

Background

The Governor's Highway Safety Program's Strategic Planning for MVC Crash and Health Data Linkage project found that data documentation was a major limiting factor to moving forward with data linkage efforts. To address this, the following standardized documentation format was applied to the key data sources for crash and health outcomes. The standardized documentation consists of two types of documents: an overview and a data dictionary. Data elements for each document type are listed below.

Document One: Overview

This provides an overview of the data source, including background of the data source, instructions on accessing the data for research and any known quality concerns with the data.

Heading	Description
Overview	General description of data source
Data owner	Owner of the data
Data description and collection criteria	Brief description of the data and the requirements for inclusion in the dataset
Type of data: source or compiled/abstracted	Whether the data were entered by the original data entry source or were abstracted or compiled from multiple sources
Are the data available to outside parties for analytical purposes?	Indicates the availability of the data for research
Process to obtain the data for research	The process for researchers to request access to the data
Website	Web address for the data source
Contacts	Key contact(s) in accessing the data
Data flowchart	A flowchart indicating the source and flow of the data
Who enters the original data	The profession(s) of the persons who enter the source data
Injury classification method	The injury classification method found in the data (e.g. ICD-10-CM, KABCO, etc.)
Collection timeframe	Timeframe for when the data was entered after event
Years available	Description of timespan for which data is available
Data history	Changes in data that might affect research
Is a data dictionary available?	Whether the data owner provided a publically available data dictionary
Dictionary	Name of the data dictionary used to populate the template
Source documentation field map	A mapping of how the labels found in the source documentation were translated into the template
Additional fields available in source documentation	Any data elements in the source documentation which were not retained in the template
Known data quality issues	Known issues with the data (e.g., race data may be incomplete or unreliable)

Document Two: Data Dictionary

The data dictionary lists all the data elements *potentially* available for that data; their availability for each project depends on the approval of the data source. The data elements included in each data dictionary depend on what was available from the source documentation.

Data element	Description	Example value
Field	Descriptive name of the field	Non-Motorist Contributing Circumstances
Field - literal	The literal element name either in the source database or in SAS	TYP_CD 8
Table or category	The table or category of the field, depending on the source documentation	Person: Non-Motorist
Description	A description of the field	The actions of the non-motorist that may have contributed to the crash.
Format	The format of the field (Alphanumeric, datetime, numeric, etc.)	Numeric
Length	The allowable length of the field in digits	3
Required (Y/N)	Whether the field is required or not	N
Unique key (Y/N)	Whether the field value is unique and can be used as a key field	N
Source comments	Comments or background from the original source beyond the definition of the field. Source documentation which provided more extensive background for each field may have multiple data elements concatenated for this field.	Source: Refer to blocks 8 - 9 on DMV-349 Form. Rationale: Important for evaluating the effect that dangerous risky non-motorist behavior has on motor vehicle crashes.
Retired (Y/N)	Whether the field is retired from use	Y
Retired Date	The date the field was retired	January 2015
Available for research	Used only for the North Carolina Trauma Registry; whether the field can ever be returned to the researcher through the data request	Y
Option list: Code	Non-descriptive code for the option available for that field	008
Option list: Code value	Descriptive value of the option	Wrong Side Of Road
Option list: Comments	Comments from source regarding that code/value pair	Max of 2