

Supplementary Online Content

eMethods. Link to download data used for case-control study of distance from home and crash risk; inclusion criteria; detailed description of derivation of study variables; and detailed results.

eResults. Tables showing distribution of cases and controls in relation to distance from home, crude odds ratios, adjusted odds ratios, and sensitivity analysis.

eReferences. References cited in eMethods.

eMethods

Data

The data used in the case-control analysis of the relationship between a driver's distance from home and crash risk were originally collected by the United States National Highway Traffic Safety Administration (NHTSA) for a study of the relationship between drivers' use of alcohol and drugs and their risk of involvement in a motor vehicle crash.¹ The data can be downloaded from the Web site of the NHTSA via the URL below. (The data are available in in SAS, SPSS, Stata, and Excel formats. The authors used the Stata version.)

<https://www.nhtsa.gov/drug-and-alcohol-crash-risk-study/drug-and-alcohol-crash-risk-study-databases>

Inclusion Criteria

Drivers were eligible for the current study if:

- They consented to participate in NHTSA's above-referenced study of the crash risk associated with alcohol and drugs (identified by variable *consent* having a value of 1).
- Were driving a car, pickup truck, van, minivan, or sport utility vehicle (*vehicle_type* in range 1-5).
- Provided a valid response to the question, "About how many miles away are you now from where you live?" (*miles_fromhome* in range 1-4).

The database included records of all 12,790 drivers contacted by the researchers in the course of the study, of which 2,004 did not consent to participate, an additional 147 were driving an ineligible or unknown type of vehicle, and an additional 25 did not report how far from home they were, leaving 10,611 who met all inclusion criteria.

Variables

Dependent variable

The dependent variable was *case* (1: Case; 0: Control), which identifies whether each observation was a case (a driver who crashed) or a control (a randomly selected driver flagged down by a police officer at the crash site 7 days after the crash and invited to participate in the study).

Main independent variable

The main independent variable was *miles_fromhome* (1: 0-5 miles; 2: 6-10 miles; 3: 11-20 miles; 4: More than 20 miles; 5,98,99: various forms of "unknown") which represents how far the driver was from home when he or she crashed or was randomly flagged down and invited to participate in the study. Drivers who were an unknown distance from home were excluded from all analyses.

Covariates

Covariates included in the analysis reported in the article are listed below, along with corresponding variable names and any necessary explanatory notes.

- *age* (whole numbers range 13-92) represents driver age in years. In multivariable analyses, the authors modeled age using terms *age* and *age-squared*, to allow for the relationship between age and crash risk to be non-monotonic. (Many studies find relationship to be roughly the shape of a "U" or a backwards "J.")
- *driver_sex* (1: male; 2: female).
- *crash_weather*; *control_weather* (1: clear; 2: cloudy; 3: light rain; 4: heavy rain; 5: light snow; 6: heavy snow; 7: fog; 8: wind; 9: other; 99: unknown). The authors derived a new variable, which they called

simply *weather*, in which cases had values from the variable *crash_weather* and controls had values from the variable *control_weather*, to indicate the weather conditions present for each individual driver. Few crashes occurred in heavy rain, snow, fog, wind, or other weather, and none of the controls for these crashes occurred under similar weather conditions. Thus, the authors included only crashes and controls where *weather* was clear, cloudy, or light rain. The presence of clouds is unlikely to influence crash risk; so the authors collapsed categories as 1: clear or cloudy; 2: light rain.

- *crash_roadway_conditions; control_roadway_conditions* (1: dry; 2: wet; 3: snow/ice; 4: slippery [mud, oil, etc.], 98: other; 99: unknown). The authors derived a new variable, *roadway_conditions*, in which cases had values from *crash_roadway_conditions* and controls had values from *control_roadway_conditions*, to indicate the roadway surface conditions present for each individual driver. Few crashes occurred in snow/ice or on slippery roads, thus the authors included only crashes and controls where *roadway_conditions* were dry or wet.
- *crash_lighting; control_lighting* (1: daylight; 2: dusk; 3: dawn; 4: dark-street lights; 5: dark-no street lights; 6: dark-street lights not functioning; 99: unknown). The authors derived a new variable, *lighting*, in which cases had values from *crash_lighting* and controls had values of *control_lighting*, and then collapsed categories as 1: daylight; 2: dawn or dusk; 3: dark.
- *vehicle_type* (1: car; 2: SUV; 3: minivan; 4: van; 5: pickup truck; 6: motorcycle; 7: other; 99: unknown). Motorcycles and vehicles of other/unknown type were excluded per the study inclusion criteria.

Matching variable

The variable *casenum* was used to identify drivers involved in the same crash as one-another and the corresponding control drivers for those cases.

Other variables

A sensitivity analysis was performed with drivers who tested positive for alcohol and/or illegal drugs excluded.

The alcohol information used in the study is contained in the variable *pbt_result_new*, which contains the driver's estimated breath alcohol concentration or blood alcohol concentration, depending on type of test administered. A driver was considered positive for alcohol if *pbt_result_new* > 0.

The drug information used in the study is contained in the variables *illegal_of* and *illegal_blood*; which indicated whether the driver's oral fluid sample and blood sample, respectively, were positive for illegal drugs. A driver was considered positive for illegal drugs if *illegal_of* = 1 or *illegal_blood* = 1.

Initially the authors derived a variable that measured years of driving experience, by subtracting the variable *age_licensed* (the age at which the driver reported obtaining his or her first license) from the variable *age*. However, unsurprisingly, this was highly correlated with the driver's age ($r^2 = 0.97$), so this variable ultimately was not included in the analysis model.

Statistical Analysis

Cases and controls were tabulated by distance from home. Crude odds ratios were estimated using conditional logistic regression with controls matched to cases by *casenum*. Adjusted odds ratios were estimated using conditional logistic regression with the covariates listed above added to the model. Crude odds ratios were estimated both using all observations and using observations with no missing values for any of the covariates included in the adjusted model. The adjusted model included complete observations only. Crude and adjusted odds ratios were also estimated with drivers positive for alcohol and/or illegal drugs excluded.

eResults

eTable 1. Distribution of distance from home among cases and controls.

	All Drivers		Complete Observations Only		Complete Observations; Alcohol & Drug-Positive Drivers Excluded	
	Controls (n=7,065)	Cases (n=3,546)	Controls (n=6,814)	Cases (n=3,362)	Controls (n=5,820)	Cases (n=2,504)
Miles from Home	<i>Column %</i>		<i>Column %</i>		<i>Column %</i>	
0 - 5	58.1	50.3	58.0	50.1	57.8	50.5
6 - 10	22.0	24.3	22.1	24.2	22.0	24.4
11 - 20	11.4	13.7	11.3	13.7	11.5	13.3
> 20	8.5	11.8	8.5	12.0	8.7	11.8

Column percentages may not add to 100.0 due to rounding.

eTable 2. Ratios of crude odds of crash involvement in relation to driver's distance from home.

Miles from Home	All Drivers		Complete Observations Only		Alcohol & Drug-Positive Drivers Excluded	
	<i>Odds Ratio (95% CI)</i>		<i>Odds Ratio (95% CI)</i>		<i>Odds Ratio (95% CI)</i>	
0 - 5	1	–	1	–	1	–
6 - 10	1.29	(1.17 - 1.44)	1.27	(1.14 - 1.42)	1.27	(1.12 - 1.44)
11 - 20	1.46	(1.29 - 1.67)	1.48	(1.29 - 1.70)	1.37	(1.17 - 1.61)
> 20	1.70	(1.47 - 1.96)	1.71	(1.48 - 1.98)	1.62	(1.36 - 1.92)

Odds ratios estimated using conditional logistic regression with controls matched to cases.

eTable 3. Ratios of adjusted odds of crash involvement in relation to driver's distance from home.

Miles from Home	All Complete Observations		Alcohol & Drug-Positive Drivers Excluded	
	<i>Odds Ratio (95% CI)</i>		<i>Odds Ratio (95% CI)</i>	
0 - 5	1	–	1	–
6 - 10	1.28	(1.15 - 1.43)	1.28	(1.12 - 1.45)
11 - 20	1.47	(1.47 - 1.69)	1.35	(1.15 - 1.60)
> 20	1.77	(1.52 - 2.06)	1.68	(1.40 - 2.00)

Odds ratios estimated using conditional logistic regression with controls matched to cases. Covariates adjusted were age, age², sex, vehicle type, rain, wet pavement, and lighting conditions.

eReferences

1. Lacey, J. H., Kelley-Baker, T., Berning, et al. (2016). *Drug and alcohol crash risk: A case-control study* (Report No. DOT HS 812 355). Washington, DC: National Highway Traffic Safety Administration.