1	4
		% within Crash /Control Status	.1%	.0%	.1%				.1%	.0%	.0%
		% of Total	.0%	.0%	.1%				.0%	.0%	.0%
	Total	**Count**	2210	5000	7210	2319	5058	7377	4529	10058	14587
		% of Total	30.7%	69.3%	100%	31.4%	68.6%	100%	31.0%	69.0%	100%

Table F-8. Weekly Mileage Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Typical Weekly Mileage	0	Count	33	1	34	27	26	53	60	27	87
		% within Crash/Control Status	1.6%	.0%	.5%	1.3%	.5%	.8%	1.5%	.3%	.6%
		% of Total	.5%	.0%	.5%	.4%	.4%	.8%	.4%	.2%	.6%
	1-49	Count	389	896	1285	364	870	1234	753	1766	2519
		% within Crash/Control Status	19.4%	18.5%	18.7%	17.9%	17.4%	17.5%	18.7%	17.9%	18.1%
		% of Total	5.7%	13.1%	18.7%	5.2%	12.4%	17.5%	5.4%	12.7%	18.1%
	50-99	Count	448	928	1376	359	944	1303	807	1872	2679
		% within Crash/Control Status	22.3%	19.1%	20.1%	17.7%	18.9%	18.5%	20.0%	19.0%	19.3%
		% of Total	6.5%	13.5%	20.1%	5.1%	13.4%	18.5%	5.8%	13.5%	19.3%
	100-199	Count	459	1180	1639	480	1181	1661	939	2361	3300
		% within Crash/Control Status	22.9%	24.3%	23.9%	23.6%	23.6%	23.6%	23.3%	24.0%	23.8%
		% of Total	6.7%	17.2%	23.9%	16.8%	16.8%	23.6%	6.8%	17.0%	23.8%
	200-299	Count	302	776	1078	341	844	1185	643	1620	2263
		% within Crash/Control Status	15.1%	16.0%	15.7%	16.8%	16.9%	16.8%	15.9%	16.4%	16.3%
		% of Total	4.4%	11.3%	15.7%	4.8%	12.0%	16.8%	4.6%	11.7%	16.3%
	300-499	Count	227	613	840	299	618	917	526	1231	1757
		% within Crash/Control Status	11.3%	12.6%	12.2%	14.7%	12.4%	13.0%	13.0%	12.5%	12.6%
		% of Total	3.3%	8.9%	12.2%	4.3%	8.8%	13.0%	3.8%	8.9%	12.6%
	500-749	Count	107	311	418	95	329	424	202	640	842
		% within Crash/Control Status	5.3%	6.4%	6.1%	4.7%	6.6%	6.0%	5.0%	6.5%	6.1%
		% of Total	1.6%	4.5%	6.1%	1.4%	4.7%	6.0%	1.5%	4.6%	6.1%
	750+	Count	41	150	191	65	191	256	106	341	447
		% within Crash/Control Status	2.0%	3.1%	2.8%	3.2%	3.8%	3.6%	2.6%	3.5%	3.2%
		% of Total	.6%	2.2%	2.8%	.9%	2.7%	3.6%	.8%	2.5%	3.2%
	Total	Count	2006	4885	6861	2030	5003	7033	4036	9858	13894
		% of Total	29.2%	70.8%	100%	28.9%	71.1%	100%	29.0%	71.0%	100%

Table F-9. Trip Origins Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Trip Origin	Subject's Home	Count	551	1550	2101	478	1278	1756	1029	2828	3857
		% within Crash/Control Status	27.3%	32.0%	30.6%	23.2%	25.5%	24.8%	25.3%	28.7%	27.7%
		% of Total	8.0%	22.6%	30.6%	6.8%	18.1%	24.8%	7.4%	20.3%	27.7%
	Other's Home	Count	375	710	1085	320	698	1018	695	1408	2103
		% within Crash/Control Status	18.6%	14.7%	15.8%	15.6%	13.9%	14.4%	17.1%	14.3%	15.1%
		% of Total	5.5%	10.4%	15.8%	4.5%	9.9%	14.4%	5.0%	10.1%	15.1%
	Liquor Store	Count	10	17	27		7	7	10	24	34
		% within Crash/Control Status	.5%	.4%	.4%		.1%	.1%	.2%	.2%	.2%
		% of Total	.1%	.2%	.4%		.1%	.1%	.1%	.2%	.2%
	Other Business	Count	221	483	704	113	374	487	334	857	1191
		% within Crash/Control Status	11.0%	10.0%	10.3%	5.5%	7.5%	6.9%	8.2%	8.7%	8.6%
		% of Total	3.2%	7.0%	10.3%	1.6%	5.3%	6.9%	2.4%	6.2%	8.6%
	Bar/ Tavern	Count	41	25	66	46	92	138	87	117	204
		% within Crash/Control Status	2.0%	.5%	1.0%	2.2%	1.8%	2.0%	2.1%	1.2%	1.5%
		% of Total	.6%	.4%	1.0%	.7%	1.3%	2.0%	.6%	.8%	1.5%
	Restaurant	Count	57	249	306	187	509	696	244	758	1002
		% within Crash/Control Status	2.8%	5.1%	4.5%	9.1%	10.2%	9.8%	6.0%	7.7%	7.2%
		% of Total	.8%	3.6%	4.5%	2.6%	7.2%	9.8%	1.8%	5.4%	7.2%
	Work	Count	448	1203	1651	440	1057	1497	888	2260	3148
		% within Crash/Control Status	22.2%	24.8%	24.1%	21.4%	21.1%	21.2%	21.8%	22.9%	22.6%
		% of Total	6.5%	17.5%	24.1%	6.2%	15.0%	21.2%	6.4%	16.2%	22.6%
	Other	Count	313	606	919	472	996	1468	785	1602	2387
		% within Crash/Control Status	15.5%	12.5%	13.4%	23.0%	19.9%	20.8%	19.3%	16.2%	17.2%
		% of Total	4.6%	8.8%	13.4%	6.7%	14.1%	20.8%	5.5%	11.0%	17.2%
	Total	Count	2016	4843	6859	2056	5011	7067	4072	9854	13926
		% of Total	29.4%	70.6%	100%	29.1%	70.9%	100%	29.2%	70.8%	100%

Table F-10. Driver Ages Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Age	**Under 21**	**Count**	240	392	632	243	347	590	483	739	1222
		% within Crash/Control Status	11.9%	8.1%	9.2%	11.8%	7.0%	8.4%	11.9%	7.5%	8.8%
		% of Total	3.5%	5.7%	9.2%	3.4%	4.9%	8.4%	3.5%	5.3%	8.8%
	21-24	**Count**	255	515	770	216	495	711	471	1010	1481
		% within Crash/Control Status	12.6%	10.6%	11.2%	10.5%	9.9%	10.1%	11.6%	10.3%	10.6%
		% of Total	3.7%	7.5%	11.2%	3.1%	7.0%	10.1%	3.4%	7.3%	10.6%
	25-34	**Count**	569	1591	2160	587	1445	2032	1156	3036	4192
		% within Crash/Control Status	28.2%	32.8%	31.5%	28.6%	29.0%	28.8%	28.4%	30.9%	30.1%
		% of Total	8.3%	23.2%	31.5%	8.3%	20.5%	28.8%	8.3%	21.8%	30.1%
	35-44	**Count**	459	1198	1657	468	1290	1758	927	2488	3415
		% within Crash/Control Status	22.7%	24.7%	24.1%	22.8%	25.8%	24.9%	22.8%	25.3%	24.6%
		% of Total	6.7%	17.5%	24.1%	6.6%	18.3%	24.9%	6.7%	17.9%	24.6%
	45-54	**Count**	246	684	930	283	809	1092	529	1493	2022
		% within Crash/Control Status	12.2%	14.1%	13.6%	13.8%	16.2%	15.5%	13.0%	15.2%	14.5%
		% of Total	3.6%	10.0%	13.6%	4.0%	11.5%	15.5%	3.8%	10.7%	14.5%
	55-64	**Count**	127	295	422	147	391	538	274	686	960
		% within Crash/Control Status	6.3%	6.1%	6.1%	7.1%	7.8%	7.6%	6.7%	7.0%	6.9%
		% of Total	1.9%	4.3%	6.1%	2.1%	5.5%	7.6%	2.0%	4.9%	6.9%
	65+	**Count**	122	170	292	112	214	326	234	384	618
		% within Crash/Control Status	6.0%	3.5%	4.3%	5.4%	4.3%	4.6%	5.7%	3.9%	4.4%
		% of Total	1.8%	2.5%	4.3%	1.6%	3.0%	4.6%	1.7%	2.8%	4.4%
	Total	**Count**	2018	4845	6863	2056	4991	7047	4074	9836	13910
		% of Total	29.4%	70.6%	100%	29.2%	70.8%	100%	29.3%	70.7%	100%

Table F-11. Marital Status Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	**Control**	**Total**	**Crash**	**Control**	**Total**	**Crash**	**Control**	**Total**
Marital Status	**Single**	**Count**	1001	2228	3229	1115	2532	3647	2116	4760	6876
		% within Crash/Control Status	49.6%	45.9%	47.0%	54.3%	50.5%	51.6%	52.0%	48.2%	49.3%
		% of Total	14.6%	32.4%	47.0%	15.8%	35.8%	51.6%	15.2%	34.1%	49.3%
	Married	**Count**	688	1963	2651	643	1687	2330	1331	3650	4981
		% within Crash/Control Status	34.1%	40.4%	38.6%	31.3%	33.6%	33.0%	32.7%	37.0%	35.7%
		% of Total	10.0%	28.6%	38.6%	9.1%	23.9%	33.0%	9.5%	26.2%	35.7%
	Living Together	**Count**	67	192	259	55	168	223	122	360	482
		% within Crash/Control Status	3.3%	4.0%	3.8%	2.7%	3.4%	3.2%	3.0%	3.6%	3.5%
		% of Total	1.0%	2.8%	3.8%	.8%	2.4%	3.2%	.9%	2.6%	3.5%
	Divorced	**Count**	160	294	454	153	425	578	313	719	1032
		% within Crash/Control Status	7.9%	6.1%	6.6%	7.4%	8.5%	8.2%	7.7%	7.3%	7.4%
		% of Total	2.3%	4.3%	6.6%	2.2%	6.0%	8.2%	2.2%	5.2%	7.4%
	Separated	**Count**	48	113	161	41	110	151	89	223	312
		% within Crash/Control Status	2.4%	2.3%	2.3%	2.0%	2.2%	2.1%	2.2%	2.3%	2.2%
		% of Total	.7%	1.6%	2.3%	.6%	1.6%	2.1%	.6%	1.6%	2.2%
	Widowed	**Count**	54	65	119	47	92	139	101	157	258
		% within Crash/Control Status	2.7%	1.3%	1.7%	2.3%	1.8%	2.0%	2.5%	1.6%	1.9%
		% of Total	.8%	.9%	1.7%	.7%	1.3%	2.0%	.7%	1.1%	1.9%
	Total	**Count**	2018	4855	6873	2054	5014	7068	4072	9869	13941
		% of Total	29.4%	70.6%	100%	29.1%	70.9%	100%	29.2%	70.8%	100%

Table F-12. Driver Gender by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Gender	**Male**	**Count**	1348	3198	4546	1598	3335	4933	2946	6533	9479
		% within Crash/Control Status	59.3%	63.9%	62.5%	66.8%	65.9%	66.2%	63.1%	64.9%	64.4%
		% of Total	18.5%	43.9%	62.5%	21.4%	44.8%	66.2%	20.0%	44.4%	64.4%
	Female	**Count**	926	1805	2731	794	1723	2517	1720	3528	5248
		% within Crash/Control Status	40.7%	36.1%	37.5%	33.2%	34.1%	33.8%	36.9%	35.1%	35.6%
		% of Total	12.7%	24.8%	37.5%	10.7%	23.1%	33.8%	11.7%	24.0%	35.6%
	Total	**Count**	2274	5003	7277	2392	5058	7450	4666	10061	14727
		% of Total	31.2%	68.8%	100%	32.1%	67.9%	100%	31.7%	68.3%	100%

Drivers with single marital status are over-represented in the total crash sample, 52.0% to 48.2%. More Fort Lauderdale drivers reported that they were single (51.6%) and divorced (8.2%) than Long Beach drivers (47.0% single, 6.6% divorced). Widowed drivers are over-represented in the crash group from both sites.

The crash and control drivers differed only slightly in terms of gender. For the total sample, males are slightly underrepresented and females are slightly over-represented in crashes. Males are underrepresented in the Long Beach crash group, 59.3% to 63.9%, but they are slightly over-represented in the Fort Lauderdale crash group, 66.8% to 65.9%.

Drivers classified their ethnicity by choosing from the following categories on a response card: African-American, Asian, Caucasian, Hispanic/Latino, Native American, Pacific Islander, Multi-ethnic or other (Table F-13). Although the proportion of African-Americans among all drivers was higher in Fort Lauderdale (23.8%) than in Long Beach (6.0%), they were underrepresented in the Fort Lauderdale crash group and over-represented in the Long Beach crash group (17.3% to 15.5%). The proportions of Asians (10.5%), Hispanics (27.5%) and Pacific Islanders (2.2%) were higher in Long Beach than in Fort Lauderdale data where there were 1.1% Asians, 10.4% Hispanics, and 0.2% Pacific Islanders. The non-specific *other* category was selected by 7.7% of the Fort Lauderdale and 2.2% of the Long Beach sample, and was over-represented in the crash group at both sites.

1.9 Education

A single question assessed education by asking how many a driver how many years of school he/she had completed (Table F-14). About a quarter of all drivers reported 12 years of schooling, and another 30% reported more than 12 but less than 16 years. A slightly higher proportion in Fort Lauderdale compared to Long Beach reported 16 years, which is assumed to mean college graduation. Only 4.0% of the drivers reported 19 or more years of education.

Table F-13. Ethnic Group Reported by Crash and Control Drivers by Site

Reported Ethnic Group		Long Beach			Fort Lauderdale			Total		
		Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
African-American	Count	349	752	1101	428	1256	1684	777	2008	2785
	% within Crash/Control Status	17.3%	15.5%	16.0%	20.7%	25.1%	23.8%	19.1%	20.4%	20.0%
	% of Total	5.1%	11.0%	16.0%	6.1%	17.8%	23.8%	5.6%	14.4%	20.0%
Asian	Count	218	502	720	31	45	76	249	547	796
	% within Crash/Control Status	10.8%	10.4%	10.5%	1.5%	.9%	1.1%	6.1%	5.5%	5.7%
	% of Total	3.2%	7.3%	10.5%	.4%	.6%	1.1%	1.8%	3.9%	5.7%
Caucasian	Count	826	1902	2728	1121	2744	3865	1947	4646	6593
	% within Crash/Control Status	41.0%	39.2%	39.8%	54.3%	54.8%	54.6%	47.8%	47.1%	47.3%
	% of Total	12.0%	27.7%	39.8%	15.8%	38.8%	54.6%	14.0%	33.3%	47.3%
Hispanic/ Latino	Count	483	1404	1887	217	522	739	700	1926	2626
	% within Crash/Control Status	24.0%	29.0%	27.5%	10.5%	10.4%	10.4%	17.2%	19.5%	18.8%
	% of Total	7.0%	20.5%	27.5%	3.1%	7.4%	10.4%	5.0%	13.8%	18.8%
Native American	Count	9	26	35	15	21	36	24	47	71
	% within Crash/Control Status	.4%	.5%	.5%	.7%	.4%	.5%	.6%	.5%	.5%
	% of Total	.1%	.4%	.5%	.2%	.3%	.5%	.2%	.3%	.5%
Pacific Islander	Count	41	112	153	4	11	15	45	123	168
	% within Crash/Control Status	2.0%	2.3%	2.2%	.2%	.2%	.2%	1.1%	1.2%	1.2%
	% of Total	.6%	1.6%	2.2%	.1%	.2%	.2%	.3%	.9%	1.2%
Multi-Ethnic	Count	26	61	87	40	73	113	66	134	200
	% within Crash/Control Status	1.3%	1.3%	1.3%	1.9%	1.5%	1.6%	1.6%	1.4%	1.4%
	% of Total	.4%	.9%	1.3%	.6%	1.0%	1.6%	.5%	1.0%	1.4%
Other	Count	62	89	151	207	338	545	269	427	696
	% within Crash/Control Status	3.1%	1.8%	2.2%	10.0%	6.7%	7.7%	6.6%	4.3%	5.0%
	% of Total	.9%	1.3%	2.2%	2.9%	4.8%	7.7%	1.9%	3.1%	5.0%
Total	Count	2014	4848	6862	2063	5010	7073	4077	9858	13935
	% of Total	29.4%	70.6%	100%	29.2%	70.8%	100%	29.3%	70.7%	100%

Table F-14. Years of School Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Years of School	0-9	Count	222	476	698	133	210	343	355	686	1041
		% within Crash/Control Status	11.0%	9.8%	10.2%	6.5%	4.2%	4.9%	8.8%	7.0%	7.5%
		% of Total	3.2%	6.9%	10.2%	1.9%	3.0%	4.9%	2.6%	4.9%	7.5%
	9-11	Count	185	299	484	223	431	654	408	730	1138
		% within Crash/Control Status	9.2%	6.2%	7.1%	10.9%	8.6%	9.3%	10.1%	7.4%	8.2%
		% of Total	2.7%	4.4%	7.1%	3.2%	6.1%	9.3%	2.9%	5.2%	8.2%
	12 High School	Count	535	1229	1764	578	1390	1968	1113	2619	3732
		% within Crash/Control Status	26.6%	25.3%	25.7%	28.2%	27.8%	27.9%	27.4%	26.6%	26.8%
		% of Total	7.8%	17.9%	25.7%	8.2%	19.7%	27.9%	8.0%	18.8%	26.8%
	13-15	Count	585	1498	2083	556	1406	1962	1141	2904	4045
		% within Crash/Control Status	29.1%	30.9%	30.4%	27.2%	28.1%	27.8%	28.1%	29.5%	29.1%
		% of Total	8.5%	21.8%	30.4%	7.9%	19.9%	27.8%	8.2%	20.9%	29.1%
	16 Bachelors	Count	265	758	1023	317	944	1261	582	1702	2284
		% within Crash/Control Status	13.2%	15.6%	14.9%	15.5%	18.9%	17.9%	14.3%	17.3%	16.4%
		% of Total	3.9%	11.0%	14.9%	4.5%	13.4%	17.9%	4.2%	12.2%	16.4%
	17-18 Masters	Count	144	419	563	154	401	555	298	820	1118
		% within Crash/Control Status	7.2%	8.6%	8.2%	7.5%	8.0%	7.9%	7.3%	8.3%	8.0%
		% of Total	2.1%	6.1%	8.2%	2.2%	5.7%	7.9%	2.1%	5.9%	8.0%
	19+ Doctorate	Count	74	171	245	86	221	307	160	392	552
		% within Crash/Control Status	3.7%	3.5%	3.6%	4.2%	4.4%	4.4%	3.9%	4.0%	4.0%
		% of Total	1.1%	2.5%	3.6%	1.2%	3.1%	4.4%	1.2%	2.8%	4.0%
	Total	Count	2010	4850	6860	2047	5003	7050	4057	9853	13910
		% of Total	29.3%	70.7%	100%	29.0%	71.0%	100%	29.2%	70.8%	100%

Drivers with 16 or more years of education were under-represented in the crash groups. In Long Beach 10.2% reported zero to nine years of education, compared to 4.9% in Lauderdale 4.9%. These individuals and those with nine to 11 years of education were over-represented in the crash group at both sites.

1.10 Employment

Questions about drivers' work hours and their type of work did not yield reliable data due to ambiguities in work descriptions and coding discrepancies. The data are not tabled. Another question asked them to classify their employment status as one of eight categories (Table F-15). Full-time employees are underrepresented and other categories are over-represented in the Long Beach crash group. The Fort Lauderdale data exhibit no large over- or under-representations.

1.11 Sleep

Because the effects of lack of sleep could add to the effects of alcohol, the questionnaire contained several items about sleep. The number of hours last slept yielded an unexpected finding (Table F-16). Drivers with seven or fewer hours of sleep were under-represented among crash drivers. This finding occurred at both sites.

1.12 Recent Alcohol Consumption

Interviewers asked drivers about their alcohol consumption during two periods, two hours prior and 30 minutes prior to the interview (Tables F-17, F-18). In Fort Lauderdale, 15.6% of all drivers indicated they had consumed alcohol within two hours. In Long Beach, 11.6% of the crash-involved drivers and 6.3% of the controls (7.8% of all drivers) admitted this drinking behavior.

Measured BACs were not always explained by the amount of alcohol that drivers reported drinking within two hours. Some of the crash-involved drivers showed a positive BAC but claimed to have had no alcohol within two hours (5.3% in Long Beach, 8.5% in Fort Lauderdale), and others showed a zero BAC but said they had drunk alcohol within two hours (1.8% in Long Beach, 3.4% in Fort Lauderdale). The discrepancies between measured BACs and reported drinking were smaller among control drivers. A small amount of alcohol could have been metabolized within two hours accounting for the zero BACs. Also, a large amount of alcohol could yield a positive BAC even though consumption ceased more than two hours previously

Responses to a question about drinking within 30 minutes of the interview differed by site. More of the Long Beach drivers admitted to alcohol use within a half hour. Drivers from both groups and at both sites denied drinking within 30 minutes and tested positive on the PBT.

1.13 Drinking Habits

Since previous research has shown that relative risk varies as a function of a person's normal drinking pattern, drivers were asked about drinking frequency and quantity. They were also asked which alcohol beverage they preferred and about their drinking and driving practices.

Table F-15. Employment Status Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	**Control**	**Total**	**Crash**	**Control**	**Total**	**Crash**	**Control**	**Total**
Employment Status	**Employed Full Time**	**Count**	1186	3424	4610	1452	3798	5250	2638	7222	9860
		% within Crash/Control Status	58.9%	70.5%	67.1%	70.8%	75.8%	74.3%	64.9%	73.2%	70.8%
		% of Total	17.3%	49.8%	67.1%	20.6%	53.8%	74.3%	18.9%	51.8%	70.8%
	Employed Part Time	**Count**	245	503	748	146	376	522	391	879	1270
		% within Crash/Control Status	12.2%	10.4%	10.9%	7.1%	7.5%	7.4%	9.6%	8.9%	9.1%
		% of Total	3.6%	7.3%	10.9%	2.1%	5.3%	7.4%	2.8%	6.3%	9.1%
	Unemployed and Looking For Work	**Count**	156	201	357	93	171	264	249	372	621
		% within Crash/Control Status	7.7%	4.1%	5.2%	4.5%	3.4%	3.7%	6.1%	3.8%	4.5%
		% of Total	2.3%	2.9%	5.2%	1.3%	2.4%	3.7%	1.8%	2.7%	4.5%
	Retired	**Count**	118	185	303	110	258	368	228	443	671
		% within Crash/Control Status	5.9%	3.8%	4.4%	5.4%	5.1%	5.2%	5.6%	4.5%	4.8%
		% of Total	1.7%	2.7%	4.4%	1.6%	3.7%	5.2%	1.6%	3.2%	4.8%
	Going To School	**Count**	191	342	533	159	200	359	350	542	892
		% within Crash/Control Status	9.5%	7.0%	7.8%	7.7%	4.0%	5.1%	8.6%	5.5%	6.4%
		% of Total	2.8%	5.0%	7.8%	2.3%	2.8%	5.1%	2.5%	3.9%	6.4%
	Homemaker	**Count**	59	119	178	24	80	104	83	199	282
		% within Crash/Control Status	2.9%	2.5%	2.6%	1.2%	1.6%	1.5%	2.0%	2.0%	2.0%
		% of Total	.9%	1.7%	2.6%	.3%	1.1%	1.5%	.6%	1.4%	2.0%
	Disabled	**Count**	45	57	102	25	45	70	70	102	172
		% within Crash/Control Status	2.2%	1.2%	1.5%	1.2%	.9%	1.0%	1.7%	1.0%	1.2%
		% of Total	.7%	.8%	1.5%	.4%	.6%	1.0%	.5%	.7%	1.2%
	Other	**Count**	14	24	38	43	84	127	57	108	165
		% within Crash/Control Status	.7%	.5%	.6%	2.1%	1.7%	1.8%	1.4%	1.1%	1.2%
		% of Total	.2%	.3%	.6%	.6%	1.2%	1.8%	.4%	.8%	1.2%
	Total	**Count**	2014	4855	6869	2052	5012	7064	4066	9867	13933
		% of Total	29.3%	70.7%	100%	29.0%	71.0%	100%	29.2%	70.8%	100%

Table F-16. Hours Last Slept Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Hours Last Slept	0-4	Count	83	289	372	95	393	488	178	682	860
		% within Crash/Control Status	4.1%	6.0%	5.4%	4.7%	7.8%	6.9%	4.4%	6.9%	6.2%
		% of Total	1.2%	4.2%	5.4%	1.3%	5.6%	6.9%	1.3%	4.9%	6.2%
	5	Count	86	354	440	71	397	468	157	751	908
		% within Crash/Control Status	4.3%	7.3%	6.4%	3.5%	7.9%	6.6%	3.9%	7.6%	6.5%
		% of Total	1.3%	5.2%	6.4%	1.0%	5.6%	6.6%	1.1%	5.4%	6.5%
	6	Count	207	924	1131	237	847	1084	444	1771	2215
		% within Crash/Control Status	10.3%	19.0%	16.5%	11.6%	16.9%	15.4%	11.0%	18.0%	15.9%
		% of Total	3.0%	13.5%	16.5%	3.4%	12.0%	15.4%	3.2%	12.7%	15.9%
	7	Count	389	1024	1413	352	1088	1440	741	2112	2853
		% within Crash/Control Status	19.4%	21.1%	20.6%	17.2%	21.7%	20.4%	18.3%	21.4%	20.5%
		% of Total	5.7%	14.9%	20.6%	5.0%	15.4%	20.4%	5.3%	15.2%	20.5%
	8	Count	744	1491	2235	719	1420	2139	1463	2911	4374
		% within Crash/Control Status	37.1%	30.7%	32.6%	35.2%	28.3%	30.3%	36.1%	29.5%	31.4%
		% of Total	10.8%	21.7%	32.6%	10.2%	20.1%	30.3%	10.5%	20.9%	31.4%
	9	Count	230	393	623	260	417	677	490	810	1300
		% within Crash/Control Status	11.5%	8.1%	9.1%	12.7%	8.3%	9.6%	12.1%	8.2%	9.3%
		% of Total	3.4%	5.7%	9.1%	3.7%	5.9%	9.6%	3.5%	5.8%	9.3%
	10+	Count	269	380	649	308	447	755	577	827	1404
		% within Crash/Control Status	13.4%	7.8%	9.5%	15.1%	8.9%	10.7%	14.2%	8.4%	10.1%
		% of Total	3.9%	5.5%	9.5%	4.4%	6.3%	10.7%	4.1%	5.9%	10.1%
	Total	Count	2008	4855	6863	2042	5009	7051	4050	9864	13914
		% of Total	29.3%	70.7%	100%	29.0%	71.0%	100%	29.1%	70.9%	100%

Table F-17. Alcohol Consumption in the Last 2 Hours Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Consumed Alcohol in the Last 2 Hours	**Yes**	**Count**	233	305	538	317	785	1102	550	1090	1640
		% within Crash/Control Status	11.6%	6.3%	7.8%	15.4%	15.7%	15.6%	13.5%	11.0%	11.8%
		% of Total	3.4%	4.4%	7.8%	4.5%	11.1%	15.6%	3.9%	7.8%	11.8%
	No	**Count**	1781	4551	6332	1737	4228	5965	3518	8779	12297
		% within Crash/Control Status	88.4%	93.7%	92.2%	84.6%	84.3%	84.4%	86.5%	89.0%	88.2%
		% of Total	25.9%	66.2%	92.2%	24.6%	59.8%	84.4%	25.2%	63.0%	88.2%
	Total	**Count**	2014	4856	6870	2054	5013	7067	4068	9869	13937
		% of Total	29.3%	70.7%	100%	29.1%	70.9%	100%	29.2%	70.8%	100%

Table F-18. Alcohol Consumption in the Last 30 Minutes Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Consumed Alcohol in the Last 30 Minutes	**Yes**	**Count**	90	109	199	68	256	324	158	365	523
		% within Crash/Control Status	4.5%	2.2%	2.9%	3.3%	5.1%	4.6%	3.9%	3.7%	3.8%
		% of Total	1.3%	1.6%	2.9%	1.0%	3.6%	4.6%	1.1%	2.6%	3.8%
	No	**Count**	1922	4745	6667	1984	4757	6741	3906	9502	13408
		% within Crash/Control Status	95.5%	97.8%	97.1%	96.7%	94.9%	95.4%	96.1%	96.3%	96.2%
		% of Total	28.0%	69.1%	97.1%	28.1%	67.3%	95.4%	28.0%	68.2%	96.2%
	Total	**Count**	2012	4854	6866	2052	5013	7065	4064	9867	13931
		% of Total	29.3%	70.7%	100%	29.0%	71.0%	100%	29.2%	70.8%	100%

Drivers most frequently responded to a question about preferred beverage by saying they were non-drinkers (Table F-19). For those who indicated a preference, almost 52% chose beer. The second choice was wine, and liquor was the third choice. The mean number of drinks per drinking day appears in Table F-20. Again, non-drinkers occupy the largest category, and the most frequent response by drinkers was a single drink.

Subjects were asked how many times they had driven within two hours of drinking alcohol during the prior 28 days (Table F-21). Because few drivers admitted to having done so more than one time, the data are tabled as Yes and No responses. The pattern differs by site and group. In Fort Lauderdale, 15.4% of the crash drivers and 23.7% of the controls stated they had not driven after drinking in the past 28 days. In Long Beach, approximately 13% of both groups denied having done so.

1.14 DUI Arrests

Drivers were asked whether they had ever been arrested for DUI (Table F-22). A higher proportion of both crash and control drivers in Long Beach reported DUI arrests. At both sites, the crash-involved drivers more often reported a DUI arrest than did the controls.

1.15 Other Drugs

Table F-23 summarizes responses about the use of drugs. The responses were relatively uniform by site and group, and 21.6% of the drivers admitted to using medicine or other drugs with slightly more such admissions from the crash-involved drivers.

1.16 Crash Site Land Use Characteristics

The research team recorded observations of land use in the area of the crash. Categories recorded were for establishments selling alcohol for on-site consumption (restaurant, tavern, bar), establishments selling alcohol for off-site consumption (package store), homes, apartment buildings, retail stores and/or industrial buildings/activities (Table F-24). On-site establishments were more numerous in Fort Lauderdale, but the number of off-site outlets was similar for the two sites. Crashes occurred more frequently in Long Beach residential neighborhoods and in retail or industrial areas in Fort Lauderdale.

1.17 BAC Distribution

Table F-25 presents the distribution of the 13,886 quantitative BAC readings that met the study's criteria.[2] The measurements were obtained by PBT breath tests, police evidentiary breath tests, and hospital blood analyses.

Overall, 86.6 % of the drivers did not have measurable BACs with somewhat more drivers at 0.00 BAC in Long Beach than in Fort Lauderdale. More control than crash drivers were at a zero BAC. Crash-involved drivers were underrepresented through 0.059% BAC.

---

[2] In order to be included, a BAC had to be measured by the project's breath testing equipment, a police evidential device or a hospital laboratory from a specimen drawn within two hours of the crash time. BACs taken after the two hour limit or with unknown measurement times were excluded from all relative risk analyses

Crash-involved drivers had a higher relative frequency of BACs ≥ 0.06% than controls, and the discrepancy increased with increasing BAC.

Table F-19. Type of Alcohol Reported by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Type of Alcohol	**Beer**	Count	665	1730	2395	675	1687	2362	1340	3417	4757
		% within Crash/Control Status	33.1%	35.6%	34.9%	33.0%	33.7%	33.5%	33.0%	34.6%	34.2%
		% of Total	9.7%	25.2%	34.9%	9.6%	23.9%	33.5%	9.6%	24.5%	34.2%
	Liquor	Count	178	428	606	207	698	905	385	1126	1511
		% within Crash/Control Status	8.8%	8.8%	8.8%	10.1%	13.9%	12.8%	9.5%	11.4%	10.8%
		% of Total	2.6%	6.2%	8.8%	2.9%	9.9%	12.8%	2.8%	8.1%	10.8%
	Wine	Count	264	776	1040	358	851	1209	622	1627	2249
		% within Crash/Control Status	13.1%	16.0%	15.1%	17.5%	17.0%	17.1%	15.3%	16.5%	16.1%
		% of Total	3.8%	11.3%	15.1%	5.1%	12.1%	17.1%	4.5%	11.7%	16.1%
	Don't Drink	Count	905	1921	2826	808	1777	2585	1713	3698	5411
		% within Crash/Control Status	45.0%	39.6%	41.2%	39.5%	35.4%	36.6%	42.2%	37.5%	38.8%
		% of Total	13.2%	28.0%	41.2%	11.4%	25.2%	36.6%	12.3%	26.6%	38.8%
	Total	Count	2012	4855	6867	2048	5013	7061	4060	9868	13928
		% of Total	29.3%	70.7%	100%	29.0%	71.0%	100%	29.1%	70.9%	100%

Table F-20. Mean Number of Drinks per Drinking Day Reported by Crash and Control Drivers by Site

Mean Number of Drinks Per Drinking Day		Long Beach			Fort Lauderdale			Total		
		Crash	**Control**	**Total**	**Crash**	**Control**	**Total**	**Crash**	**Control**	**Total**
0	**Count**	1104	2445	3549	937	2075	3012	2041	4520	6561
	% within Crash/Control Status	55.6%	50.9%	52.3%	46.6%	41.7%	43.1%	51.1%	46.2%	47.6%
	% of Total	16.3%	36.0%	52.3%	13.4%	29.7%	43.1%	14.8%	32.8%	47.6%
1	**Count**	438	1131	1569	483	1226	1709	921	2357	3278
	% within Crash/Control Status	22.1%	23.5%	23.1%	24.0%	24.6%	24.4%	23.0%	24.1%	23.8%
	% of Total	6.5%	16.7%	23.1%	6.9%	17.5%	24.4%	6.7%	17.1%	23.8%
2	**Count**	243	650	893	319	824	1143	562	1474	2036
	% within Crash/Control Status	12.2%	13.5%	13.2%	15.9%	16.5%	16.3%	14.1%	15.1%	14.8%
	% of Total	3.6%	9.6%	13.2%	4.6%	11.8%	16.3%	4.1%	10.7%	14.8%
3-4	**Count**	160	419	579	202	679	881	362	1098	1460
	% within Crash/Control Status	8.1%	8.7%	8.5%	10.0%	13.6%	12.6%	9.1%	11.2%	10.6%
	% of Total	2.4%	6.2%	8.5%	2.9%	9.7%	12.6%	2.6%	8.0%	10.6%
5-7	**Count**	29	128	157	55	146	201	84	274	358
	% within Crash/Control Status	1.5%	2.7%	2.3%	2.7%	2.9%	2.9%	2.1%	2.8%	2.6%
	% of Total	.4%	1.9%	2.3%	.8%	2.1%	2.9%	.6%	2.0%	2.6%
8-11	**Count**	11	31	42	16	31	47	27	62	89
	% within Crash/Control Status	.6%	.6%	.6%	.8%	.6%	.7%	.7%	.6%	.6%
	% of Total	.2%	.5%	.6%	.2%	.4%	.7%	.2%	.4%	.6%
Total	**Count**	1985	4804	6789	2012	4981	6993	3997	9785	13782
	% of Total	29.2%	70.8%	100%	28.8%	71.2%	100%	29.0%	71.0%	100%

Table F-21. Reports of Driving After Drinking in Last 28 Days by Crash and Control

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Driven Within 2 Hours of Drinking	**Yes**	**Count**	1749	4217	5966	1726	3824	5550	3475	8041	11516
		% within Crash/Control Status	87.3%	87.0%	87.1%	84.6%	76.3%	78.7%	86.0	81.6	82.8
		% of Total	25.5%	61.5%	87.1%	24.5%	54.2%	78.7%	25.0	57.8	82.8
	No	**Count**	254	632	886	314	1187	1501	568	1819	2387
		% within Crash/Control Status	12.7%	13.0%	12.9%	15.4%	23.7%	21.3%	14.0	18.4	17.2
		% of Total	3.7%	9.2%	12.9%	4.5%	16.8%	21.3%	4.1	13.1	17.2
	Total	**Count**	2003	4849	6852	2040	5011	7051	4043	9860	13903
		% of Total	29.2%	70.8%	100%	28.9%	71.1%	100%	29.1%	70.9%	100%

Table F-22. Reports of Prior DUI Arrests by Crash and Control Drivers by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Arrested For DUI?	**Yes**	**Count**	183	363	546	141	243	384	324	606	930
		% within Crash/Control Status	9.1%	7.5%	8.0%	6.9%	4.8%	5.4%	8.0%	6.1%	6.7%
		% of Total	2.7%	5.3%	8.0%	2.0%	3.4%	5.4%	2.3%	4.4%	6.7%
	No	**Count**	1826	4487	6313	1909	4769	6678	3735	9256	12991
		% within Crash/Control Status	90.9%	92.5%	92.0%	93.1%	95.2%	94.6%	92.0%	93.9%	93.3%
		% of Total	26.6%	65.4%	92.0%	27.0%	67.5%	94.6%	26.8%	66.5%	93.3%
	Total	**Count**	2009	4850	6859	2050	5012	7062	4059	9862	13921
		% of Total	29.3%	70.7%	100%	29.0%	71.0%	100%	29.2%	70.8%	100%

Table F-23. Reports of Use of Medicine or Drugs by Crash and Control Drivers by

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
Used Medicine or Drugs	**Yes**	**Count**	464	967	1431	470	1110	1580	934	2077	3011
		% within Crash/Control Status	23.1%	20.0%	20.9%	23.0%	22.1%	22.4%	23.0%	21.1%	21.6%
		% of Total	6.8%	14.1%	20.9%	6.7%	15.7%	22.4%	6.7%	14.9%	21.6%
	No	**Count**	1547	3877	5424	1577	3903	5480	3124	7780	10904
		% within Crash/Control Status	76.9%	80.0%	79.1%	77.0%	77.9%	77.6%	77.0%	78.9%	78.4%
		% of Total	22.6%	56.6%	79.1%	22.3%	55.3%	77.6%	22.5%	55.9%	78.4%
	Total	**Count**	2011	4844	6855	2047	5013	7060	4058	9857	13915
		% of Total	29.3%	70.7%	100%	29.0%	71.0%	100%	29.2%	70.8	100%

Table F-24. Land Use Characteristics at Crash Locations by Site

Characteristic Visible at Crash Site		Long Beach	Fort Lauderdale
Establishment for On-site Alcohol Consumption	**Count**	241	449
	% within Site	17.0%*	30.9%
Establishment for Off-site Alcohol Consumption	**Count**	341	327
	% within Site	24.0%	22.5%
Homes	**Count**	793	444
	% within Site	55.9%	30.6%
Apartment Buildings	**Count**	715	392
	% within Site	50.4&	27.0%
Retail Stores	**Count**	647	875
	% within Site	45.6%	60.3%
Industry	**Count**	77	104
	% within Site	5.4%	7.2%

Table F-25. Crash and Control Drivers' BAC by Site

			Long Beach			Fort Lauderdale			Total		
			Crash	Control	Total	Crash	Control	Total	Crash	Control	Total
BAC Level (Interval)	0.00	Count	1706	4442	6148	1611	4268	5879	3317	8710	12027
		% within Crash/Control Status	82.9%	91.8%	89.2%	80.3%	85.6%	84.1%	81.6%	88.7%	86.6%
		% of Site Total	24.7%	64.4%	89.2%	23.0%	61.0%	84.1%	23.9%	62.7%	86.6%
	.001 - .019	Count	36	117	153	50	192	242	86	309	395
		% within Crash/Control Status	1.7%	2.4%	2.2%	2.5%	3.9%	3.5%	2.1%	3.1%	2.8%
		% of Site Total	.5%	1.7%	2.2%	.7%	2.7%	3.5%	.6%	2.2%	2.8%
	.020 - .039	Count	28	86	114	50	201	251	78	287	365
		% within Crash/Control Status	1.4%	1.8%	1.7%	2.5%	4.0%	3.6%	1.9%	2.9%	2.6%
		% of Site Total	.4%	1.2%	1.7%	.7%	2.9%	3.6%	.6%	2.1%	2.6%
	.040 - .059	Count	22	62	84	41	131	172	63	193	256
		% within Crash/Control Status	1.1%	1.3%	1.2%	2.0%	2.6%	2.5%	1.5%	2.0%	1.8%
		% of Site Total	.3%	.9%	1.2%	.6%	1.9%	2.5%	.5%	1.4%	1.8%
	.060 - .079	Count	21	46	67	48	87	135	69	133	202
		% within Crash/Control Status	1.0%	1.0%	1.0%	2.4%	1.7%	1.9%	1.7%	1.4%	1.5%
		% of Site Total	.3%	.7%	1.0%	.7%	1.2%	1.9%	.5%	1.0%	1.5%
	.080 - .099	Count	17	23	40	38	26	64	55	49	104
		% within Crash/Control Status	.8%	.5%	.6%	1.9%	.5%	.9%	1.4%	.5%	.7%
		% of Site Total	.2%	.3%	.6%	.5%	.4%	.9%	.4%	.4%	.7%
	.100 - .119	Count	22	19	41	21	28	49	43	47	90
		% within Crash/Control Status	1.1%	.4%	.6%	1.0%	.6%	.7%	1.1%	.5%	.6%
		% of Site Total	.3%	.3%	.6%	.3%	.4%	.7%	.3%	.3%	.6%
	.120 - .139	Count	26	17	43	24	20	44	50	37	87
		% within Crash/Control Status	1.3%	.4%	.6%	1.2%	.4%	.6%	1.2%	.4%	.6%
		% of Site Total	.4%	.2%	.6%	.3%	.3%	.6%	.4%	.3%	.6%
	.140 - .159	Count	31	10	41	29	8	37	60	18	78
		% within Crash/Control Status	1.5%	.2%	.6%	1.4%	.2%	.5%	1.5%	.2%	.6%
		% of Site Total	.4%	.1%	.6%	.4%	.1%	.5%	.4%	.1%	.6%
	.160 - .199	Count	51	11	62	57	18	75	108	29	137
		% within Crash/Control Status	2.5%	.2%	.9%	2.8%	.4%	1.1%	2.7%	.3%	1.0%
		% of Site Total	.7%	.2%	.9%	.8%	.3%	1.1%	.8%	.2%	1.0%
	.200+	Count	98	4	102	38	5	43	136	9	145
		% within Crash/Control Status	4.8%	.1%	1.5%	1.9%	.1%	.6%	3.3%	.1%	1.0%
		% of Site Total	1.4%	.1%	1.5%	.5%	.1%	.6%	1.0%	.1%	1.0%
	Total	Count	2058	4837	6895	2007	4984	6991	4065	9821	13886
		% of Site Total	29.8%	70.2%	100%	28.7%	71.3%	100%	29.3%	70.7%	100%

This Page Intentionally Left Blank

Appendix G

Additional Relative Risk Analysis Results

List of Tables in Appendix G

Table G-1. Log Linear Summary Table of Differences in Participation Rates

Source	*df*	**Chi-square**	***p***
Main effect of case-control Status on participation level	2	213.3	< .0001
Main effect of site on Participation level	2	1.40	0.50
Site x case-control status Interaction on participation level	2	65.6	< .0001

Table G-2. Sensitivity and Specificity of PAS in Detecting Alcohol-Positive Drivers (Long Beach and Fort Lauderdale Combined)

(*N* = 12,587)

I. Specificity:

Number of PAS Bars	Percentage with Positive BAC Values
0	4.75
1	80.2
2	84.0
3 or more	97.4

II. Sensitivity:

BAC Level	Percentage with Positive PAS Readings
0.0	0.55
.001-.019	29.5
.020-.039	53.5
.040-.059	76.6
.060-.079	77.5
.080-.099	86.1
.100-.119	91.2
.120-.139	89.9
.140-.159	91.9
.160-.199	90.3
.20+	94.3

Table G-3. Log Linear Analysis of Variance Table for PAS-Participation Level Relationship (Singlet Terms Omitted)*

Source	***df***	**Chi-square**	***P*****
Case-control status on PAS	3	22.2	< .0001
Case-control status on participation level	2	2.79	.25
Participation level on PAS	6	53.2	< .0001
Case-control status x participation level interaction on PAS	5***	12.9	.024

*Singlet terms fitted but omitted from display.

**Values subject to distortion due to the number of cells with small *N*s.

***Note: Reflects a parameter restriction due to a 0 frequency in the data matrix of the uncollapsed table, which entered PAS as 4 levels. The text table collapsed the 4 levels into 3 by combining levels 1 and 2.

Table G-4. Logistic Regression – Reduced All Covariates Model (Retention $p \leq .05$)

Covariate	Regression coefficient	*p* value	Adjusted odds ratio
Widowed	.536	.0006	1.71
Grammar school education	.240	.006	1.27
Ethnicity	--	--	--
Latino	-.274	< .0001	.76
Other ethnic	.245	.0021	1.28
Employment status	--	--	--
Unemployed	.420	< .0001	1.52
Student	.205	.038	1.23
Other/disabled	.385	.006	1.47
Military or police	.412	.003	1.52
Trip origin	--	--	--
Own home	-.196	.0005	.82
Other home	.146	.034	1.16
Other	.197	.004	1.22
Vehicle type	--	--	--
2 door	-.223	< .0001	.80
Convertible	.769	.007	2.16
Mean hours sleep	--	--	--
0-4	-.654	< .0001	.52
5-7	-.348	< .0001	.71
Hours last slept	--	--	--
0-4	-.602	< .0001	.55
5-7	-.382	< .0001	.68
Hours awake	-.00084	< .0001	.999
Drive after drinking	-.031	.014	.97
Drink days per month	-.017	< .0001	.98
Drinks per setting	--	--	--
3-4	-.295	.0002	.74
5-7	-.384	.022	.68
Deception Index (1=yes)	.511	< .0001	1.67
Gender	-.257	< .0001	.77
Gender x site	.361	< .0001	1.43
Age	--	--	--
Under 21 (ref.)	--	--	1.0
21-24	-.301	.004	.74
25-34	-.522	< .0001	.59
35-44	-.629	< .0001	.53
45-54	-.695	< .0001	.50
55+	-.441	< .0001	.64
BAC^1	-9.29	.025	--
BAC^2	312.9	< .0001	--
BAC^3	-878.3	< .0001	--

Cox-Snell R2 = .096; likelihood ratio chi-square = 1094.6; *df* = 33; $p < .0001$.

Note: On categorical variables, the coefficients represent deviations from the average of the logits of the referent group and all other categories that were deleted including the original referent group.

Dunlap and Associates, Inc.
110 Lenox Avenue
Stamford, Connecticut 06906-2300
(203) 323-8464   *Fax* (203) 964-0799