Building a MVC injury system of linked data: Lessons learned & questions answered about pedestrian injuries

January 26, 2022
Introductions

Dr. Anna E. Waller, ScD

Dr. Katherine (Katie) J. Harmon, PhD
Acknowledgments

Funding:
Funding was provided by the Centers for Disease Control and Prevention, the NC Governor’s Highway Safety Program, Collaborative Sciences Center for Road Safety, and NC Safe Routes to School.

Data attribution & disclaimer:
NC DETECT is a statewide public health syndromic surveillance system, funded by the NC Division of Public Health (NC DPH) Federal Public Health Emergency Preparedness Grant and managed through collaboration between NC DPH and UNC-CH Department of Emergency Medicine’s Carolina Center for Health Informatics. The NC DETECT Data Oversight Committee does not take responsibility for the scientific validity or accuracy of methodology, results, statistical analyses, or conclusions presented.
Outline

• Part I: Why Is Data Linkage Important?
  – Overview of Data Linkage in North Carolina
  – What Did We Learn from Data Linkage?
    • Building Relationships Across Organizations
    • Data Quality
    • Methodology
    • NC Transportation Safety & Public Health Data Dashboard
    • Selected Results from NC Pedestrian and Bicycle Crash Injury Data Linkage Study

• Part II: Challenges to Sustainable Data Linkage
  – Obtaining/Maintaining Stakeholder Support
  – Obtaining Funding
  – Technical Issues
Part I: Why is Data Linkage Important?

Brief Overview of Data Linkage in North Carolina
NC Data Linkage Objectives

Goal: Establish an integrated statewide MVC injury surveillance system

Integrated MVC and Health Information has the potential to:
- Improve safety outcome analysis and evaluation,
- Expand research activities,
- And inform policy and safety programs.

Why Link Crash Data with Other Data Sources?
Most data sources are limited in scope; by linking multiple data sources, we create a much richer dataset that can then be used to answer important questions.
Part I: Why is Data Linkage Important?

What Did We Learn from Data Linkage?

North Carolina Crash Injury Surveillance System (NC-CISS): Using linked data to understand the true impact of motor vehicle crashes

The goal of the project is to link crash report data with health outcome data to provide a more complete picture of the circumstances and outcomes associated with motor vehicle crash injuries in North Carolina.

We first looked at 2018 crash reports issued by law enforcement.

131,242 people were reported injured in crashes in 2018

However, when linked with emergency department data, 33,338 people were reported injured or died in the crash reports

Using crash data alone underestimates the number of people who seek healthcare for motor vehicle crash injuries.

Linking crash data and emergency department data provides a more realistic estimate of the burden of motor vehicle crash injury on our community.

For more information, go to the CCHI Transportation & Health Data website or if you have any questions, contact NC-CISS at cchi@unc.edu

North Carolina Data Integration for Motor Vehicle Crash Injury Research: The Long Road Ahead

Katherine J. Harmon, PhD
UNC Highway Safety Research Center
Katherine Petilos, PMP, MPS
Anna E. Wailer, ScD
UNC-Chapel Hill Center for Health Informatics

Background

Motor vehicle crashes (MVCs) are one of the leading causes of death and non-fatal injuries. CCRP people were killed and 1,333,157 people were non-fatally injured in North Carolina MVCs in 2018.

The North Carolina Traffic Records Coordinating Committee (NCTRCC) has an interest in a statewide MVC injury surveillance system. The ability to integrate safety information from a variety of sources has the potential to improve safety research and inform policy and safety programs.

Table 1: Results with Determinant Linkage

<table>
<thead>
<tr>
<th>Project Description</th>
<th>Data Sources</th>
<th>Linkage Fields Used</th>
<th>Results of Linkage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Exposure</td>
<td>External Data</td>
<td>Medical Hospital</td>
<td>Death certificate</td>
</tr>
<tr>
<td>Primary Exposure</td>
<td>External Data</td>
<td>Medical Hospital</td>
<td>Death certificate</td>
</tr>
<tr>
<td>Chief Complaint</td>
<td>Emergency Data</td>
<td>Medical Hospital</td>
<td>Death certificate</td>
</tr>
<tr>
<td>Chief Complaint</td>
<td>Emergency Data</td>
<td>Medical Hospital</td>
<td>Death certificate</td>
</tr>
<tr>
<td>Disposition</td>
<td>Emergency Data</td>
<td>Medical Hospital</td>
<td>Death certificate</td>
</tr>
<tr>
<td>Disposition</td>
<td>Emergency Data</td>
<td>Medical Hospital</td>
<td>Death certificate</td>
</tr>
</tbody>
</table>

Table 2: Project Details

<table>
<thead>
<tr>
<th>Project</th>
<th>Description</th>
<th>Data Sources</th>
<th>Linkage Fields Used</th>
<th>Results of Linkage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>Crash report</td>
<td>Crash report</td>
<td>Death certificate</td>
<td>Death certificate</td>
</tr>
<tr>
<td>Case 2</td>
<td>Crash report</td>
<td>Crash report</td>
<td>Death certificate</td>
<td>Death certificate</td>
</tr>
</tbody>
</table>

Recommendations

1. Add a unique variable to the crash report that identifies if the crash was reported to the NC-TVRC.
2. Include a unique personal identifier in all MVC data sources.
3. Improve capture of emergency medical treatment records.
4. NC-TVRC data needs to be standardized.

Conclusion

NC-TVRC can generate much more information with the data sources that are available for integration with MVC data. These health outcome data sources provide a more detailed overview of MVC injuries as compared to the crash report data.

https://cchi.web.unc.edu/transportation-health-data/
Collaboration Is Essential

• Project Staff
  – Investigators
  – Program managers
  – Statisticians
• Data owners
• Data users
  – State/Local departments of transportation
  – State/Local health departments
  – Investigators
• Community and advocacy groups
• Funders

North Carolina Department of Health and Human Services
  • Injury and Violence Prevention Branch
  • State Center for Health Statistics
  • Communicable Disease Branch

North Carolina Department of Transportation
  • Governor’s Highway Safety Program
  • Traffic Records Coordinating Committee

University of North Carolina – Chapel Hill
  • Carolina Center for Health Informatics
  • Highway Safety Research Center
  • Injury Prevention Research Center
Data Quality – MVC Data Sources Identified & Documented

- Crash Data (NC DMV)
- Pedestrian/Bicycle Crash Data (HSRC/NCDOT)
- EMS data (NC OEMS)
- Emergency Department Data (NC DETECT)
- Hospital Discharge Data (SCHS)
- Hospital Claims Data (UNC Sheps)
- North Carolina Trauma Registry (NCTR)
- BCBS/Medicaid Claims Data (UNC Sheps)
- Death Registration Data (SCHS)
- Medical Examiners Reports (OCME)
- Fatality Analysis Reporting System (NHTSA)
- Highway Safety Information System (FHWA)

Linkage Methods

We investigated four different linkage methodologies but focused on deterministic linkage.

<table>
<thead>
<tr>
<th>Linkage methods</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchical</td>
<td>Matches records using a set of pre-defined shared identifiers over multiple passes or “cascades”; allows some flexing with matching variables (age +/- 1 year)</td>
</tr>
<tr>
<td>deterministic linkage w/ fuzzy matching</td>
<td>Recursive partitioning trees Matches records using a calculated ‘distance’ between linkage variables</td>
</tr>
<tr>
<td>Probabilistic linkage</td>
<td>Matches records based on a pre-assigned probability that the match is correct (e.g. Linksolv)</td>
</tr>
<tr>
<td>Hand review</td>
<td>Matches records through manual review</td>
</tr>
</tbody>
</table>

Strengths:
- Easy to explain to a multi-disciplinary audience,
- High quality results,
- Fast,
- And replicable in many applications.

Challenge:
A sufficient & representative match rate.

The public health impact of pedestrian injury

Pedestrians are at an increased risk of severe injury and death, as compared to other road users. Pedestrians include any person not using a motorized vehicle in a public area. This definition includes people walking, using skateboards, roller skates, hoverboards and standing electric scooters, as well as persons using wheelchairs and assistive mobility devices. For a comprehensive definition of pedestrian, see the Safe States Report under Resources. Everyone is a pedestrian at some point in their day regardless of whether their trip mainly includes driving, cycling, or using public transit. Our data focuses only on pedestrians injured due to motor vehicle crashes; other types of pedestrian injuries are not included.

The societal burden of pedestrian injuries and fatalities remains substantial and data limitations continue to make surveillance difficult. The data provided here will assist local practitioners in determining trends and disparities among pedestrians injured across NC to support program development.

Pedestrian injury and fatality has been increasing since the late 2000s. Data below shows recent trends in pedestrian injuries and fatalities since 2016.

Average Number of Pedestrian Injuries Reported in NC crash and ED Visit Data, by Month: Oct 1, 2010 – Sept. 30, 2015

On average, >90 pedestrian injuries observed in the ED visit data, per month
Average Number of Bicyclist Injuries Reported in NC crash and ED Visit Data, by Month: Oct 1, 2010 – Sept. 30, 2015

On average, >580 bicyclist injuries observed in the ED visit data, per month
KABCO Does Not Always Provide an Accurate Assessment of Pedestrian Injury Severity

<table>
<thead>
<tr>
<th>Police assigned injury severity (KABCO)</th>
<th>Serious or fatal injury (based on clinical assessment)</th>
<th>Non-serious injury (based on clinical assessment)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>K: Killed</td>
<td>206 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>A: Disabling injury</td>
<td>437 (89%)</td>
<td>53 (11%)</td>
</tr>
<tr>
<td>B: Evident injury</td>
<td>1,431 (50%)</td>
<td>1,440 (50%)</td>
</tr>
<tr>
<td>C: Possible injury</td>
<td>488 (16%)</td>
<td>2,523 (84%)</td>
</tr>
<tr>
<td>O: No injury</td>
<td>20 (12%)</td>
<td>141 (88%)</td>
</tr>
<tr>
<td>Total</td>
<td>2,582 (38%)</td>
<td>4,157 (62%)</td>
</tr>
</tbody>
</table>
Pedestrian Injury Severity was Highest for Children and Older Adults

Frequency of serious pedestrian injuries, by age group: NC, 2010-2015

- **0-4**: 46% serious or fatal, 54% nonserious
- **5-9**: 43% serious or fatal, 57% nonserious
- **10-14**: 37% serious or fatal, 63% nonserious
- **15-19**: 39% serious or fatal, 61% nonserious
- **20-24**: 33% serious or fatal, 67% nonserious
- **25-34**: 36% serious or fatal, 64% nonserious
- **35-44**: 33% serious or fatal, 67% nonserious
- **45-54**: 40% serious or fatal, 60% nonserious
- **55-64**: 38% serious or fatal, 62% nonserious
- **65-74**: 48% serious or fatal, 52% nonserious
- **75+**: 54% serious or fatal, 46% nonserious
One Possible Explanation for the Higher Injury Severity Observed among NC Child Pedestrians was the Frequency of Head Injuries

24% of adults 25-44 years of age were diagnosed with head injuries.

39% of children 0-14 years of age were diagnosed with head injuries.

Male pedestrians (31%) and pedestrians struck at estimated speeds >35 MPH (40%) were also more likely to be diagnosed with head injuries, as compared to females (24%) and pedestrians struck at speeds ≤35 MPH (25%).

https://www.roadsafety.unc.edu/research/projects/2019r22/
Part II: Challenges to Sustainable Data Linkage
Challenges – Stakeholder Support

• **Obtaining stakeholder support** is rewarding but not always successful.
  – Accept that you will win some battles and lose others.

• **Maintaining stakeholder engagement** can also be challenging.

• While having a broad base of support has clear strengths, when you have many stakeholders, there may be a reluctance among stakeholders to feel *ownership* of the system.
Challenges - Funding

• Despite efforts to build a sustainable data linkage system, we have been **unable to secure continued funding**.
• Large data linkage systems are not conducive to annual, low-budget, proof-of-concept proposals. Support is needed for:
  – Staff with expertise in transportation safety, public health, statistics, and program management;
  – System maintenance and ongoing linkage efforts;
  – Computers and software;
  – The examination of specific research question;
  – And the dissemination of data and research products.
• To date, most states have been **unable (or unwilling) to pay** for ongoing, sustainable data linkage.
Challenges - Technical

• **Data acquisition** is a time-consuming process (weeks to months).

• Cleaning and linking large datasets is *methodologically complex*.
  – Requires considerable technical skills.
  – Linking more than one year of data can take hours.
  – A single change to a dataset structure can result in onerous changes to linkage methodology.

• Linking to **protected health information** impacts the analysis and dissemination of data.
  – Health data are typically aggregated to protect patients’ anonymity.
Despite Challenges, Data Linkage is Worth It

- Data linkage provides an opportunity to:
  - Build relationships;
  - Improve data quality;
  - Develop methodology and technical capacity;
  - Characterize the burden of pedestrian/bicycle crash injury across multiple data sources;
  - Evaluate pedestrian/bicycle safety decisions, programs/interventions, and policies;
  - Educate decision-makers and the public;
  - And address equity issues around safety and access.
Questions?

Contact information:
Anna Waller, ScD
Research Professor
Director, Carolina Center for Health Informatics
Department of Emergency Medicine
awaller@med.unc.edu

Contact information:
Katie Harmon, PhD
Research Associate
UNC HSRC
919.962.0745
harmon@hsrl.unc.edu