Motor Vehicle Crash and Health Outcome Data Integration Recommendations for North Carolina

August 20, 2020

Principal Investigator: Anna Waller, ScD

Co-Principal Investigator: Katherine Harmon, PhD, MPH

Project Manager: Erika Redding, MSPH

Co- Authors:

    Mike Dolan Fliss, PhD, MPS, MSW

    Seth LaJeunesse, CAGS, MCRP

    Katherine Peticolas, MPP

    Sharon Schiro, PhD

    Libby Thomas, MS

Background and Purpose

The North Carolina (NC) Governor's Highway Safety Program (GHSP), a program within the NC Department of Transportation (DOT), has the stated mission of "zero deaths on North Carolina roadways." As part of this mission, GHSP funded the Carolina Center for Health Informatics (CCHI), within the University of North Carolina (UNC) School of Medicine, to link health outcome data with police crash report data to improve motor vehicle crash (MVC) injury surveillance in the state. While this project has specifically focused on health outcomes data, it is important to note that our linkage work can also inform future crash linkage projects with other data sources such as road inventory data.

Since 2015, the MVC Injury Data Linkage Project has linked police crash report data with death certificate, EMS, trauma registry, emergency department (ED) visit, and hospital encounter data (combined ED visit and hospitalization data). Through these linkages, the project team has identified the following recommendations to data linkage and crash data collection and reporting that, if implemented, would result in an improvement to crash data quality, integration, and transportation safety research. The upcoming transition of the NC crash data system to a fully electronic crash (e-crash) data system provides a unique opportunity for implementing many of these recommendations.

For each linkage study and subsequent report, our team developed a list of recommendations regarding methods for improving future data linkage efforts and transportation safety research. The purpose of this report is to synthesize these recommendations. Below, is a consolidated list of key recommendations, as determined by our research team, to improve both data linkage and MVC research. These recommendations have been organized according to subject area.

Data Linkage Recommendations

1. **General Recommendations for the Improvement of Crash-Health Outcome Data Integration**

    1.1.   Include a ***common unique personal identifier*** on all crash, EMS, ED visit, hospital, death, and trauma records. Currently, no common unique personal identifier is collected on all relevant data sources. Having a single common identifier would improve the ease and accuracy of data integration. Researchers would be able to track a patient from crash until death or injury. Ideally, this unique personal identifier would not be tied to any personally identifiable information, such as driver's license number, social security number, or medical record number. Such an identifier would also enable self-linkage within data sources, such as ED visits to multiple hospital facilities by the same patient following the same or multiple crashes; a trauma transfer for major injury is an example of a situation where this would be useful.

1.2. Have a ***single organization perform all data linkages***. Ultimately, the aim of the MVC Injury Data Linkage Project is to develop a fully integrated MVC injury surveillance system. This can only be achieved with a central organization performing all linkages and serving as both a warehouse and as a "honest broker" for the linked data. Input and guidance from the data owners will be necessary to determine linkage methods, assess linkage quality, and ensure data security and de-identification rules post linkage are met.

1.3. Thoroughly and transparently ***document the methods*** used to perform data linkage. Documentation is extremely important for the purposes of evaluation and for informing others about the process used to link the data, including the rationale behind methodologic decisions. This is especially true when the personnel who performed the linkage are no longer available to provide explanation.

2. Recommendations for Changes to the NC Division of Motor Vehicles (DMV) 349 Crash Report Form

For the Improvement of Data Linkage

2.1. Add a ***unique identifier for all persons involved in a crash*** (see recommendation 1.1) to the NC crash report form. Currently, the NC DMV-349 form contains only a unique identification number for the crash. Researchers and data analysts must therefore create their own person-level identification using the method of their selection (often through some combination of the crash identification number, vehicle number, and person number). This can generate confusion when sharing crash data among persons and organizations. In addition, creating a unique personal identifier in the crash data that is ***common to other relevant data sources*** is the first step in developing a mechanism for being able to track individuals across multiple data systems, including healthcare data systems.

2.2. Add a ***yes/no variable*** to the NC crash report form to indicate if ***EMS responded to the scene***. Currently, EMS response is captured on the NC DMV-349 form by a free text variable; therefore, there is no uniformity in the manner in which the responding EMS agency is reported. In future linkage analysis, it would be beneficial to have a binary variable (yes/no) indicating whether or not EMS reported to the scene for every crash report.

2.3. Add a ***yes/no variable*** to the NC crash report form identifying if each person involved in a crash was ***treated/attended/transported by EMS***. While the variable described in 2.2 would indicate if EMS responded to the scene, we are currently still unable to tell which individuals were attended by EMS at any given crash because EMS response is coded at the crash-level instead of the individual-level. Coding at the individual level, in addition to the crash-level, would allow us to make distinctions between individuals.

2.4. Improve the ***collection of information*** of ***destination hospital*** for persons injured in motor vehicle collisions. The NC DMV 349 form contains a free text field for "Injured Taken by EMS to (47)". This variable identifies the hospital to which the

3

injured person is transported by EMS. This variable could be of use for future data linkages; however, at the present, this variable is often left blank. In addition, when police officers provide a destination hospital, there is considerable variation in hospital designation, including varying names and spellings for the same hospital. Additionally, providing a drop-down pick-list of North Carolina hospital options could further improve the quality and utility of this data element.

2.5.   Improve the *collection of information* for the variable "Zip" on the NC crash report form. The NC DMV 349 form contains a free text field for "Zip". This variable identifies the ***ZIP code of residence*** for the individuals involved in the motor vehicle collision(s). This is a key linkage variable. While this variable is rarely missing in the crash report data, it sometimes contains invalid, incomplete, or incorrect information, such as ZIP codes containing letters, less than five digits, or values outside of existing US postal codes. The ZIP code entered should be required to meet standard 5-digit ZIP format. Providing a drop-down pick-list of most likely ZIP code options (filtered by city or county to reduce the number of options) could further improve the quality and utility of this data element.

2.6.   Improve the *collection of information* for the variable "City" on the NC crash report form. The NC DMV 349 form contains a free text field for "City". This variable identifies the ***city of residence*** for the individuals involved in the motor vehicle collision(s). This is a key linkage variable. While this variable is rarely missing in the crash report data, there is considerable variation in city of residence designation, including misspellings and abbreviations. Providing a drop down pick-list of most likely City options could improve the quality and utility of this data element.

2.7.   Improve the *collection of information* for the variable "DOB" on the NC crash report form. The date of birth is a key linkage variable. Data entry errors could be prevented by denying the entry of future dates of birth or dates of birth that do not match the age. Commonly entered dates of birth used as a work-around to enter an unknown date of birth, such as 1/1/01, could also be verified.

2.8.   Improve the *collection of information* for the variable "Age" on the NC crash report form. Errors could be prevented by validating the age with the date of birth or by disallowing entry of ages less than 0 or greater than 120. Unusual ages such as 117 should be verified (3,498 crash records from 2018 had an age of 117). Allow ages of 0 for babies under 1 year old.

2.9.   ***Review the NC DMV 39 form at regular intervals*** (e.g. five-year intervals) for incorporating changes designed to improve crash data collection

2.10.   Add an *ethnicity variable* to the NC crash report form, instead of combining ethnicity with race.

For the Improvement of Transportation Safety Research

2.11. ***Add or modify an existing variable*** to incorporate additional information about ***drug impairment***. Currently, the NC DMV 349 form contains three variables related to alcohol/drug impairment: "Alcohol/Drugs Suspected (37)", "Alcohol/Drugs Test Status (38)", and "Alcohol/Drugs Test Results (39)". While these variables provide important information regarding suspected alcohol/drug involvement, far less information is collected for drugs other than alcohol. For drugs other than alcohol, common drug types (e.g. marijuana, cocaine, prescription opioids) should be included under "Alcohol/Drugs Test Results (39)" rather than in the crash narrative.

    2.11.1. There are additionally a number of other fields that could be updated as part of a comprehensive update of the DMV 349 form. For example, the occupant/non-motorist protection field could be expanded to include fields for different types of child restraints (rear-facing, forward-facing, booster seat, etc) so that correct use can be better determined. Any effort to update the DMV-349 should involve the various transportation task forces who could advise on their specific fields (Occupant Protection Task Force, Impaired Driving Task Force, Older Driver Committee, TRCC, etc).

2.12. ***Modify existing variable "Vehicle Style (Type) (41)"*** to incorporate ***new vehicle styles/types***. The NC DMV 349 form has not been updated for more than a decade; therefore, it does not accommodate new and emerging modes of transportation (i.e."micromobility" devices) such as electric scooters (e-scooters) and electric bicycles (e-bikes). In addition, there is a need for more thorough descriptions of common "pedestrian conveyances", such as wheelchairs, skateboards, roller skates, hoverboards, and other wheeled devices that may be used on trafficways.

    2.12.1. The Pedestrian and Bicycle Information Center (PBIC) is in the process of developing recommendations for categorizing micromobility devices in crash data, based on work performed by SAE International, the National Highway Traffic Safety Administration (NHTSA), the American National Standards Institute (ANSI), among other organizations. These recommendations will be released within the upcoming year by PBIC: http://www.pedbikeinfo.org

2.13. Add a ***yes/no variable*** to the NC crash report form to indicate if the vehicle involved was part of a ***motor vehicle rideshare***. There have been widespread reports of motor vehicle rideshare (e.g. Uber, Lyft) involvement in an increasing number of crashes, especially pedestrian crashes. However, in NC, it is nearly impossible to identify and monitor trends regarding motor vehicle rideshares. Therefore, a yes/no variable indicating that the vehicle involved was part of a commercial rideshare is needed.

2.14. ***Add a variable for county of residence*** to the NC crash report form. The current NC DMV 349 form contains a variable for county of crash, but no variable for the county of residence of persons involved in motor vehicle collisions. County of residence is a valuable variable for data linkage with health data, which is more likely to accurately report county of residence than ZIP code of residence.

3. Recommendations for Changes to Data Collected by the NC Office of EMS

For the Improvement of Data Linkage

3.1. Add a ***common unique identifier for all persons involved in a motor vehicle crashes to all EMS records*** (see recommendation 1.1)***.*** Currently, the data captured by NC OEMS contain a unique identifier; however, this identifier is not transferable to crash or other health outcome data sources. Therefore, the creation of a unique identifier common to all relevant data sources would improve the ease and accuracy of linkage. This unique identifier should be used for data integration purposes only and should not be tied to personally identifiable information, such as medical record number.

3.2. ***Harmonize key linkage fields (e.g. EMS agency name, hospital name, etc.)*** across crash-related data sources where possible, and if not, consider whether these core fields can be harmonized after the fact. Though many data systems will be unchanging for the foreseeable future, database designers should, when developing new systems or upgrading old ones, attempt to harmonize core linkage fields like demographic and personal identifiers. These fields can be used in lieu of or in supplement to common identifiers across these datasets and are easier to use when they are pre-harmonized or harmonization is possible.

For the Improvement of Transportation Safety Research

3.3. ***Improve quality*** of health outcome data captured by NC OEMS, including Glasgow Coma Scale score, Revised Trauma Score, and patient symptoms. We found that NC EMS data did not capture as many health outcome data elements useful for transportation safety research as anticipated. In addition, many of the health outcome data elements captured, such as Glasgow Coma Scale score, Revised Trauma Score, patient symptoms, etc. had high levels of missingness or contained vague/unclear responses (e.g. a patient symptom of "traumatic injury").

4. Recommendations for Changes to Data Collected by the NC Hospitals

4.1. Add a ***common unique identifier for all persons involved in a motor vehicle crashes to the hospital patient records*** (see recommendation 1.1)***.*** Currently, the data captured by NC hospitals contain a unique identifier; however, this identifier is not transferable to crash or other health outcome data sources. Therefore, the creation of a unique identifier common to all relevant data sources would improve the ease and accuracy of linkage. This unique identifier should be

used for data integration purposes only and should not be tied to personally identifiable information, such as medical record number.

5. **Recommendations for Changes to Data Collected by the NC Healthcare Association/NC DETECT**

   5.1. Add a ***common unique identifier for all persons involved in a motor vehicle crashes to the NC DETECT system*** (see recommendation 1.1), not included in the limited dataset*.* Currently, the data captured by NC DETECT for ED visits contain a unique identifier; however, this identifier is not transferable to crash or other health outcome data sources. Therefore, the creation of a unique identifier common to all relevant data sources would improve the ease and accuracy of linkage. This unique identifier should be used for data integration purposes only and should not be tied to personally identifiable information, such as medical record number.