README for "Drug and Alcohol Crash Risk: A Case-Control Study [Supporting Datasets]" dataset.
Office of Behavioral Safety Research (BSR), National Highway Transportation Safety Administration (NHTSA), U.S. Department of Transportation (USDOT)
2024-01-20

LINKS TO DATASET

A. Dataset homepage link:
    https://doi.org/10.21949/1529970

SUMMARY OF DATASET

This dataset supports the conclusions of the report "Drug and Alcohol Crash Risk: A Case-Control Study." The study used a "case-control" design to estimate the risk of crashes involving drivers using drugs, alcohol or both. Data was collected in Virginia Beach, Virginia, for 20 months. The study obtained biological measures on more than 3,000 crash drivers at the scenes of the crashes, and 6,000 control (comparison) drivers. Control drivers were recruited one week after the crashes at the same time, day of week, location, and direction of travel as the crash-involved drivers. Data included 10,221 breath samples, 9,285 oral fluid samples, and 1,764 blood samples. Oral fluid and blood samples were screened and confirmed for the presence of alcohol and drugs. The crash risk associated with alcohol and other drugs was estimated using odds ratios that indicate the probability of a crash occurring over the probability that such an event does not occur. If a variable (alcohol and/or drugs) is not associated with a crash, the odds ratio for that variable will be 1.00. A higher or lower number indicates a stronger relationship between the probability of a crash occurring and the presence of that variable (alcohol and/or drugs in the driver). Confidence intervals (CIs) of an odds ratio indicate the range in which the true value lies—with 95 percent confidence. Alcohol: Alcohol was the largest contributor to crash risk. The unadjusted crash risk estimates for alcohol indicated drivers with a breath alcohol concentrations (BrACs) of .05 grams per 210 liters (g/210L) are 2.05 times more likely to crash than drivers with no alcohol. For drivers with BrACs of .08 g/210L, the unadjusted relative risk of crashing is 3.98 times that of drivers with no alcohol. When adjusted for age and gender, drivers with BrACs of .05 g/210L are 2.07 times more likely to crash than drivers with no alcohol. The adjusted crash risk for drivers at .08 g/210L is 3.93 times that of drivers with no alcohol. Drugs: Unadjusted drug odds ratio estimates indicated a significant increase in crash risk. For the active ingredient in marijuana, delta-9-tetrahydrocannabinol (THC), this yielded an unadjusted odds ratio of 1.25. However, after adjusting for gender, age, race/ethnicity, and alcohol, there was no indication that any drug significantly contributed to crash risk. The adjusted odds ratios for THC were 1.00, 95 percent CI [.83, 1.22], indicating no increased or decreased crash risk. Odds ratios for antidepressants were .86, 95 percent CI [.56, 1.33]; narcotic analgesics were 1.17, 95% percent drugs as an overall category were .99, 95 percent CI [.84, 1.18], and prescription and over-the-counter medications were 1.02, 95 percent CI [.83, 1.26]. Alcohol and Drugs: Analyses found no statistically significant interaction effects when drivers were positive for both alcohol and drugs. Although initial analyses suggested that the combination of alcohol and other drugs were contributors to increased crash risk, additional analyses adjusting for other risk factors indicated no significant effect. When both alcohol and other drugs were consumed, alcohol alone was associated with crash risk. NTL staff has reviewed the data. While researchers gathered names, ages, and other demographic information from participants, those sensitive data have either been excluded from the final dataset, anonymized, or pose only a small chance of re-identification. Further, while data collection took place over 20 months, the "Year" that any crash occurred is not revealed, further reducing re-identification risk. NTL staff feel the risk of re-identification of study participants from this dataset is low, but not zero.
The .ZIP folder of datasets and supporting documentation is 20 MB in size. The ZIP contains files in the following formats: .CSV files which can be opened with any text editor; .TXT files which can be opened with any text editor; .PDF files that can be opened with any PDF reader; .DOCX files that can be opened in Microsoft Word and some web-based programs; .SAV files which can be opened with IBM SPSS statistical software; .SAS and .sas7bdat files which can be opened with SAS statistical software; .XLSX files which can be opened with Microsoft Excel and other

spreadsheet programs; and, .JSON files which can be opened with text editors or metadata editing programs.

TABLE OF CONTENTS

A. GENERAL INFORMATION
----

0. Title of Dataset:
   Drug and Alcohol Crash Risk: A Case-Control Study [Supporting Datasets]

1. Description of Dataset:
   This dataset supports the conclusions of the report "Drug and Alcohol Crash Risk: A Case-Control Study." The study used a "case-control" design to estimate the risk of crashes involving drivers using drugs, alcohol or both. Data was collected in Virginia Beach, Virginia, for 20 months. The study obtained biological measures on more than 3,000 crash drivers at the scenes of the crashes, and 6,000 control (comparison) drivers. Control drivers were recruited one week after the crashes at the same time, day of week, location, and direction of travel as the crash-involved drivers. Data included 10,221 breath samples, 9,285 oral fluid samples, and 1,764 blood samples. Oral fluid and blood samples were screened and confirmed for the presence of alcohol and drugs. The crash risk associated with alcohol and other drugs was estimated using odds ratios that indicate the probability of a crash occurring over the probability that such an event does not occur. If a variable (alcohol and/or drugs) is not associated with a crash, the odds ratio for that variable will be 1.00. A higher or lower number indicates a stronger relationship between the probability of a crash occurring and the presence of that variable (alcohol and/or drugs in the driver). Confidence intervals (CIs) of an odds ratio indicate the range in which the true value lies—with 95 percent confidence. Alcohol: Alcohol was the largest contributor to crash risk. The unadjusted crash risk estimates for alcohol indicated drivers with a breath alcohol concentrations (BrACs) of .05 grams per 210 liters (g/210L) are 2.05 times more likely to crash than drivers with no alcohol. For drivers with BrACs of .08 g/210L, the unadjusted relative risk of crashing is 3.98 times that of drivers with no alcohol. When adjusted for age and gender, drivers with BrACs of .05 g/210L are 2.07 times more likely to crash than drivers with no alcohol. The adjusted crash risk for drivers at .08 g/210L is 3.93 times that of drivers with no alcohol. Drugs: Unadjusted drug odds ratio estimates indicated a significant increase in crash risk. For the active ingredient in marijuana, delta-9-tetrahydrocannabinol (THC), this yielded an unadjusted odds ratio of 1.25. However, after adjusting for gender, age, race/ethnicity, and alcohol, there was no indication that any drug significantly contributed to crash risk. The adjusted odds ratios for THC were 1.00, 95 percent CI [.83, 1.22], indicating no increased or decreased crash risk. Odds ratios for antidepressants were .86, 95 percent CI [.56, 1.33]; narcotic analgesics were 1.17, 95% percent drugs as an overall category were .99, 95 percent CI [.84, 1.18], and prescription and over-the-counter medications were 1.02, 95 percent CI [.83, 1.26]. Alcohol and Drugs: Analyses found no statistically significant interaction effects when drivers were positive for both alcohol and drugs. Although initial analyses suggested that the combination of alcohol and other

drugs were contributors to increased crash risk, additional analyses adjusting for other risk factors indicated no significant effect. When both alcohol and other drugs were consumed, alcohol alone was associated with crash risk.

NTL staff has reviewed the data. While researchers gathered names, ages, and other demographic information from participants, those sensitive data have either been excluded from the final dataset, anonymized, or pose only a small chance of re-identification. Further, while data collection took place over 20 months, the "Year" that any crash occurred is not revealed, further reducing re-identification risk. NTL staff feel the risk of re-identification of study participants from this dataset is low, but not zero.

The .ZIP folder of datasets and supporting documentation is 20 MB in size. The ZIP contains files in the following formats: .CSV files which can be opened with any text editor; .TXT files which can be opened with any text editor; .PDF files that can be opened with any PDF reader; .DOCX files that can be opened in Microsoft Word and some web-based programs; .SAV files which can be opened with IBM SPSS statistical software; .SAS and .sas7bdat files which can be opened with SAS statistical software; .XLSX files which can be opened with Microsoft Excel and other spreadsheet programs; and, .JSON files which can be opened with text editors or metadata editing programs.

2.A Dataset homepage link:
    https://doi.org/10.21949/1529970

3. Authorship Information:
    Principal Data Creator or Data Manager Contact Information
        Name: Amy Berning
            ORCiD: https://orcid.org/0000-0002-0614-2793
            Office of Behavioral Safety Research (BSR), National Highway Transportation Safety Administration (NHTSA), U.S. Department of Transportation
            Address: 1200 New Jersey Ave SE, Washington D.C. 20590
            Email: amy.berning@dot.gov

    Organizational Contact Information
        Name: National Transportation Library Data Curator
        Institution: National Transportation Library, Office of Information and Library Sciences, Bureau of Transportation Statistics, U.S. Department of Transportation
            Address: 1200 New Jersey Ave SE, Washington D.C. 20590
            Email: ntldatacurator@dot.gov

4. Date of data collection and update interval:
    Data was collected over 20 months. However, the Years are not disclosed, to help protect privacy. No updates

5. Geographic location of data collection:
    The state of Virginia, in the United States, at Virginia Beach, as noted in the report "Drug and Alcohol Crash Risk: A Case-Control Study" https://doi.org/10.21949/1525791

6. Information about funding sources that supported the collection of the data:
    DTNH22-06-C-00040

```
-----------------------------------------------------------------
B. SHARING/ACCESS & POLICIES INFORMATION
-----------------------------------------------------------------
```

0. Recommended citation for the data:
   U.S. Department of Transportation, National Highway Transportation Safety Administration (NHTSA), Office of Behavioral Safety Research (BSR). (Last updated 2024-01.) Drug and Alcohol Crash Risk: A Case-Control Study [Supporting Datasets]. https://doi.org/10.21949/1529970

1. Licenses/restrictions placed on the data:
   These data are in the Public Domain.

2. Was data derived from another source?:
   No

3. This dataset and its documentation were created and shared to meet the requirements enumerated in the U.S. Department of Transportation's "Plan to Increase Public Access to the Results of Federally-Funded Scientific Research" Version 1.1 <<https://doi.org/10.21949/1524190>> and guidelines suggested by the DOT Public Access website <<https://doi.org/10.21949/1503647>>, in effect and current as of January 20, 2024.


```
-----------------------------------------------------------------
C. DATA & RELATED FILE OVERVIEW
-----------------------------------------------------------------
```

1. Filenames
Data Files List
   A.1. Filename: NHTSA_BSR_DACRS_2016_DATA_csv.csv
      Short description: 12,790 data records, with 529 variables in header row. .CSV version, opens with any text editor.
   A.2. Filename: NHTSA_BSR_DACRS_2016_DATA_xlsx.xlsx
      Short description: 12,790 data records, with 529 variables in header row. .XLSX version, opens with Mircosoft Excel or other spreadsheet program.
   A.3. Filename: NHTSA_BSR_DACRS_2016_DATA_sav.zip
      Short description: 12,790 data records. .SAV version, opens with IBM SPSS statistical software.
   A.4. Filename: NHTSA_BSR_DACRS_2016_DATA_sas.zip
      Short description: 12,790 data records. .SAS version, opens with SAS statistical software.
   A.5. Filename: NHTSA_BSR_DACRS_2016_DATA_dta.zip
      Short description: 12,790 data records. .DTA version, opens with Stata statistical software.

README
   B. Filename: _NHTSA_BSR_DACRS_2016_README_20240120_1200.txt
      Short description:
      The README.txt file that includes human-readable information about the data, variable definitions, contact information, and other contextual information. The file you are reading now.

   C. Filename: _NHTSA_BSR_DACRS_2016_README_20240120_1200.pdf
      Short description:
      A PDF version of the README.txt file that includes human-readable information about the data, variable definitions, contact information, and other contextual information.

METADATA
   D. Filename: NHTSA_BSR_DACRS_2016_METADATA_20240120_1200.json
      Short description:
      The machine-readable .json metadata file based on DCAT-US (Project Open Data) metadata schema v1.1.

DATA MANAGEMENT PLAN
   E. Filename: NHTSA_BSR_DACRS_2016_DMP_20240120_1200.txt
      Short description:
      The human-readable data management plan associated with this dataset.

   F. Filename: NHTSA_BSR_DACRS_2016_DMP_20240120_1200.pdf
      Short description:
      A pdf version of the human-readable data management plan associated with this dataset.

DATA DICTIONARY
   G. Filename: NHTSA_BSR_DACRS_2016_DATADictionary.pdf
      Short description:
      This document defines the 529 variables collected or generated for this dataset.

REPORTS AND OUTPUTS
   H. Filenames:
        NHTSA_BSR_DACRS_2016_REPORT_812355.pdf;
NHTSA_BSR_DACRS_2016_REPORT_812117_Traffic_Safety_Facts.pdf

2. Are there multiple versions of the dataset? No

D. METHODOLOGICAL INFORMATION
----

1. Description of methods used for collection/generation of data:
   For Methodological information, please read pages 13 to 30 in the file
NHTSA_BSR_Wash_Marijuana_2015_REPORT_812299.pdf, "Drug and Alcohol Crash Risk: A Case-Control Study:
Methodology" This file is also available in ROSA P at https://doi.org/10.21949/1525791

E. DATA-SPECIFIC INFORMATION
----

1. NHTSA_BSR_DACRS_2016_DATA_csv.csv (as well as .xlsx, .sav, .sas, and .dta formats).

A. Notes on table structure: see Data Dictionary file NHTSA_BSR_DACRS_2016_DATADictionary.pdf

B. Number of variables:
   529 variables

C. Number of cases/rows:
  12,790 rows

D. Each row represents:
  1 vehicle involved in a crash or control survey in Virginia Beach during the study period.

E. Data Dictionary/Variable List:
  1. The full Data Dictionary is in file: NHTSA_BSR_DACRS_2016_DATADictionary.pdf

F. Missing data codes:
   #NULL!   See Data Dictionary
   Empty Cell   See Data Dictionary

F. CHANGE LOG / UPDATE LOG

This _NHTSA_BSR_DACRS_2016_README_20240120_1200.txt file was created on 2024-01-20 by Leighton L Christiansen https://orcid.org/0000-0002-0543-4268, Data Curator, leighton.christiansen@dot.gov

[Note changes or update to the readme.txt file, e.g.:]

2024-01-20: Original file created