Catalyzing Advancements via Data Linkage: New Jersey Traffic Safety Outcomes Program Data Warehouse

Melissa R. Pfeiffer, MPH
Center for Injury Research and Prevention
Traffic Records Forum
August 2017
Graduated Driver Licensing Decal Law
Effect on Young Probationary Drivers

Allison E. Curry, PhD, Melissa R. Pfeiffer, MPH, Russell Localio, PhD, Dennis R. Durbin, MD

**Background:** Decal laws have been implemented internationally to facilitate police enforcement of graduated driver licensing (GDL) restrictions (e.g., passenger limit, nighttime curfew) but have not been evaluated. New Jersey implemented the first decal law in the U.S. on May 1, 2010.
NJ Traffic Safety Outcomes (NJ-TSO) Data Warehouse
NJ Traffic Safety Outcomes (NJ-TSO) Data Warehouse

- Traffic Citations: 2004-2014
- Birth certificate data: 1979-2000
- Death certificate data: 2004-2014
- Licensing data: 2004-2014 (N = 10 million)
- Crash data: 2004-2014 (N = 6 million)
- Census data
- Hospital discharge data: 2004-2014
- CHOP electronic health record data: 1987-1997 (N = 148,000)
NJ Traffic Safety Outcomes (NJ-TSO) Data Warehouse

Licensing Data 2004 – 2014
(N = 10 million)

Crash Data 2004 – 2014
(N = 6 million)
Hierarchical Data Linkage

Licensing Data 2004 – 2014 (N = 10 million)

Crash Data 2004 – 2014 (N = 6 million)

Pass 1: Exact DLN
Pass 2: DOB, First Name (FN), Last Name (LN), crash date
Pass 3: DOB, FN, LN, Middle Initial (MI), zip
Pass 4: DOB, FN, LN, DLN, crash year
Pass 5: DOB, FN, LN, DLN, sex, street number, zip, crash date

98%! 5 Passes
Young driver crash rates by licensing age, driving experience, and license phase

Allison E. Curry\textsuperscript{a,b,*}, Melissa R. Pfeiffer\textsuperscript{a}, Dennis R. Durbin\textsuperscript{a,b,c}, Michael R. Elliott\textsuperscript{d,e}
NJ Traffic Safety Outcomes (NJ-TSO) Data Warehouse

- **Traffic Citations**: 2004-2014
- **Census data**: (2004-2014)
- **Birth certificate data**: 1979-2000
- **License data**: 2004-2014 (N=10 million)
- **Crash data**: 2004-2014 (N=6 million)
- **Death certificate data**: 2004-2014
- **Hospital discharge data**: 2004-2014
- **CHOP electronic health record data**: 1987-1997 (N=148,000)
### Crash-related citations

<table>
<thead>
<tr>
<th>Driver charge entry</th>
<th>Coded value</th>
</tr>
</thead>
<tbody>
<tr>
<td>“SPEEDING 61 MPH IN A 25 MPH?”</td>
<td>11/Speeding</td>
</tr>
<tr>
<td>“SPEEDING (70/50)”</td>
<td>11/Speeding</td>
</tr>
<tr>
<td>“CARELESS/UNSAFE”</td>
<td>17/Careless driving</td>
</tr>
<tr>
<td>“TOO MANY PASSENGERS/PROV. DL”</td>
<td>34/GDL passenger violation</td>
</tr>
<tr>
<td>“WRONG WAY ON ONE-WAY ST.”</td>
<td>1/Other moving violations</td>
</tr>
<tr>
<td>“PARKED IN NO PARKING ZONE BBB?”</td>
<td>88/All other violations</td>
</tr>
</tbody>
</table>
Compliance With and Enforcement of Graduated Driver Licensing Restrictions

Allison E. Curry, PhD, Melissa R. Pfeiffer, MPH, Michael R. Elliott, PhD

Introduction: Graduated Driver Licensing (GDL) is the most effective strategy to reduce the burden of young driver crashes, but the extent to which young intermediate (newly licensed) drivers comply with, and police enforce, important GDL passenger and night-time restrictions is largely unknown. Population-level rates of intermediate drivers’ compliance were estimated as well as police enforcement among crash-involved drivers who were noncompliant.
With Linked Citation Data

Enforcement higher among younger drivers

![Bar chart showing the proportion of GDL citations issued to younger drivers. The chart compares passenger enforcement and night enforcement, with ages ranging from 17 to 20 years old. The data indicates a higher proportion of citations among younger drivers.]
NJ Traffic Safety Outcomes (NJ-TSO) Data Warehouse

Traffic Citations 2004-2014

Birth certificate data 1979-2000

Death certificate data 2004-2014

Licensing data 2004-2014 (N= 10 million)

Crash data 2004-2014 (N= 6 million)

Census data

Hospital discharge data 2004-2014

CHOP electronic health record data 1987-1997 (N=148,000)
Young driver licensing: Examination of population-level rates using New Jersey’s state licensing database

Allison E. Curry\textsuperscript{a,b,*}, Melissa R. Pfeiffer\textsuperscript{a}, Dennis R. Durbin\textsuperscript{a,b,c}, Michael R. Elliott\textsuperscript{d,e}, Konny H. Kim\textsuperscript{a}
CHOP electronic health record data 1987-1997 (N=148,000)

Census data

Traffic Citations 2004-2014

Licensing data 2004-2014 (N= 10 million)

Birth certificate data 1979-2000

Hospital discharge data 2004-2014

Crash data 2004-2014 (N= 6 million)

Death certificate data 2004-2014

NJ Traffic Safety Outcomes (NJ-TSO) Data Warehouse

Birth certificate data

Hospital discharge data

Death certificate data

Traffic Citations

Licensing data

Crash data

Census data

Hospital discharge data
EHR Data Linkage

- True match rate (correct matches/original matches): 99.95%
- False non-match rate (correct matches not found): 1.5%

Pass 1: DOB, First Name (FN), Last Name (LN), zip
Pass 2: DOB, FN, LN, Middle Initial (MI), sex
Pass 3: DOB, FN, LN, MI, zip, city
Pass 4: DOB, FN, LN, MI, zip, city among twins
Pass 5: DOB, FN, LN, street number, zip
Pass 6: DOB, LN, street number, zip
With Linked Electronic Health Record Data

Motor Vehicle Crash Risk Among Adolescents and Young Adults With Attention-Deficit/Hyperactivity Disorder

Allison E. Curry, PhD, MPH; Kristina B. Metzger, PhD; Melissa R. Pfeiffer, MPH; Michael R. Elliott, PhD; Flaura K. Winston, MD, PhD; Thomas J. Power, PhD
With Linked Electronic Health Record Data

Teens with ADHD get licensed at lower rates and later

Figure 2. Inverse Kaplan-Meier Curves Depicting Cumulative Probability of Licensure by Sex and Attention-Deficit/Hyperactivity Disorder (ADHD) Status

<table>
<thead>
<tr>
<th>Age, y</th>
<th>Without ADHD</th>
<th>Females</th>
<th>Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.0</td>
<td>8323</td>
<td>3232</td>
<td>7542</td>
</tr>
<tr>
<td>17.5</td>
<td>3232</td>
<td>1931</td>
<td>1933</td>
</tr>
<tr>
<td>18.0</td>
<td>1931</td>
<td>1474</td>
<td>1503</td>
</tr>
<tr>
<td>18.5</td>
<td>1474</td>
<td>1194</td>
<td>1194</td>
</tr>
<tr>
<td>19.0</td>
<td>1194</td>
<td>996</td>
<td>1030</td>
</tr>
<tr>
<td>19.5</td>
<td>996</td>
<td>876</td>
<td>926</td>
</tr>
<tr>
<td>20.0</td>
<td>876</td>
<td>787</td>
<td>848</td>
</tr>
<tr>
<td>20.5</td>
<td>787</td>
<td>763</td>
<td>763</td>
</tr>
<tr>
<td>21.0</td>
<td>763</td>
<td>763</td>
<td>763</td>
</tr>
</tbody>
</table>

No. at risk
- Females without ADHD: 8323, 3232, 1931, 1474, 1194, 996, 876, 787, 763
- Males without ADHD: 7542, 3189, 1933, 1503, 1194, 1030, 926, 848, 763
- Females with ADHD: 691, 368, 265, 185, 149, 119, 104, 91, 80
- Males with ADHD: 1788, 1059, 722, 546, 445, 377, 328, 276, 246
Longitudinal study of driver licensing rates among adolescents and young adults with autism spectrum disorder

Allison E Curry¹, Benjamin E Yerys¹,², Patty Huang¹ and Kristi B Metzger¹
NJ Traffic Safety Outcomes (NJ-TSO) Data Warehouse

- Traffic Citations 2004-2014
- Birth certificate data 1979-2000
- Death certificate data 2004-2014
- Licensing data 2004-2014 (N= 10 million)
- Crash data 2004-2014 (N= 6 million)
- Census data
- Hospital discharge data 2004-2014
- CHOP electronic health record data 1987-1997 (N=148,000)
Incorporation of Additional Data Sources

• LinkSolv

• Combine data sets by year of birth
  • 1985 file: all license, driver, occupant, birth, death, hospital records of people born in 1985

<table>
<thead>
<tr>
<th>Source</th>
<th># of records</th>
</tr>
</thead>
<tbody>
<tr>
<td>License</td>
<td>140,000</td>
</tr>
<tr>
<td>Driver</td>
<td>133,000</td>
</tr>
<tr>
<td>Occupant</td>
<td>24,000</td>
</tr>
<tr>
<td>Birth</td>
<td>107,000</td>
</tr>
<tr>
<td>Death</td>
<td>1,000</td>
</tr>
<tr>
<td>Hospital</td>
<td>430,000</td>
</tr>
</tbody>
</table>

Example: Birth year 1985

• Single file match
  • Connect all records for same individual
## Linkage Data Elements

(P = Primary, S = Secondary)

<table>
<thead>
<tr>
<th>Data Element</th>
<th>License</th>
<th>Crash</th>
<th>Birth</th>
<th>Death</th>
<th>Hosp/ED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name (first, last, middle initial)</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Date of birth</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Geography of residence</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Sex</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Social Security Number</td>
<td>P</td>
<td>P</td>
<td></td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Date of death</td>
<td>P</td>
<td></td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Event date</td>
<td>P</td>
<td>P</td>
<td></td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Event time</td>
<td>P</td>
<td></td>
<td></td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Father's last name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Event location</td>
<td>S</td>
<td></td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>E-codes/seat position, vehicle type</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>Transport to hospital/hospital code</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>ICD-9 diagnosis codes</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>Marital status</td>
<td>S</td>
<td></td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Race and ethnicity</td>
<td>S</td>
<td></td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
</tbody>
</table>
Evaluation and Completion

- Visual inspection of each pass
- Estimate false positive rate (sum of false match probabilities for each pair / total number of pairs)
- Single imputed warehouse via maximum a posteriori estimation
Uses of Hospital Data

• Injury: crash report vs hospital data
  • Concordance
  • Misclassification

• Incidence of crash with injury

• Comorbid conditions
Uses of Birth Certificate Data

• Improve quality of race/ethnicity data
Uses of Death Certificate Data

• Improve quality of demographic data
  • Race/ethnicity
  • Marital status
  • Education

• Enhance injury information with crash-related deaths

• Adjust denominator
Analytic Data Elements

• Underlying cause of death and date of death (death)
• Motor vehicle crash-related injuries and deaths (hospital/ED, death)
• Injury severity data (hospital/ED)
• Economic costs associated with MVC-related hospitalizations (hospital/ED)
• Patient conditions diagnosed before MVCs (hospital, EHR)
• Geocoded residential address information (license, crash, perhaps hospital)
Population Subgroups

- License holders who did not crash
- People with diagnosed conditions, whether or not they crashed
- Unlicensed drivers who crashed in NJ
  - Including connected records for those who crashed >1 time
Potential for Growth
Replication in Other Jurisdictions

Potential Challenges

**Federal Drivers’ Privacy Protection Act of 1994, 18 U.S.C. §2721**
(b) Permissible Uses [of personal information from state motor vehicle records], (5) For use in research activities, and for use in producing statistical reports, so long as the personal information is not published, redisclosed, or used to contact individuals

**N.J.S.A 39:2-3.4c(5)**
For use in educational initiatives, research activities, and for use in producing statistical reports, so long as the personal information is not published, redisclosed, or used to contact individuals and, in the case of educational initiatives, only to organ procurement organizations as aggregated, non-identifying information.
Replication in Other Jurisdictions

Potential Challenges

- Legal barriers
- Administrative barriers
- Not intended for research purposes
- Non-matching identifiers
Acknowledgements
Funding

- National Science Foundation Center for Child Injury Prevention Studies
- National Institute of Child Health and Human Development R01HD079398, R03HD073248-02 and R21HD092820
- CHOP Foerderer Grant of Excellence
- State Farm Insurance Company
- AAA Foundation for Traffic Safety 4035-51129
- New Jersey Division of Highway Traffic Safety
- CHOP Pediatric Development Fund
Car Crash Research: Utilizing the New Jersey Traffic Safety Outcomes Data Warehouse

This line of car crash research at CIRP aims to advance traffic safety research and associated epidemiologic methods through novel administrative data linkages. Led by Allison E. Curry, PhD, MPH, the research team established a comprehensive data warehouse that includes the full licensing, citation, and crash history of every New Jersey driver between 2004 and 2014.

To provide further information about these NJ residents and drivers, we also linked zip-code level indicators from the Census and electronic health records for 150,000 pediatric patients of the CHOP Healthcare Network. We are also currently in the process of linking statewide birth certificate data, death certificate data, and hospital and emergency department discharge data.

The New Jersey Traffic Safety Outcomes (NJ-TSO) Data Warehouse allows us to fill numerous important gaps in car crash research. Crash reports contain data only on the events occurring just prior to the crash, the crash event itself, and the immediate conditions of those involved in the crash. While helpful, this information only covers a few minutes in the lifespan of a crash. The NJ-TSO Data Warehouse provides us with the ability to increase the study period from minutes to potentially decades. With these data, we can better study the effects of car crashes and to develop evidence-based tactics to prevent them.
Teendriversource.org

For families

For educators

For communities

We can all help keep teens safe on the road. Whether you're a teen, parent, policymaker, educator, or researcher, this site will support your efforts with free information and downloadable resources. Together we can make a difference.

How can we help you today? Pick the group that best describes you

1. I educate & support teens
2. I educate & support parents
3. I educate & support policymakers
4. I am a teen
5. I am a parent/guardian
6. I am a researcher

Enhancing Education To Keep Teens Safe On The Road

Teaching Techniques To Parents And Community Educators

Reference Guide

Teach Your Teen To Be A Smart Passenger

Set House Rules and Reduce Crash Risks For Your Teen

Ride Like A Friend

Don't ride with teens. Nearly half of teens who are killed in car crashes are killed in crashes with other teens. Ride like a friend, and let others know who's important. We'll all just saving a life-respect can save lives. But it can. Help us find solutions. And spread the word.

New Driver Headquarters

Just Chill

Just Chill

Just Chill

Just Chill

Just Chill

Just Chill
States with comprehensive GDL programs generally have lower fatality rates than those with less robust ones. Further reductions in teen crash rates will rely on strengthening specific provisions of GDL, as well as greater enforcement of and compliance with these provisions.

Recent research* from The Children’s Hospital of Philadelphia (CHOP) assessed the effect of decals (or vehicle identifiers) and suggests US states should consider decals as a potential way to enhance the effectiveness of their GDL programs.

Using a linked dataset of New Jersey’s (NJ) licensing and crash

http://injury.research.chop.edu/education-advocacy-tools/image-gallery#Teen Driver and Teen Passenger Safety Images
NJ-TSO Research Team
Allison E. Curry (PI),
Melissa R. Pfeiffer, Kristina B. Metzger, Meghan E. Kirk

Contact Information:
Melissa R. Pfeiffer, MPH
Center for Injury Research and Prevention
The Children's Hospital of Philadelphia

pfeifferm@email.chop.edu


