Harmonization of 42 Years of Fatality Analysis Reporting System (FARS) Datasets

Nicholas R. Sachs and John H. Holmes, PhD, FACE, FACMI

DEPARTMENT of BIOSTATISTICS EPIDEMIOLOGY & INFORMATICS

UNIVERSITY of PENNSYLVANIA INJURY SCIENCE Center

Background

- Every year, 30,000+ people die in car crashes across the United States
- Researchers have analyzed policy changes and its effect on nationwide crash rates and comparisons between states, but there is a dearth of literature on analyzing differences within specific states.
- In order to answer questions regarding state-level issues leading to fatal crashes, the creation of a longitudinal, high quality dataset is necessary.

Data Source

- The National Highway Traffic Safety Administration (NHTSA) publishes online datasets containing information on all fatal car crashes that occur on public roadways each year.
- Titled the Fatality Analysis Reporting System (FARS), it is comprised of various datasets such as the accident, vehicle, and person.
- In order to ensure validity and accuracy, the NHTSA audits the data through internal consistency checks periodically.
- The Accident, Person, and Vehicle datasets from 1975-2017 were utilized for this dataset.

Harmonization Process Map

Utilized FARS Analytical User's Manual to create variable data dictionary to understand how coding schemas changed from past to present

↓

Recoded all variables using Stata v.16 according to the 2017 FARS coding schema

↓

Appended older accident, person, vehicle datasets to 2017 versions to create a single, harmonized accident, person, and vehicle dataset

↓

Performed a one-to-many merge of the person and vehicle datasets to the accident dataset

Assumptions & Limitations of Dataset

- The data gathered is retrospective. Any missing or incorrect information gathered at the time of the crash is unable to be either discovered or corrected.
- Due to the decisions made by the authors during recoding, nuances in the data may be lost or altered in order to adhere to modern coding schemas and facilitate longitudinal analysis.

Conclusions & Future Directions

- Harmonization of decades of fatal car crash data allows for researchers and policymakers to analyze statewide and national trends in fatal crashes.

- Future directions:
  - Compare crash rates and characteristics across counties in Pennsylvania to determine modifiable factors to reduce road fatalities
  - Investigate effects of state and national transportation policies on fatal crashes over time
  - Deep temporal phenotyping of crash events using complex systems models