

**North Carolina Linkage Study for Motor Vehicle Crashes Involving Pedestrians,
Bicyclists, and Motorcyclists: Description of Key Linkage Variables**

Erika M. Redding, MSPH¹
Katherine Harmon, PhD, MPH¹
Anna Waller, ScD²
Mike Dolan Fliss, PhD, MSW, MPH^{3,4}

Statement of Funding: Funding for this project was provided by the NC Governors Highway Safety Program.

Acknowledgements: Our team would also like to thank our project's advisors and staff, Kathy Peticolas, Clifton Barnett, Dennis Falls, and Amy Ising from the Carolina Center for Health Informatics; Steve Marshall from the UNC Injury Prevention Research Center; Bevan Kirley, Nancy Lefler, Eric Rodgman, and Laura Sandt from the UNC Highway Safety Research Center; and Alan Dellapenna from the NC Division of Public Health.

¹ University of North Carolina Highway Safety Research Center

² Carolina Center for Health Informatics, Department of Emergency Medicine, University of North Carolina School of Medicine

³ University of North Carolina Injury Prevention Research Center

⁴ North Carolina Department of Health and Human Services, Injury Prevention Branch

Table of Contents

Background	3
Purpose	3
Overview of Linkage Methods.....	Error! Bookmark not defined.
Key Linkage Variables	3
Table 1. Key linkage variables across all crash linkage pilot projects	4

Background

The North Carolina (NC) Governor's Highway Safety Program (GHSP), a program within the NC Department of Transportation (NC DOT), has the stated mission of "zero deaths on North Carolina (NC) roadways." As part of this mission, in 2016, the GHSP funded the Carolina Center for Health Informatics (CCHI) at the University of North Carolina at Chapel Hill (UNC-CH) to link health outcome data with police crash report data to improve motor vehicle crash (MVC) injury surveillance in the state. In year 1 of the project, we organized a group of key data owners and users as part of our MVC Injury Stakeholders Group and performed an [evaluation](#) of previously linked police-reported crash (crash) data with NC emergency medical services (EMS data). In years 2-4, we performed several demonstration projects linking NC crash data, managed by the UNC Highway Safety Research Center (HSRC), with health outcome data sources. These health outcome data sources included [hospital encounter](#) (emergency department and inpatient encounters), [emergency department visit](#), trauma registry, and Medicaid claims data.

Purpose

In this report we have developed a consolidated list of all key linkage variables used in all linkage projects performed as part of the GHSP-funded MVC Injury Data Linkage Project. This list can be used by researchers and practitioners that wish to duplicate our data linkage results as well as identify integral components of each dataset for linkage.

Included Linkage Projects and Variables

In this report we have included the key linkage variables from the following linkage pilot projects:

- 2017 North Carolina Health Care Association (NHCA) and HSRC crash data, pedestrians and bicyclists, only (funding year 2)
- 2017 NC DETECT emergency department (ED) data and HSRC crash data, pedestrians and bicyclists, only (funding year 3)
- 2018 NC Trauma Registry and HSRC crash data, all road user types (funding year 4)
- 2018 NC Medicaid Claims and HSRC crash data, pedestrians, bicyclists, and motorcyclists, only (funding year 4)

Overview of Linkage Methods

All studies used deterministic linkage methods, with the emergency department, trauma registry, and Medicaid linkages using hierarchical deterministic/cascade linkage methods. The trauma registry linkage also used "fuzzy" cleaning to facilitate linkage. More detailed descriptions of linkage methodologies are provided in the full study reports posted on the CCHI Transportation & Health Data [website](#).

A comprehensive list of all variables used for linkage can be found in **Table 1**.

Table 1. Comprehensive list of key linkage used for GHSP funded linkage projects

Data Source	Project Date	Variable Type	Variable Definition
North Carolina Health Care Association (NCHA)	August, 2019	Sex	Crash victim sex
		Age	Age of crash victim in years
		Zip Code	First 5 digits of crash victim's zip code of residence
		Street Address (first five characters)	First 5 characters of crash victims street address
NC DETECT Emergency Department (ED) Visit Data	August, 2019	Date of Birth	Crash victim date of birth
		5-digit Zip Code of Residence	First 5 digits of crash victim's zip code of residence
		Age	Crash victim age
		Sex	Crash victim sex
		ICD-10-CM injury mechanism codes	Code for a pedestrian/bicycle crash injury and a chief complaint keyword of a pedestrian/bicycle crash.
North Carolina Trauma Registry	Project is Ongoing	Age	Non-numeric value of crash victim age
		Date of Birth	Crash victim date of birth
		Gender	Crash victim gender
		Race and Ethnicity	Crash victim race/ethnicity
		Injury Existence	Identified existence of injury for crash victim
		Injury Severity (True/False)	Severity of injuries experienced by crash victims
		Crash (True/False)	Is the injury due to a crash or is the event a crash
		Position in Vehicle	Seating position of victim at time of crash
		Residence state	Crash victim state of residence
		Residence County	Crash victim county of residence
		Residence Zip	All five or first three of residential zip code
		Residence City	Crash victim city of residence
		Crash/Injury State	State in which crash injury occurred
		Crash/Injury County	County in which crash injury occurred
		Crash and Injury dates	Month and day of crash/injury date
		Crash and Hospital admission date	Crash victim hospital admission date
		Hospital	Harmonized facility names. Compared to both facility on record and referring facility names
North Carolina Medicaid Claims Data	September, 2020	Personal identification number*	This includes crash person-identification number (HRSC data) Medicaid identification number (Medicaid data)
		Crash Case Number*	The case number associated with the involved crash
		Date of Birth*	Crash victim date of birth
		Address*	Crash victim address including city, zip code, and state

* Variables were used for linkage but removed from analytic dataset for the protection of potentially identifiable personal information