North Carolina Linkage Study for Motor Vehicle Crashes Involving Pedestrians and Bicyclists

Police Crash Report Data & NC DETECT Emergency Department Visit Data

Katherine J. Harmon, PhD[1], Katherine Peticolas, PMP[2], and Anna E. Waller, ScD[2]

August 2019

[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

Funding

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

Acknowledgements

We would like to acknowledge our Project Team members: Clifton Barnett, Alan Dellapenna, Jr., Dennis Falls, Tony Fernandez, Seth LeJeunesse, Amy Ising, Nancy Lefler, Steve Marshall, Eric Rodgman, Laura Sandt, Sharon Schiro, and Libby Thomas. In addition, we would like to acknowledge Lana Deyneka & Zach Faigen from the Communicable Disease Branch of the North Carolina Division of Public Health who provided permission for the NC DETECT emergency department visit data used in the data linkage.

Data Attribution & Disclaimer

NC DETECT is a statewide public health syndromic surveillance system, funded by the 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.

Table of Contents

Table of Tables

Table of Figures

List of Abbreviations

CCHI	Carolina Center for Health Informatics
DMV	Division of Motor Vehicles
DOT	Department of Transportation
ED	Emergency Department
GHSP	Governor's Highway Safety Program
HSRC	Highway Safety Research Center
MVC	Motor vehicle crash
NC	North Carolina
NC DETECT	North Carolina Disease Event Tracking and Epidemiologic Collection Tool
NC DMV	North Carolina Division of Motor Vehicles
NC DOT	North Carolina Department of Transportation
NC DPH	North Carolina Division of Public Health
NHTSA	National Highway Traffic Safety Administration.
UNC	University of North Carolina

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 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 MVC (motor vehicle crash) injury surveillance in the state. While the ultimate objective of the MVC Injury Data Linkage Project is to improve injury surveillance for all MVCs, most of our initial project activities have centered on pedestrian and bicycle crashes. We selected these types of MVC injuries for the following reasons:

1) The incidence of NC pedestrian fatalities has increased over the last five years and is of special concern to many of our stakeholders;
2) Project Team expertise;
3) The 2,500 annual pedestrian/bicycle crashes represent a more manageable number of records for linkage (as compared to ~275,000 total annual MVCs);
4) And since pedestrians and bicyclists are more likely to be injured, they are more likely to seek treatment at an emergency department than motor vehicle occupants.

Purpose

In this report, we describe a pilot project in which we linked police crash report and NC DETECT emergency department visit data. In Aim 1, we describe the methods used to link police-reported pedestrian and bicycle crash records to emergency department visit records for the 2017 calendar year. In Aim 2, we briefly describe the results of the data linkage between these two data sources. The aims for this project are listed below.

Aims

Aim 1: Describe the methods used to link statewide police crash report and NC DETECT emergency department visit data for pedestrian/bicycle crash injuries.

Aim 2: Describe the results of the pedestrian/bicycle crash-emergency department visit data linkage.

Aim 1 - Methods Used to Link Statewide Police Crash Report and NC DETECT Emergency Department Visit Data

Data Sources

We received 2017 pedestrian/bicycle crash injury data from two separate sources: NC police crash reports and NC DETECT. Table 1 describes the sources used in the data linkage study.

We received the 2017 police crash report data file from the UNC Highway Safety Research Center (HSRC). The police crash report data are owned by the NC Division of Motor Vehicles (NC DMV). Each year, NC DMV provides a copy of the crash data file to UNC HSRC (excluding the names of the persons involved in the crash). UNC HSRC pulled all records in which the crash report indicated that a pedestrian or bicyclist was involved in the crash. We used a broad case definition to capture as many pedestrian and bicycle crashes as possible. Typically, a team within UNC HSRC codes all pedestrian/bicycle crashes; however, this process is labor intensive, and the 2017 pedestrian/bicycle crash data file was not available at the time of linkage. Therefore, the pedestrian/bicycle crash data file used in our data linkage analyses may contain misclassified cases. Upon release, please refer to the HSRC data for the "official" 2017 total of pedestrian/bicycle crashes in NC (http://www.pedbikeinfo.org/pbcat_nc).

Prior to linkage, we cleaned and processed the crash data. For several records, the crash report coded the crash as both pedestrian and bicycle related. For these crashes, we coded the crash as involving a bicycle, rather than a pedestrian. In addition, for each crash, only one report is produced. Therefore, we had to disaggregate information for all individual pedestrians and cyclists. From 4,069 crash reports, we abstracted information for 4,241 individual pedestrians and cyclists.

We received the 2017 emergency department visit data file from the NC Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT). NC DETECT is owned by the NC Division of Public Health (NC DPH) and operated in collaboration with CCHI. NC DETECT collects emergency department visit data from all 24/7 acute-care hospital-affiliated civilian emergency departments in North Carolina. In 2017, 126 emergency departments submitted data to NC DETECT. These data are made available for public health surveillance as part of a statewide mandate (§130A-480) enacted by the NC General Assembly in 2005. To learn more about NC DETECT, visit the following website: https://ncdetect.org. We obtained all NC DETECT emergency department visits with one or more ICD-10-CM injury diagnosis codes and/or injury mechanism codes (ICD-10-CM codes that started with a "S", "T", "V", "W", "X," or "Y"). We did not restrict our analyses to visits with bicycle

and pedestrian injury mechanism codes. In NC, injury mechanism codes are not mandated by state law and there is a large variation in reporting across facilities. The 2017 NC DETECT emergency department visit file contained 1,077,925 visits that met our criteria (21% of all 2017 ED visits in NC DETECT).

Table 1. Description of Data Sources Used in Crash-Emergency Department Data Linkage Pilot Project

Data	Description	Source
Police Crash Reports[1,2]	All 2017 NC crash reports for MVCs involving pedestrians/bicyclists.	NC DMV (provided by UNC HSRC)
NC DETECT emergency department visit data	All 2017 emergency department visits containing an injury diagnosis code and/or an injury mechanism code in *any* position	NC DETECT/NC DPH

Abbreviations: NC, North Carolina; UNC, University of North Carolina; MVC, motor vehicle crash

[1]Pedestrian/bicyclist involved MVCs were identified using the variables "Person Type" (Field 22 on the NC DMV 349 form), "Vehicle Type" (Field 41 on the NC DMV 349 form), "First Harmful Event (Field 52 on the NC DMV 349 form), and "Most Harmful Event" (Field 56 on the NC DMV 349 form). For crash reports that characterized the collision as involving both a pedestrian and a bicyclist, the record was classified as involving a bicyclist.

[2]In NC, police are only required to complete a crash report if the crash occurred on a publicly maintained road or public vehicle access road (e.g. "traffic-related"); however, for pedestrian/bicycle crashes, police sometimes complete reports that do not meet these criteria. We included all records regardless of whether the crash met the criteria of being "traffic-related".

Exclusions

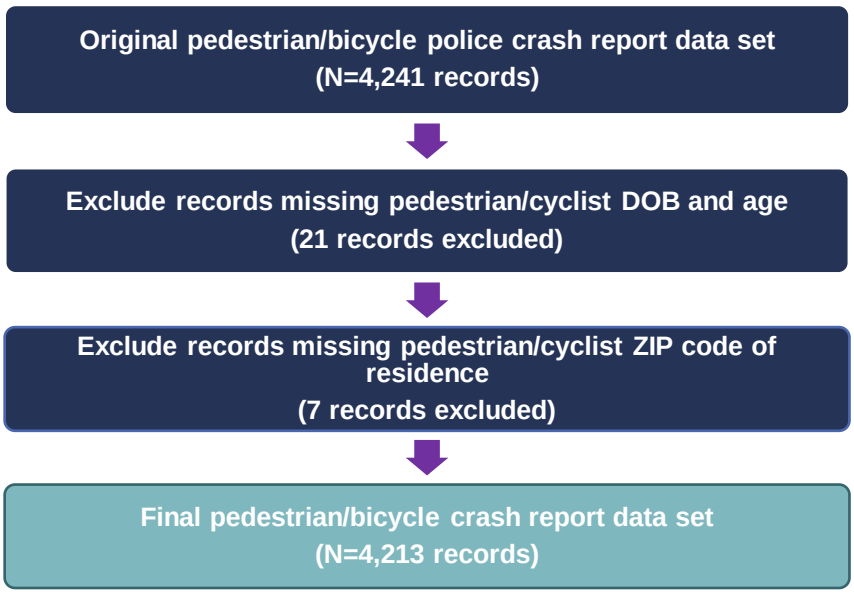
Prior to linkage, we excluded all crash records and emergency department visits that were missing values for variables that we deemed critical for linkage.

Police Crash Report Data Exclusions:

We excluded all records missing values for the following variables:

1. Date of birth AND age
2. And 5-digit ZIP code of residence.

Figure 1. Police Crash Report Data Exclusions



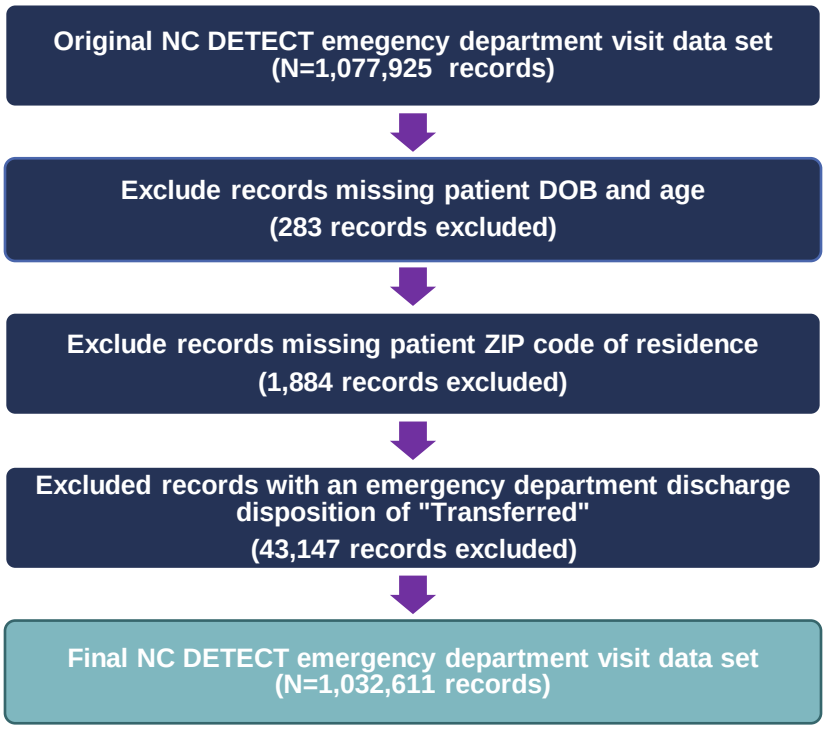
NC Emergency Department Visit Data Set Exclusions:

We excluded all records missing values for the following variables:

1. Date of birth AND age
2. And 5-digit ZIP code of residence.

In addition, we excluded all emergency department visits with a discharge disposition of "transferred". In NC DETECT emergency department visit data, we can only track individual patients returning to the same hospital or, in some instances, returning to a different hospital within the same healthcare system, using the Internal Tracking Identification Number (a unique identification number created by CCHI to track patients anonymously). Therefore, when patients are transferred to different hospitals and healthcare systems, it is unlikely that we can track these patients across facilities. In order to reduce the likelihood of one-to-multiple matches, we decided to exclude transfers under the assumption that we would likely capture the visit at the terminal facility.

Figure 2. NC DETECT Emergency Department Visit Data Exclusions



Linkage Algorithm

For the data linkage, we used hierarchical deterministic methods. Hierarchical deterministic linkage matches data according to a list of predefined variables in a stepwise fashion. For a match to occur, the two data sources must have the exact same values for the linkage variables. However, if a match does not occur during the first linkage step, certain linkage criteria are relaxed (while other linkage criteria are tightened) and a second round of matching commences. If matching does not occur during the second round of matching, more rounds of matching can be performed. For this pilot project, we performed two rounds of linkage. For both rounds of linkage, we limited the time window from the date/time of crash to the date/time of emergency department visit to 336 hours (14 days). For the first round of linkage, we linked records based on patient date-of-birth and 5-digit ZIP code of residence.

<u>Linkage Variables (Exact Match)</u>

1. Date-of-Birth
2. And 5-digit ZIP code of residence.

Other Criteria

3. Emergency department visit date/time occurred after crash date/time
4. And emergency department visit date/time less than or equal to 336 hours (14 days) after crash.

During the first round of data linkage, we linked 2,013 pedestrian/bicycle crash records to NC DETECT emergency department records.

For the second round of linkage, we used age rather than patient date-of-birth for the data linkage. Since patient age is less specific than patient date-of-birth, we only linked crash records to NC DETECT emergency department visit records with a pedestrian/bicycle crash injury ICD-10-CM injury mechanism code (Table 2) and/or a pedestrian/bicycle crash injury keyword located in the Chief Complaint (Table 3). In addition, records also had to match on patient sex.

Linkage Variables (Exact Match)

1. Sex,
2. Age,
3. And 5-digit ZIP code of residence.

Other Criteria

5. Emergency department visit had to contain a pedestrian/bicycle crash injury mechanism code and/or keyword,
6. Emergency department visit date/time occurred after crash date/time,
7. And emergency department visit date/time less than or equal to 336 hours (14 days) after crash.

During the second round of data linkage, we linked an additional 61 records.

Table 2. Pedestrian/Bicycle Crash Injury ICD-10-CM Injury Mechanism Codes[1]

ICD-10-CM Code	Description
Pedestrian ICD-10-CM Codes	
V00 (.01-.898)	Pedestrian conveyance accident
V01 (.00-.99)	Pedestrian injured in collision with pedal cycle
V02 (.00-.99)	Pedestrian injured in collision with two- or three-wheeled motor vehicle
V03 (.00-.99)	Pedestrian injured in collision with car, pick-up truck or van
V04 (.00-.99)	Pedestrian injured in collision with heavy transport vehicle or bus
V05 (.00-.99)	Pedestrian injured in collision with railway train or railway vehicle
V06 (.00-.99)	Pedestrian injured in collision with other nonmotor vehicle
V09 (.00-9)	Pedestrian injured in other and unspecified transport accidents
Cyclist ICD-10-CM Codes	
V10 (.0-.9)	Pedal cycle rider injured in collision with pedestrian or animal
V11 (.0-.9)	Pedal cycle rider injured in collision with other pedal cycle
V12 (.0-.9)	Pedal cycle rider injured in collision with two- or three-wheeled motor vehicle
V13 (.0-.9)	Pedal cycle rider injured in collision with car, pick-up truck or van
V14 (.0-.9)	Pedal cycle rider injured in collision with heavy transport vehicle or bus
V15 (.0-.9)	Pedal cycle rider injured in collision with railway train or railway vehicle
V16 (.0-.9)	Pedal cycle rider injured in collision with other nonmotor vehicle
V17 (.0-.9)	Pedal cycle rider injured in collision with fixed or stationary object
V18 (.0-.9)	Pedal cycle rider injured in noncollision transport accident
V19 (.0-.9)	Pedal cycle rider injured in other and unspecified transport accidents
Y93.55	Activity at time of health condition – bike riding

[1]The NC DETECT emergency department visit data set used in this pilot project contained up to 37 fields for ICD-10-CM codes; we considered an emergency department visit to be pedestrian/bicycle crash injury-related if it contained one or more injury mechanism codes in any one of the 37 available fields.

Table 3. Pedestrian/Bicycle Crash Injury Keywords

Pedestrian/Cyclist Keywords in Chief Complaint
Pedestrian Keywords
Contains: 'PEDESTRIAN', 'PED STRUCK', 'PEDS STRUCK', 'PED VS MVC' , ' PEDS VS MVC' , 'PED VS CAR' , 'PEDS VS CAR' ,'HIT BY CAR' , 'STRUCK BY VEHICLE' , 'STRUCK BY CAR' , 'PEDVSCAR' , 'PEDESTRAIN' , 'HIT BY VEHICLE' , 'RAN OVER BY CAR' , 'PEDESTRIAN STRUCK' , 'PED VS TRUCK' , 'PEDS VS TRUCK' , 'PED VS MOTORIZED VEHICLE' , or 'PEDS VS MOTORIZED VEHICLE' **And Not:** 'MOPED', 'SCOOTER', 'PEDAL', 'BICYCLE', or 'BIKE
Cyclist Keywords
Contains: 'BICYCLE', 'BIKE', ' PEDAL', 'BICYCLE ACCIDENT', or 'BICYCLIST' **And Not:** MOTOR CYCLIST', 'SCOOTER', 'MOTORCYCLE', 'PEDAL PULSE', 'PEDAL EDEMA', 'PEDAL PULSES', 'MOPED', 'DIRT BIKE', 'MOTOR BIKE', 'MOTORBIKE', 'DIRTBIKE', 'CAR OR BIKE', or 'PEDESTRIAN'

Results of Initial Linkage

We matched 2,074 crash and emergency department visit records. However, 196 crash records matched to multiple hospital encounters. Of these, 103 crash records matched to more than one emergency department visit record because the patient returned to the same healthcare system within the 14-day study period. Since the return visit was within the study period, we assumed that this return visit was related to the original pedestrian/bicycle crash. Therefore, we retained these visits in our analysis data set.

The remaining 93 crash records that matched to two or more emergency department visit records did not appear to be return or follow-up visits. For these matches, it appeared that a single crash record was linking to two or more emergency department visits made by two or more individual patients with different Internal Tracking Identification Numbers. After performing a manual review of these one-to-multiple matches, we determined that most of these matches fell into one of the following two scenarios:

1. The patient returned to a different hospital for follow-up medical care, so received a different Internal Tracking Identification Number (as mentioned previously, we can only track patients that return to the same hospital system in NC DETECT)
2. Or the patient was transferred from one hospital to another hospital, so received a different Internal Tracking Identification Number. While we had assumed that excluding all patient records with an emergency department disposition of "Transferred' would have eliminated this problem, it appears that not all emergency department visit records have the correct disposition (based on our review).

For these one-to-multiple matches, our first approach was to select the emergency department visit record closest in time to the actual crash; however, we altered our strategy when we realized that sometimes the second visit contained more detailed patient information. Therefore, we created a strategy designed to capture the record with the most information from one-to-multiple matches. For each record, we created an indicator variable (0/1) for time since crash ($\leq$6 hours), presence of bicycle/crash injury mechanism code, presence/absence of pedestrian/bicycle crash injury keyword in chief complaint, and number of diagnosis codes (>4 codes; Table 4). We then summed the "0's" and "1's" across the indicator variables and selected the record with the greater value in each one-to-multiple grouping. For one-to-multiple groupings in which the records contained the same values, we selected the record closest in time to the crash. In the examples provided in Figure 3, we would select record "A" from the first and second pairs.

There were no matches in which an emergency department visit matched to more than one crash record.

Table 4. Selection Criteria for One-to-Many Matches

Variable	Indicator Definition
Time since crash	If time since crash less than or equal to 6 hours, then Indicator Variable 1 = 1; Else Indicator Variable 1 =0;
Presence of pedestrian/bicycle crash injury keyword in chief complaint	If chief complaint contains a pedestrian/bicycle crash injury keyword, then Indicator Variable 2 = 1; Else Indicator Variable 2 = 0;
Presence of pedestrian/bicycle crash injury ICD-10-CM injury mechanism code	If one or more diagnosis code fields contains a pedestrian/bicycle crash injury mechanism code, then Indicator Variable 3 = 1; Else Indicator Variable 3 = 0;
Number of ICD-10-CM diagnosis codes	If number of diagnosis codes greater than 4 (median number of diagnosis codes), then Indicator 4 = 1; Else Indicator 4 = 0;
Indicator variable "total" (Indicator 5)	Sum of Indicator Variables 1-4

Figure 3. Example of One-to-Many Record Selection[1]

Pair 1

	Internal Tracking ID	Time Since Crash (Hours)	Chief Complaint	Dx 1	Dx 2	Dx 3	Dx 4	Dx 5	Ind 1	Ind 2	Ind 3	Ind 4	Ind 5
A	123	1	FELL OF BIKE	V13.4XXA	S00.00XA	S00.30XA	S80.00XA		1	1	1	0	3
B	345	8	LEG PAIN	S72.301A	F10.920	F41.9			0	0	0	0	0

Pair 2

	Internal Tracking ID	Time Since Crash (Hours)	Chief Complaint	Dx 1	Dx 2	Dx 3	Dx 4	Dx 5	Ind 1	Ind 2	Ind 3	Ind 4	Ind 5
A	678	1	MVC VS PEDESTRIAN	V04.19XA	S36.119A				1	1	1	0	3
B	912	2	PEDESTRIAN	S36.116A	S30.1XXA	S30.0XXA	S36.031A	I95.9	1	1	0	1	3

[1]The information presented in this figure was created by the authors for illustrative purposes; the information presented does not reflect actual patient information.

Aim 2: Results of the Pedestrian/Bicycle Crash-Emergency Department Visit Data Linkage

Description of Linked Pedestrian/Bicycle Crash-NC DETECT Emergency Department Visit Data Set

Our final linked pedestrian/bicycle crash-NC DETECT emergency department visit data set contained 1,972 observations, of which 1,870 observations were incident emergency department visits (Table 5). We linked 44% of pedestrian/bicycle crash records to NC DETECT emergency department visit records. The proportion of records that linked was slightly higher for pedestrians (45%) than for cyclists (43%).

Table 5. Frequency of Linked NC DETECT Pedestrian/Bicycle Crash Injury-Related Emergency Department Visits

Number of linked NC DETECT emergency department visits during 14-day study period	Frequency	
	N	%
One	1,870	94.8%
Two	97	4.9%
Three	5	0.3%
Total Number of Visits	**1,972**	**100.0%**

Manual Review of Linked Pedestrian/Bicycle Crash-NC DETECT Emergency Department Visit Data Set

We manually reviewed 100 linked crash-emergency department visit records to assess the accuracy/validity of the linkage. We considered a linkage to be a "true match" if the emergency department visit contained a keyword and/or an injury mechanism code for a pedestrian/bicycle crash injury. We considered a linkage to be a "false match" if the emergency department visit record contained keywords and/or injury mechanism codes for injury mechanisms other than pedestrian/bicycle crash injuries. We categorized a linkage as an "unknown match" if the linked emergency department visit record contained a non-specific chief complaint and was missing an injury mechanism code or contained a non-specific injury mechanism code (such as V87.7XXA-Person injured in collision between other specified motor vehicles [traffic], initial encounter). For

records that we classified as “false” or “unknown matches”, we reviewed the record through the NC DETECT web portal to obtain additional information about the emergency department visit. This step was important, for two records classified as falls contained information in the triage note indicating that the visit was related to pedestrian/bicycle crash injuries. After our review, we categorized 87 records as “true matches”, 1 record as a “false match”, and 12 records as “unknown matches”. The one “false match” was for treatment for a tick bite.

We also attempted to locate 100 unlinked crash records in emergency department visit data through the NC DETECT web portal. We could have proceeded in the opposite direction (unlinked pedestrian/bicycle crash injury-related emergency department visits to crash records), but this would have been much more challenging given the lack of detail about the crash provided in the emergency department visit records.

Table 6 displays the results of this review. We located about a third of unlinked pedestrian/bicycle crash records in NC DETECT. For over one-half of the unlinked records located in NC DETECT, the reason for non-linkage was related to differences in ZIP codes between the two data sources. However, most of these non-linkages matched on City of Residence (we did not include City of Residence in our NC DETECT emergency department visit data set). For future linkages between police crash report and NC DETECT emergency department visit data, we recommend including City of Residence in the emergency department visit data set and adding City of Residence to the linkage algorithm. For the remaining non-linkages, the most common reason was a lack of injury diagnoses in the emergency department visit records. Most of these records contained diagnosis codes for pain (e.g. M79.661-Pain in right lower leg) or contained no diagnosis codes at all (common with patients who “Leave without Medical Advice”). Since we only requested NC DETECT emergency department visits with one or more injury diagnosis codes, these visits were not included in our emergency department visit data set. However, most of the emergency department visits missing injury diagnoses contained one or more pedestrian/bicycle crash injury keywords in the Chief Complaint or Triage Note. For future linkages, we recommend including pedestrian/bicycle crash injury keywords as part of the initial NC DETECT emergency department visit data pull.

Table 6. Results of Manual Review of Unlinked Pedestrian/Bicycle Crash Records

KABCO (Injury Severity)	Located in NC DETECT[1]				
	Yes		No		Total
	n	%	n	%	n
Killed (**K**)	1	14.3%	6	85.7%	7
Serious Injury (**A**)	2	50.0%	2	50.0%	4
Minor Injury (**B**)	15	48.4%	16	51.6%	31
Possible Injury (**C**)	11	25.6%	32	74.4%	43
No injury (**O**)	1	8.3%	11	91.7%	12
Unknown	1	33.3%	2	66.7%	3
Total	**31**	**31.0%**	**69**	**69.0%**	**100**

1Located through the NC DETECT web portal, not the NC DETECT data set used in analyses. Data accessed through the NC DETECT web portal contain more patient information, some of which is sensitive and is not shared with researchers.

Based on our manual record reviews, we calculated sensitivity and specificity (among other evaluation statistics). Our current linkage algorithm is highly specific but lacks sensitivity (Table 7). It is likely that by incorporating the recommendations described in the preceding paragraph, we will be able to improve sensitivity. Based on our linkage algorithm evaluation statistics, we feel that our linkage algorithm was moderately successful.

Table 7. Match Status and Linkage Algorithm Evaluation Statistics

Match Status			
	n		n
True Positive (a)	87	False Positive (c)	1
False Negative (b)	31	True Negative (d)	69

Measure	Formula	Statistic	95% CI
Sensitivity	$\frac{a}{a+b}$	73.7%	(64.8%-81.4%)
Specificity	$\frac{d}{c+d}$	98.6%	(92.3%-100.0%)
Positive Predictive Value (PPV)	$\frac{Sensitivity}{1-Specificity}$	98.9%	(93.8%-100.0%)
Negative Predictive Value (NPV)	$\frac{1-Sensitivity}{Specificity}$	69.0%	(59.0%-77.9%)
Accuracy	$\frac{(a+d)}{(a+b+c+d)}$	83.0%	(76.8%-88.1%)
Cohen's Kappa[1,2]	$\frac{(p_o-p_e)}{(1-p_e)}$	0.67	(0.56-0.77)

[1] p_o =Observed agreement (identical to accuracy).

[2] p_e = Probability of chance agreement.

Tables 8 and 9 display the results of the crash-NC DETECT emergency department visit linkage for pedestrians and cyclists, respectively. For both tables, the columns contain summary statistics for unlinked pedestrian/bicycle crash records, linked pedestrian/bicycle crash- emergency department visit records, and unlinked pedestrian/bicycle crash-injury related emergency department visits. Rather than display the summary statistics for all NC DETECT emergency department visit records that did not link, we selected only emergency department visits with a pedestrian/bicycle crash injury mechanism or keyword. We selected this subset because it is likely to be more comparable to pedestrian/bicycle crash records than the full data set. In addition, the NC DETECT emergency department visit subset provides some indication of the number of pedestrians and cyclists who seek treatment for injuries not reported in the police crash report data. However, it is important to use caution when interpreting the unlinked emergency department visit records, as there is no way of knowing if the injury warranted a police crash report (e.g. involved a motor vehicle and occurred on a public roadway). For example, we would expect that many of the bicycle crash injury-related emergency department visits would not involve a motor vehicle.

Based on the results displayed in Tables 8 and 9, unlinked pedestrian/bicycle crash records were more likely to involve minor injuries ("C's" and "O's") and deaths ("K's") than linked crash records. Unlinked crash records were also more likely to involve nighttime and early morning crashes than linked crash records. For pedestrian/bicycle crash injury-related emergency department visits, patients with linked records were more likely to have a disposition of "Admitted to Hospital/Died" and more likely to arrive by ambulance than unlinked emergency department visits.

Table 8. Comparison of Linked and Unlinked Pedestrian Crash Injury Records

Pedestrian Characteristics	Linkage Status of Crash Report Data[1]					
	Unlinked Crash Records		Linked Records		Unlinked ED Visit Records[2]	
	N	%	N	%	N	%
Sex						
Female	611	37.5%	559	42.0%	826	38.1%
Male	1,019	62.5%	773	58.0%	1,343	61.9%
Age Group						
0-9	65	3.9%	80	5.8%	97	4.5%
10-19	236	14.0%	203	14.7%	286	13.1%
20-29	347	20.6%	279	20.2%	493	22.6%
30-39	277	16.4%	186	13.4%	430	19.7%
40-49	231	13.7%	210	15.2%	324	14.9%
50-59	283	16.8%	191	13.8%	306	14.0%
60-69	168	10.0%	133	9.6%	156	7.2%
≥70	81	4.8%	101	7.3%	87	4.0%
Time of Crash/Emergency Department Visit[3]						
0:00-5:59	155	9.2%	103	7.4%	284	13.0%
6:00-11:59	370	21.9%	330	23.9%	390	17.9%
12:00-17:59	569	33.7%	507	36.7%	716	32.9%
18:00-23:59	594	35.2%	443	32.0%	789	36.2%
KABCO						
Fatal Injury (K)	157	9.7%	37	2.8%		
Serious Injury (A)	148	9.1%	141	10.6%		
Minor Injury (B)	467	28.9%	587	44.1%		
Possible Injury (C)	658	40.7%	515	38.7%		
No Injury (O)	188	11.6%	51	3.8%		
Race/Hispanic Ethnicity						
White, Non-Hispanic	791	48.9%	603	46.1%		
Black, Non-Hispanic	676	41.8%	537	41.0%		
Hispanic Ethnicity	93	5.8%	101	7.7%		
Other Race/Ethnicity	56	3.5%	68	5.2%		
Emergency Department Disposition						
Admitted to Hospital/Died			262	19.5%	226	10.7%
Discharged from ED			1,069	79.5%	1,847	87.3%
Other Disposition			14	1.0%	43	2.0%
Mode of Transport						
Ambulance			943	74.8%	839	42.8%
Walk-in/Other Mode			317	25.2%	1,122	57.2%
TOTAL	**1,688**	**100.0%**	**1,383**	**100.0%**	**2,179**	**100.0%**

Missing:
Sex: 51 linked ED visit-crash records, 58 unlinked crash records, and 10 unlinked ED visit records missing sex.
KABCO: 52 linked ED visit-crash records and 70 unlinked crash records missing KABCO.
Race/Hispanic ethnicity: 74 linked ED visit-crash records and 72 unlinked crash records missing race/Hispanic ethnicity.
Emergency department disposition: 38 linked ED visit-crash records and 63 unlinked ED visit records missing emergency department disposition.
Mode of transport: 123 linked ED visit-crash records and 218 ED visit records missing mode of transport.
[1]Missing values are excluded from column totals; due to rounding, column totals may not sum to 100%.
[2]Unlinked pedestrian/bicycle crash injury-related ED visits identified by presence of pedestrian/bicycle crash injury mechanism codes or keywords.
[3]Time of crash used for linked ED visit-crash records and unlinked crash records; time of ED visit used for unlinked ED visit records.

Table 9. Comparison of Linked and Unlinked Bicycle Crash Injury Records

Cyclist Characteristics	Linkage Status of Crash Report Data[1]					
	Unlinked Crash Records		Linked Records		Unlinked ED Records[2]	
	N	%	N	%	N	%
Sex						
Female	155	24.1%	113	23.4%	1,585	26.4%
Male	489	75.9%	369	76.6%	4,427	73.6%
Age Group						
0-9	29	4.4%	29	6.0%	1,494	24.8%
10-19	117	17.9%	86	17.7%	1,754	29.1%
20-29	174	26.6%	97	19.9%	681	11.3%
30-39	99	15.1%	58	11.9%	541	9.0%
40-49	60	9.2%	50	10.3%	535	8.9%
50-59	92	14.0%	99	20.3%	546	9.0%
60-69	64	9.8%	54	11.1%	344	5.7%
≥70	20	3.1%	14	2.9%	139	2.3%
Time of Crash/Emergency Department Visit[3]						
0:00-5:59	31	4.7%	20	4.1%	296	4.9%
6:00-11:59	131	20.0%	112	23.0%	851	14.1%
12:00-17:59	276	42.1%	210	43.1%	2,347	38.9%
18:00-23:59	217	33.1%	145	29.8%	2,540	42.1%
KABCO						
Fatal Injury (K)	30	4.7%	6	1.2%		
Serious Injury (A)	34	5.3%	29	6.0%		
Minor Injury (B)	210	33.0%	248	51.5%		
Possible Injury (C)	271	42.6%	181	37.6%		
No Injury (O)	91	14.3%	18	3.7%		
Race/Hispanic Ethnicity						
White, Non-Hispanic	364	57.4%	247	51.5%		
Black, Non-Hispanic	200	31.5%	193	40.2%		
Hispanic Ethnicity	42	6.6%	29	6.0%		
Other Race/Ethnicity	28	4.4%	11	2.3%		
Emergency Department Disposition						
Admitted to Hospital/Died			63	13.3%	364	6.1%
Discharged from ED			404	85.2%	5,534	92.8%
Other Disposition			7	1.5%	63	1.1%
Mode of Transport						
Ambulance			337	76.2%	1,086	19.5%
Walk-in/Other Mode			105	23.8%	4,473	80.5%
TOTAL	**655**	**100.0%**	**487**	**100.0%**	**6,034**	**100.0%**

Missing:
Sex: 5 linked ED visit-crash records, 11 unlinked crash records, and 22 unlinked ED visit records missing sex.
KABCO: 7 linked ED visit-crash records and 19 unlinked crash records missing KABCO.
Race/Hispanic ethnicity: 7 linked ED visit-crash records and 21 unlinked crash records missing race/Hispanic ethnicity.
Emergency department disposition: 13 linked ED visit-crash records and 73 unlinked ED visit records missing emergency department disposition.
Mode of transport: 45 linked ED visit-crash records and 475 ED visit records missing mode of transport.
[1]Missing values are excluded from column totals; due to rounding, column totals may not sum to 100%.
[2]Unlinked pedestrian/bicycle crash injury-related ED visits identified by presence of pedestrian/bicycle crash injury mechanism codes or keywords.
[3]Time of crash used for linked ED visit-crash records and unlinked crash records; time of ED visit used for unlinked ED visit records.

Recommendations for Future MVC Crash-NC DETECT Emergency Department Visit Data Linkages

Based on the results of the pilot project, we developed the following recommendations to help improve future MVC data linkage efforts for the state of North Carolina.

1. Improve the collection of information for the variable "Destination of the injured person". The NC DMV 349 form contains a free text field for "Destination of the injured person". This variable identifies the hospital at which the injured pedestrian or cyclist sought clinical care. This variable could be used 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. For example, "CFVM", "CAPE FEAR VALLEY MEDICAL", and "CAPE FEAR VALLEY FAYETTEVILLE" all refer to the same hospital. For police departments with electronic crash report forms, including a "pick list" of local hospitals could improve the data quality of this variable.
2. Include a unique personal identifier on all MVC injury data sources. Currently, there is no common unique personal identifier on police crash report and NC DETECT emergency department visit data. Therefore, we are unable to verify the accuracy of the linkage results. In addition, the inclusion of a unique patient identifier would help simplify the linkage process.
3. Incorporate City of Residence into linkage algorithm: Patient ZIP code of residence sometimes differed from the ZIP code of residence documented on the crash records. Therefore, we recommend including City of Residence (available in the police crash reports and NC DETECT emergency department visit data) as part of the linkage algorithm.
4. Add Triage Note to NC DETECT emergency department visit data set: While the Chief Complaint contains important information about the patient, it often reflects the nature of the injury ("BACK PAIN") rather than the mechanism of injury ("BICYCLE CRASH"). However, the Triage Note often contains much more detailed information about the patient, including injury mechanism. Having the ability to search the Triage Note for MVC injury keywords would improve our confidence that our matched records are "True Matches". Since Triage Note is a protected field (and not shared with researchers), we recommend working closely with personnel at NC DPH to incorporate elements of the Triage Note into our data linkage activities.

5. Use pedestrian/bicycle crash injury keywords as part of initial NC DETECT emergency department visit data pull: We assumed that all (or nearly all) emergency department visit records for pedestrian/bicycle crash injuries would have at least one diagnosis code for a traumatic injury. However, this assumption was incorrect; especially for patients with less severe injuries. Therefore, we recommend including keywords as part of our initial data request for MVC injury-related emergency department visits. The list of keywords can be expanded beyond pedestrian/bicycle crash injury keywords (e.g. "MVC", "MOTORCYCLE CRASH", "MOTOR VEHICLE CRASH").

Contact Information

If you have any questions about this report, please contact:

Anna Waller, ScD, Principle Investigator[1], awaller@med.unc.edu
Katie Harmon, Co-Principal Investigator[2], harmon@hsrc.unc.edu
Kathy Peticolas, Project Coordinator[1], kathy_peticolas@med.unc.edu

Carolina Center for Health Informatics
Department of Emergency Medicine, University of North Carolina at Chapel Hill
Floor 1, 100 Market Street
Chapel Hill, NC 27516

University of North Carolina Highway Safety Research Center
Suite 300, 730 Martin Luther King Jr., Blvd.
Chapel Hill, NC 27599