U.S. Department
of Transportation

National Highway
Traffic Safety
Administration

PB98-163603

People Saving People
http://www.nhtsa.dot.gov

Research Note

July 1997

Fatal Crash Involvements--What are the Odds?

Ezio C. Cerrelli

Introduction

The risk of being involved in a fatal crash varies considerably depending on specific crash characteristics, i.e., those of the driver, vehicle and environment. [For more on male and female fatal crash involvement by age group, see ***Crash Data and Rates for Age-Sex Groups of Drivers, 1994***, Research Note, October 1995.] A generally accepted measure of the risk of a fatal crash is the ratio between the number of fatal crashes associated with a specific set of conditions and the amount of travel performed under the same conditions.

To calculate risk values the appropriate values for the numerator and the denominator of the ratio must be available. In the case of fatal crash risk, the difficulties in obtaining highly detailed classifications of motor vehicle travel places a severe limitation on any attempt to estimate the risk of incurring a fatal crash under specific conditions.

Presently, travel data are not available at an adequate level of detail to permit estimating the contribution of any specific factor on the occurrence of a fatal crash. For example, the fatal involvement rates for male and female drivers, based on the total travel performed by the two groups of drivers, have, as an underlying assumption, that travel characteristics are the same for both male and female drivers. In other words, one must assume that the proportion of daytime travel is the same for male and female drivers. This assumption, therefore, must also be made for the proportion of high vs. low speed travel, day of week, type of crash, etc. While some of these assumptions may be valid, others are may be highly unlikely.

In the absence of detailed data on travel under specific conditions, the purpose of this study was to determine how the risk of being involved in a fatal crash, conditional on the occurrence of a police-reported crash, changes as a function of key crash characteristics. In this case, general assumptions are not needed, as crash involvements can be, in general, classified in as many ways as fatal involvements.

While the results do not provide a measure of the risk of a fatal involvement, they do show how the value of this risk varies as a function of key crash characteristics. Although the results are expressed as Odds Ratios, they can be considered to be approximately equal to risk ratios since the probability of a crash being fatal is very small.

Data and Analysis

Data from the Fatality Analysis Reporting System (FARS) and the General Estimate System (GES) files for the years 1993, 1994, and 1995 were used in this analysis. Only drivers of passenger cars and light trucks and vans were included in the study. Crashes involving pedestrians or pedalcyclists and crashes involving more than two vehicles were excluded. GES observations originate from a sample survey and therefore are associated with a sampling weight variable. The final data set for this study contained 99,139 fatal involvements and 207,909 unweighted GES crash involvements representing, when weighted, an estimated 28,071,714 involvements.

Seven variables were retained for each observation in the final data set described above. These variables were classified as:

- Driver Sex, i.e., Female or Male
- Driver Age, i.e., Younger (less than 65) or Older (greater than or equal to 65)
- Speed Limit, i.e., Low (less than 40mph), Moderate (40 thru 50mph), High (55mph), Highest (above 55mph)
- Hour of Day, i.e., as Daytime (6am to 9pm) or Nighttime (9pm to 6am)
- Day of Week,, i.e., Weekday or Weekend (6pm Friday through 6am Monday)

- Vehicle Type, i.e., Passenger Car (PC) or Light Trucks and Vans (LTVs)

- Type of Crash, i.e, Single Vehicle, Rear End, Side, or Head-on

Other variables or variable characteristics initially included in the analysis were omitted when found not to have a significant effect on the risk value. For example, for driver age, drivers under 25 years old were separated from the younger driver group, however estimated risk values were no greater for drivers in this age group compared to drivers aged 25-64. As result, drivers under 25 years of age were combined with drivers aged 25-64 to form a one group.

Survey Data Analysis software (SUDAAN™) was used with the FARS and GES data to estimate a statistical model to describe the variation in the risk value for the selected variables. SUDAAN™, developed at the Research Triangle Institute, is designed to perform statistical analysis of survey data, e.g., GES data. SUDAAN™ is appropriate for statistical analyses of databases containing both census and survey data, in this case, FARS (census) and GES (survey).

The final statistical model using SUDAAN™ provides an estimate of the change in the risk of a fatal involvement, given a crash, for each of the seven variables. The results for any variable can also be assumed to be independent of the effect of the remaining variables, as the interactions among the seven variables were excluded from the final statistical model as these were found not to have a significant contribution. The model can be used to provide estimates of the direct contribution of each individual variable to the risk of fatal involvement and for any combination of the remaining variables. Standard errors for each risk estimate are also provided in the following results.

Results

The GES cases are divided into twelve strata containing about 100 probability sampling units (PSU's). In order to proceed with the SUDAAN analysis, all FARS cases were assigned to a separate stratum and assigned individual PSU numbers.

The results of the SUDAAN analysis can be represented by a set of odds ratios, each representing the relative odds of being involved in a fatal crash, once involved in a police-reported crash, as a function of the characteristics of each of the seven variables. [Please note that driver involvement in a fatal crash does not imply that the driver dies in the crash.] The results of the analysis are shown in the following table:

Estimated Odds of Driver Fatal Crash Involvement by Key Driver and Crash Characteristics

Factor	Odds	95% Confidence Interval
Male/Female	1.33	1.29 - 1.38
Older/Younger	2.59	2.37 - 2.84
Moderate/Low Speed Limit	2.83	2.40 - 3.33
High/Low Speed Limit	7.17	5.56 - 9.24
Highest/Low Speed Limit	8.91	6.04 - 13.19
Weekend/Weekday	1.19	1.08 - 1.16
Nighttime/Daytime	1.91	1.79 - 2.03
LTV/PC	1.29	1.20 - 1.40
Side/Rear Collision	4.70	4.35 - 5.08
Single Vehicle/Rear Collision	8.09	6.93 - 9.45
Head On/Rear Collision	74.60	64.52 - 86.20

In the above table, the relative odds represent the ratio of the odds of a fatal involvement of the first subgroup vs. the lowest risk group, i.e. male vs. female, higher speed limit vs. lower, etc. For example, the risk of fatal crash involvement for male drivers is 1.33 compared to female drivers, so that male drivers have an estimated 33% greater risk of fatal crash involvement, given a crash, than female drivers. The last column of the table presents the confidence interval for the estimated risk value and provides a measure of the statistical variability of the estimate. Finally, the risk value or odds for each variable provides a measure of its contribution to the outcome of the crash, independently of the effect of the remaining factors.

Findings

The major contribution of this study is the quantification of key driver and crash factors, rather than the identification of these factors. That is, it was

generally accepted that the variables used in the analysis (hour of day, type of crash) were important in terms of their impact on crash outcome. Several of the variables, e.g., type of crash and speed limit, were found to have fairly high odds ratios while other variables, e.g., driver age and driver sex were found to have much lower estimated odds ratios. As such, this study may be the first attempt at this type of analysis based on a large body of data, and therefore, comparable findings from other studies may not be available.

In brief, the study finds that among all drivers involved in crashes the odds for male drivers to be involved in a fatal crash are 1.33 times that for female drivers.

Drivers in the older age group, 65 years old and above, display odds that are over two and one-half (2.59) times that for their younger counterparts.
Being involved in a crash during nighttime hours increases the odds of the crash being fatal by almost two times (1.91) compared to daytime crashes.

Relative to crashes on roadways with 'Low Speed' limits, the odds for crashes on 'Moderate' speed limits are 2.83 times as great, while the odds on "High' posted sped limit are 7.17 times as great. Crashes that occur on 'Highest' posted speeds have odds which are 8.91 times as great as those at "Low Speed' limits.

Being the driver of an LTV increases the odds of the involvement being fatal by 29 percent relative to drivers of passenger cars.

Drivers involved in a 'Side' collision have odds of being involved in a fatal crash that are 4.7 times that for drivers involved in 'Rear' collisions. The same odds for 'Single Vehicle' collisions are 8.1 times that for 'Rear' collisions. By far the factor that has the most severe effect on the odds is the occurrence of a 'Head On' collision. In this case the odds are 75 times as great as for 'Rear' collisions.

Overall, the odds of fatal involvement varies considerably as a function of these selected factors. The odds ratios presented in this paper can be considered to be independent and can be used to calculate the odds ratio for any specific combination of these factors versus its counterpart. For example, the odds for a Male driver, in a Light Truck, at Night, on a High (55 mph) speed road, involved in a Single Vehicle crash are about 190 to 1 *(1.33 x 1.29 x 1.91 x 7.17 x 8.09)* relative to a Female driver, in a Passenger Car, in Daytime, on a Low (under 40 mph) speed road, involved in Rear collision. The odds ratios for any other combination of these factors can be calculated similarly.

These results may serve a variety of purposes, among which the identification of those factors with high odds ratios, for example, in the ability to estimate the severity of a crash for EMS response and resource deployment.

For additional copies of this research note, please call (202) 366-4198 or fax your request to (202) 366-7078. For questions regarding the data reported in this research note, contact Ezio Cerrelli [202-366-5358] of the National Center for Statistics and Analysis. This research note and other general information on highway traffic safety may be accessed by Internet users at http://www.nhtsa.dot.gov/people/ncsa.