U.S. Department of Transportation
Federal Motor Carrier Safety Administration

CCFP Crash Causal Factors Program

STUDY PLAN

CCFP HEAVY-DUTY TRUCK STUDY

2025

TABLE OF CONTENTS

BACKGROUND

Overview

What Is the CCFP Heavy-Duty Truck Study?

The Crash Causal Factors Program (CCFP) was created in 2022 to comply with the Infrastructure Investment and Jobs Act of 2021 (Section 23006) [1], which mandates a "study of commercial motor vehicle crash causation." The CCFP is a widescale, detailed, multi-phased crash research program that conducts data collection and analysis to:

- » **IDENTIFY THE FACTORS** that contribute to crashes involving commercial motor vehicles (CMVs), including large trucks and buses.
- » **INFORM COUNTERMEASURES** to prevent these crashes.
- » **ESTABLISH A FOUNDATION** for continued data collection, sharing, and analysis.

Through a series of robust studies, the CCFP pursues a nuanced understanding of crashes involving CMVs so that policymakers, law enforcement agencies, regulators, and other interested parties can implement effective crash prevention measures and improve safety on our Nation's roadways. In addition to producing datasets and analytical reports, the CCFP will establish a foundation for continued data collection and analysis that the Federal Motor Carrier Safety Administration (FMCSA) and States and other jurisdictions can routinely use to optimize their CMV safety activities.

The CCFP expands on the original Large Truck Crash Causation Study (LTCCS) that concluded in 2003. The LTCCS collected an unprecedented sample of data on almost 1,000 injury and fatal crashes involving large trucks. However, regulatory and industry changes, along with an increase in fatal large truck and bus crashes in the last few years, have created the need for a new study. The CCFP takes these changes into account to provide an updated understanding of how changes in technology, vehicle safety, driver behavior, and roadway design since 2003 have altered driver and vehicle performance. Additionally, the CCFP incorporates the lessons learned from the LTCCS and will move forward with a phased approach and studies focused on specific vehicle populations and/or crash severities.

Currently, the CCFP has identified initial three phases of research:

- » **PHASE 1: CCFP HEAVY-DUTY TRUCK STUDY**
- » **PHASE 2: CCFP MEDIUM-DUTY TRUCK STUDY**
- » **PHASE 3: CCFP BUS STUDY**

Studies in Phase 2 and 3 are planned, but their scope and execution depend on future funding.

This document details the sampling plan for the first phase study, the CCFP Heavy-Duty Truck Study. Formerly called the Large Truck Crash Causal Factors Study, the Heavy-Duty Truck Study aims to identify the contributing factors in fatal crashes involving Class 7/8 trucks.

Study Purpose

The Heavy-Duty Truck Study ("the study") seeks to collect data to better understand the factors that contribute to fatal crashes involving heavy-duty trucks in order to develop effective countermeasures. A wide range of research questions have been drafted for the study covering over 80 individual crash factors [1].

Factors include those related to the:

- » Driver
- » Carrier and work conditions
- » Vehicle design
- » Vehicle maintenance
- » Vehicle load
- » Roadway design
- » Weather
- » Road condition
- » Presence of work zones
- » Other vehicles and persons involved in the crash

This Background section reviews the considerations, lessons learned, and issues related to study's sampling. The purpose of this review is to inform the sample design for the study.

Comparison to the LTCCS Sample Design

The LTCCS collected data on crashes from 24 geographic data collection areas in 17 States, which included the four geographic regions (West, Midwest, Northeast, and South) that constitute the contiguous United States (U.S.). To be included in the sample, a crash needed to involve at least one large truck with a gross vehicle weight rating (GVWR) of more than 10,000 pounds (i.e., a medium-duty or heavy-duty truck) and result in at least one fatality or one incapacitating or non-incapacitating but evident injury. Data from a total of 963 sample crashes were collected in the LTCCS.

The sampling design of the current Heavy-Duty Truck Study has much in common with the LTCCS:

- » The design is intended to obtain a nationally representative sample of crashes involving large trucks in the U.S.

» The sample uses regional stratification to ensure geographic diversity.

» The sample collection period will be approximately two years.

This study has several key differences from the LTCCS:

» This study will only sample crashes involving Class 7 and 8 trucks; that is, trucks with a gross vehicle weight rating of more than 26,000 pounds (future phases of the CCFP will focus on other vehicle types) [2].

» The study design will use data and resources from existing crash investigation teams of States and local jurisdictions (S/Js) to do the on-site investigations, rather than having separate on-site investigation teams established by the project, like LTCCS did [3].

» This study will focus on crashes in which there was at least one fatality; a convenience sample of serious injury crashes will be incorporated only if partnering S/Js are collecting this data [4].

» The study prioritizes the inclusion of teams experienced in crash reconstruction to improve data collection and better determine specific crash causal factors [5].

» This study does not include stratification by population density of the location of the crash (i.e., central cities vs. well-populated counties vs. groups of sparsely populated counties). See Representativeness: National, Stratified by Region, page 21.

» To attain more data for analysis and national estimates, this study will sample 2,000 fatal crashes (see Sample Size: 2,000 Crashes, page 21); the LTCCS aimed for 1,000 fatal or serious injury crashes.

The rationale behind these differences is documented in page 17.

Critiques of the LTCCS Sampling

Both the U.S. Centers for Disease Control and Prevention (CDC) and the Transportation Research Board (TRB) conducted reviews of the LTCCS research method, including the sampling. These reviews have implications on the sampling design for this study.

CDC documented its review in a report to U.S. Congress's Committee on Appropriations [6]. TRB formed the Committee for Review of the Federal Motor Carrier Safety Administration's Truck Crash Causation Study to provide oversight and advice concerning the LTCCS. This included a review, authored by J. Hedlund of Highway Safety North concerning the usefulness of LTCCS data for determining the relationship between crash risk and driver, vehicle, or environmental factors [7].

CDC and TRB also identified weaknesses in the LTCCS sampling plan, specifically regarding sample size and the reliance on the induced exposure to determine crash risk. In contrast, both CDC and TRB praised the LTCCS sampling plan for collecting extensive data from a nationally representative sample relatively rapidly.

As a successor to the LTCCS, this study models its sampling plan after the LTCCS, addressing the weaknesses in the LTCCS while preserving its strengths. The CDC and TRB assessments of the LTCCS are detailed below along with their implications for this study for the following issues:

» Sample size

» Induced exposure

» Sample representativeness

Sample Size

The study aims to answer research questions covering over 80 individual crash factors [1] by collecting and analyzing a sample of 2,000 fatal crashes. This section evaluates how the increased sample size criteria of 2,000 crashes will impact statistical analyses performed to answer the study's research questions and determine causal factors of crashes involving heavy-duty trucks.

Importance of Marginals in Statistical Power

To use quantitative analyses to argue that a factor contributes to crash risk, one must, at a minimum, establish a correlation between the factor and crashes. In the simplest form, the factor is a binary variable that represents whether a condition is present or not in the situation (e.g., for the factor of intoxication, a driver may be classified as either intoxicated or not). In this case, one must be able to separate crashes into four cross-tabulated groups. Given that the study samples fatal crashes, the analysis for a factor using only the study's data requires the four groups shown in Table 1.

Table 1. Schematic of crash groups necessary to establish whether a factor correlates with crash risk

		Relevance		Total
		Relevant	Irrelevant	
Factor	Has Condition	a	b	a + b
	Lacks Condition	c	d	c + d
Total		**a + c**	**b + d**	**a + b + c + d**

Table 1 illustrates the "induced exposure" method of formulating comparison groups favored in the LTCCS, where the crashes are first divided by "relevance" [8] (see Reliance on Induced Exposure, p10 for more details on this method):

» **Groups a and c**: Crashes where the factor is "relevant" (crashes of a type where the factor reasonably has opportunity to affect crash probability, if it is present), versus

» **Groups b and d**: Crashes where the factor is "irrelevant" (these crashes are used to derive the "base rate" or exposure of the factor's condition among trucks in operation).

U.S. Department of Transportation
Federal Motor Carrier Safety Administration

Crash Causal Factors Program

For the relevant and irrelevant groups, crashes are then divided into those that do or do not have the factor's condition (Groups *a* and *b*, and Groups *c* and *d*, respectively). With data divided as such, one can compare the rate of the condition among relevant crashes, a/(a + c), with the exposure rate for trucks in operation b/(b + d) to judge if the factor correlates with crashing. It is statistically straightforward to calculate both the strength of the correlation (e.g., odds ratio) and the statistical significance of the correlation.

A similar matrix of four groups as shown in Table 1 is needed to determine if a factor relates to crash severity.

Table 1 relates to sampling because, assuming the factor contributes to (relevant) crashes, the probability of achieving statistical significance (i.e., the statistical power) tends to depend most on the smallest marginal (i.e., a+c, b+d, a+b, or c+d). If any of the four marginals are too small in raw number (e.g., 10), then the study is unlikely to significantly detect the factor's effect even if the effect is strong (e.g., odds ratio of 4).

Criticism of the LTCCS Sample Size

Although the LTCCS's sample size of 1,000 crashes appears to be a substantial number, CDC [6] and TRB [7] assert that some potrential factors in crashes (e.g., a tire blow-out or triple combination trucks) may have marginals as low as 1% of all crashes, meaning they were present in only 10 crashes, which may provide insufficient statistical power. Neither CDC nor TRB directly state it, but power is also problematic if a *relevance* marginal is too small. For example, power may be insufficient if a factor can be unambiguously classified as irrelevant (or relevant) for only 1% of crashes.

Implications for Sampling

The study plans to double the LTCCS's target sample size from 1,000 to 2,000. While this represents a meaningful increase in power, it is probably not sufficient to deal with marginals on the order of 1%. Power analyses done for the CCFP [9] determined that a sample size of 2,000 only has enough power $(1 - \beta = 0.8)$ for a 1% factor marginal if the effect of the factor is exceptionally strong—an odds ratio of at least 4.00. For weak effects, such as an odds ratio of 1.2 (as presented in a hypothetical case in the TRB report [7]), a sample size of at least 96,000 is necessary for a 1% factor marginal. Such a target sample size is unacceptable, even if there were sufficient funding to obtain it. The study is scoped to include only fatal crashes involving heavy-duty trucks (Class 7 and 8). With approximately 3,500 such crashes per year nationwide [10] (Table 8, page 45) an *exhaustive* sample of 96,000 crashes would take 27 years to obtain. By then, the results from the study would be based on out-of-date data and any countermeasures long overdue.

Mitigation Possibilities

A conventional approach to dealing with small marginals is to over-sample the less-frequent condition of the factor. If a condition of interest is naturally only present in 1% of cases, and one is limited to a sample size of 2,000, one can deliberately seek out cases with the condition to raise the proportion, ideally to 50%. In other words, one uses a quasi-experimental method, rather than a naturalistic

observational method. National rates for the factor might be accurately estimated from a portion of crash data that was not over-sampled for the condition. With 50% of the sample possessing each condition of a binary factor, a sample size of 2,000 has the power to detect effects with odds ratios of at least 1.5 [9].

However, this approach is unsuitable for the study due to several limitations:

» Firstly, the study seeks to analyze more than 80 factors. Achieving high representation of all these factors in the sample is not feasible.

» Secondly, the values of many factors in a crash is not known until the data are collected. There is simply no way to "target" crashes with a specific condition for data collection.

» Thirdly, even if the crashes could be targeted, they may simply be too rare nationwide to achieve sufficiently high representation. For example, given 3,500 fatal crashes per year and a condition present in 1% of them, exhaustive inclusion of all of them in a sample size of 2,000 would require almost 6 years to reach 10% and 29 years to reach 50% of the sample.

Given these limitations, the sampling strategy cannot practically address concerns raised by CDC [6] and TRB [7] about sample size. A quantitative study of naturally occurring fatal crashes, such as that employed by the study, is not a suitable method to analyze rarely occurring conditions among heavy-duty trucks.

Rare conditions are better studied with alternative methods, such as with simulator or test-track randomized control trial experiments, where a factor's conditions can be artificially introduced, or with the use of proxy measures or events such as close-calls, which are naturally more numerous, or with a qualitative in-depth study of the rare conditions when they are present in crashes (e.g., as seen in National Transportation Safety Board (NTSB) investigations).

The CCFP could utilize such alternative methods in follow-up research after completing the study. Data from the study, especially investigation findings, could be used to prioritize factors for further study. Perhaps the CCFP could use the same data-collection organization established for the Heavy-Duty Truck Study for such follow-up studies. For example, in the future the CCFP could work with its S/J partners from the Heavy-Duty Truck Study to collect data on proxy events or to conduct specialized qualitative studies of certain kinds of crashes.

Priorities and Rare Conditions

From one perspective, rare conditions may not be a problem for the Heavy-Duty Truck Study. The purpose of the study is to collect data to drive the development of countermeasures to reduce the rate of fatal crashes. Rare conditions tend to contribute little to the overall national fatality rate precisely because they are so rare. Simply put, the fewer trucking operations with a particular crash-inducing condition present, the smaller the impact of eliminating that factor's role in crashes. Given the proportion of operations with the condition, and the degree the condition increases crash risk, one can calculate the maximum perfect-world impact of eliminating a factor's role in crashes [9]. For example, if

1% of operations have a condition and the condition increases the chance of crashing by 20% (i.e., an odds ratio of 1.2, as in the TRB [7] example that required a sample size of at least 96,000), then perfectly eliminating the factor's role in crashes would reduce the national crash rate by 0.045% [9]. That translates to approximately one less death every seven months nationwide. While a case can be made to justify such countermeasures, it seems likely there are more serious factors FMSCA should focus on first. A sample size of 2,000 will have enough statistical power to detect significant effects for factors whose perfect elimination would reduce national crash rates by 2.5% or more [9]. That might be sufficient for identifying today's most serious factors to counteract.

The implications of marginal rates for the study are thus not necessarily on its sampling design, but on the broader issues of its purpose, scope, method, and research questions. If a condition has a very low marginal frequency, then a two-year naturalistic study of fatal crashes cannot quantitatively assess the factor's impact regardless of the sampling design. Many of the 80-plus factors listed in the draft research questions for the study [1] have known frequencies (e.g., as observed in the LTCCS). A review of these factors may check the marginals. If the marginal is very low, then such factors cannot be adequately evaluated with the study's quantitative analyses, and its assessment may be deferred for follow-up research. This can simplify the overall study design and allow resources to focus on the factors that are more likely to yield conclusive results.

Reliance on Induced Exposure

Suitable Research Questions

Regarding the induced exposure techniques, TRB notes, "Their usefulness depends on whether there is a suitable control group of crashes where the feature being examined has no effect" (page 1) [7]. For some factors, it may not be possible to cleanly separate crashes into those where a factor is and is not relevant.

TRB includes its own detailed analysis of the factors that lend themselves to induced exposure and those that do not. Many of these factors are among those in the study's draft research questions [1]. The factors that lend themselves to the induced exposure technique include:

» Driver qualification, training, licensing, and assessment

» Driver alertness, hours of service, and fatigue (assuming a reliable measure of these factors can be found—TRB is skeptical of post-crash driver self-reporting)

» Employment structure and working environment ("fairly crude," (page 23) as TRB puts it)

» Vehicle maintenance, defects, and inspections

» Vehicle design and load

While not explicitly analyzed by TRB, the following factors from the research questions also lend themselves to the induced exposure technique if one employs the same logic:

» Carrier characteristics

» Driver health and driving history

These factors lend themselves to induced exposure because crash investigations may reasonably differentiate crashes where the driver or the truck did or did not exhibit an action that led to the crash. In crashes where the driver and truck exhibited no such action, one can argue (not without opposition) that none of the above could have contributed to the crash, and thus the percent of these crashes with each condition of these factors indicates the exposure of the factor's conditions for trucking operations in general.

On the other hand, TRB concludes the following factors, which are also among those in the research questions, are not suitable for analysis with the induced exposure technique:

» Vehicle control/speed leading up to and in crash event

» Roadway design and infrastructure

» Work zone presence

While not explicitly analyzed by TRB, weather and road conditions are probably also unsuitable for analysis with the induced exposure technique. For all the above factors, it is difficult to distinguish crashes where the factor could and could not have contributed. What variable indicates the factor is irrelevant? For example, among all crashes on sharp curves or in the rain, what measure separates crashes where the curve or the rain could impact likelihood from crashes from where they cannot?

A few research questions concern relative frequencies of the types of crashes involving heavy-duty trucks and different types of other parties (light vehicles, vulnerable road users, construction workers). The study is eminently suitable for obtaining these statistics, and they may be useful for prioritizing follow-up research. However, they cannot in themselves identify causal factors in crashes, nor can they allow valid estimation of the number of crashes addressable by measures directed at heavy-duty trucks and their drivers [7]. For example, the data may show that 30% of truck-pedestrian fatalities occur in darkness. Without knowing the proportion of time that pedestrians near in-motion trucks are in darkness versus in the light one cannot conclude if darkness is a risk factor.

Implications for Research Questions

From a sampling perspective, the solution is to forego the induced exposure technique and include in the sample plan the collection of detail data on trucks undergoing routine operations. However, that is not doable given the restrictions of the study. At this time, one cannot expect S/Js to collect data on trucks that are just going down the road in order to get the exposure rates for all variables. Perhaps after the study, follow-up work with S/J partners could collect such valuable, highly detailed exposure data.

Barring a solution in sampling, the implications is that the CCFP should take a careful look at each factor in its research question and determine (and document) how the factor's role in crashes will be evaluated. If induced exposure is to be used:

» What variable can defensibly serve to identify relevant and irrelevant crashes?

» Does the variable have adequately large marginals given the limited sample size?

» What other data should be collected to statistically control for confounds in such a variable (e.g., long haul versus local operations)?

» If there is no suitable variable, can one be created and collected? For example, one case [11] used the LTCCS's "critical reason" for the crash to divide crashes into relevant and not, which was likely more valid for their purposes than using stopped versus moving trucks. Even if no variable can added to provide true exposure estimates, perhaps a variable may be added to provide enlightening comparisons of factor conditions.

» For any new variable added to the Heavy-Duty Truck Study, how should such data be collected (e.g., through investigator analysis and professional opinion, interviews with company management)?

» If no suitable variable can be included in the study's data, is there another specific source (a separate dataset) that can provide useful exposure data?

» If no exposure or comparison data can be obtained to assess a factor, can the factor's contribution to crashes be assessed by some other means?

If no means can be provided to determine a factor's contribution to the study, then such a factor is a candidate for removal from the study. This would allow resources to focus on the factors that are more likely to yield conclusive results.

Sample Representativeness

Both CDC [6] and TRB [7] acknowledge that the sampling plan for the LTCCS achieved an unbiased, nationally representative sample of crashes involving large trucks. Such representative sample is helpful for estimating the prevalence of certain types of crashes, which corresponds to the study's draft research questions (e.g., regarding vulnerable road users) [1]. The LTCCS achieved a representative sample by randomly selecting small geographic data-collection areas corresponding to the "Primary Sampling Units" of the National Highway Traffic Safety Administration's (NHTSA's) Crash Investigation Sampling System (CISS).

The Heavy-Duty Truck Study uses the same general sampling approach as used in the LTCCS, including by employing regional stratification. However, it is not practical for the study to seek representativeness through random sampling of data-collection areas. This is due to restrictions related to the combination of:

- » S/J investigative capability

- » Stratification by U.S. region

- » Meeting target sample size

The combined effect of these restrictions on the practicality of random sampling is detailed in the sections below.

S/J Investigative Capability

The study seeks to collect data on 2,000 fatal crashes involving Class 7 and 8 trucks over the span of two years, and it intends to collect these data through a partnership with S/Js. The selection of S/J partners for sample data collection is driven by the following considerations:

- » The S/J's investigative capability to capture crash data accurately and completely

- » The importance of representing all major regions of the contiguous U.S. and both rural and urban areas

- » The likelihood of meeting the target sample size of 2,000 crashes in two years

- » The representativeness of the sample

The accuracy and completeness of the crash data will depend on the investigative capability of the S/J partners that collect the data. S/J data collection, combined with interviews, can match the accuracy and completeness achieved by the LTCCS using its project-specific on-site investigators. The University of Michigan Transportation Research Institute (UMTRI's) combined State crash records with interview data from the Trucks Involved in Fatal Accidents survey (TIFA), whose data collection overlapped with the LTCCS [12]. With TIFA data coming from interviews and State data coming mostly from police crash reports and some crash reconstructions, UMTRI was able to accurately derive many of the key variable values the LTCCS determined with its own investigators.

However, UMTRI uncovered some possible issues with relying on interviews and State data. Moreso than the LTCCS, UMTRI was unable to assign known values to the variables, as shown in Table 2. For UMTRI, the more detail sought, the less certain value assignments could be. For example, UMTRI found it especially difficult to distinguish between driver recognition and driver decision as a critical reason in crashes [12].

Table 2. Percent of crashes coded as "Unknown" for UMTRI using TIFA and State data versus LTCCS

Variable	UMTRI	LTCCS
Critical Reason	12.2%	1.2%
Critical Event	9.0%	0.0%
Vehicle with Right of Way	11.8%	0.4%

Crash Type	8.2%	0.0%
Pre-event Movement	9.8%	0.0%
Avoidance Maneuver	30.2%	2.9%

Data completeness is related to variations in the quality of data provided by States. Most States proved unable to provide crash reconstructions to UMTRI, and police crash reports across States varied in level of detail. Police forces that are not primarily concerned with traffic safety (e.g., those of major metropolitan areas) appear less likely to provide complete information. Two States did not submit data to UMTRI at all [12]. Thus, while partnering with States and other jurisdictions can be a viable alternative to project-specific investigation teams, it is crucial that the CCFP seek out S/Js with the inclination and investigative capability to supply complete and accurate data.

Several factors measure a S/J's overall capability to provide complete and accurate data:

» The existence of a CMV crash reconstruction team and strength of the CMV crash reconstruction program

» The existence of an information technology (IT) infrastructure which supports electronic storage and sharing of crash reports

» The existence of crash investigation policies which support timely and accurate collection of crash information

The CCFP has conducted information collection surveys designed to assess the overall capability and willingness of S/Js, specifically States, to participate in the study. The results of this survey and the selection of States are documented in Sampling Frame, , page 31. The implication for sample representativeness are documented in Weighting (page 53) and Coverage Bias Measurement and Correction (page 55).

Stratification by U.S. Region

In selecting S/Js, the CCFP needs to ensure that the sample represents all major regions of the contiguous U.S. in order to ensure a broad, nationally representative sample of crashes and to control for variability. As with the LTCCS, this can be accomplished by stratifying the sample.

In selecting S/Js to partner in the study, a balance must be struck between achieving stratification and acceptable levels of investigative capability of the selected S/J. For example, there may be only a few States from a region with acceptable crash reconstruction capabilities. As a result, all of these States are recruited and any others in the region are systematically excluded from the study.

Meeting Target Sample Size

The target sample size is 2,000 fatal crashes, incrementally increasing the statistical power of the LTCCS. With approximately 3,500 fatal crashes per year involving Class 7/8 trucks, the study seeks to sample 29% of all fatal crashes in each year of data collection. If one excludes fatal crashes in S/Js

with weaker investigation capabilities, the proportion of available crashes to sample likely exceeds 60% (most States do not appear to be able to supply crash reconstruction data [12]). Subject this to further constraints from stratifying the sample, and it seems likely that the study will need to collect all crash data from the most capable S/Js. That is, S/J investigative capability and stratification will dictate the sampling of crashes.

Implications for Sample Representativeness

The Heavy-Duty Truck Study cannot practically achieve representativeness by using random sampling like the LTCCS used. Once restricted to crashes in high-capability S/Js stratified across regions, there will not be a reasonable number of potential "sampling units" to sample from. The study will simply take all of the most capable States per region.

This has implications for the representativeness of the sample. To some degree stratification assures a certain level of representativeness. However, there is the potential for bias in any variable that is correlated with S/J investigative capability. The management of this issue is detailed in Coverage Bias Measurement and Correction page 55.

Summary

The Heavy-Duty Truck Study is the first phase in the CCFP's series of studies to understand crashes involving CMVs. In expanding on prior work of the LTCCS, the study seeks to incorporate the LTCCS's strengths while addressing weaknesses documented in critiques of the LTCSS. Table 3 summarizes these issues regarding sampling design.

Table 3. Heavy-Duty Truck Study capabilities regarding issues in LTCCS critiques

Issue	Heavy-Duty Truck Study Capability	Restrictions on Capability
Sample Size	Sample size double that of LTCCS, incrementally increasing statistical power. Sample size sufficient for major factors in crashes.	Sample size limited by: • Regulatory and budget restrictions (page 17) • Time restrictions: two years (page 18) • Vehicles: heavy-duty trucks (page 19) • Outcomes: fatalities only (page 20)
Induced Exposure Reliance	Limited use of induced exposure compared to the LTCCS and pursuit of alternative methods.	Methods limited by: • Outcomes: fatalities only (page 20) • Data collection workforce: pre-existing S/J crash investigation agencies (page 24)
Sample Represent-ativeness	Continued use of stratification and weighting. Selection of sampling areas depends on S/J investigative capability rather than chance.	Random sampling not feasible due: • Vehicles: heavy-duty trucks (page 19) • Outcomes: fatalities only (page 20) • Data collection workforce: pre-existing S/J crash investigation agencies (page 24)

As shown in Table 3, the capacity for the Heavy-Duty Truck Study to address issues is constrained by restrictions of the study, as listed in the last column. The next section details the restrictions and their rationale. Follow-up CCFP studies for specific research questions may not be subject to such restrictions.

RESTRICTIONS

The purpose of this section is to define and explain the restrictions that impact the study's sampling and analysis efforts.

Regulatory and Budget Restrictions

Definition: In consideration of mandates, appropriated funds, and current information resources, the CCFP will seek to collect and analyze a new dataset of crashes involving CMVs.

RATIONALE

A provision under Section 23006 of the Infrastructure Investment and Jobs Act drives the design of activities within the CCFP. The provision directs a study of crashes involving CMVs to accomplish two objectives:

» To determine contributing factors to crashes that involve commercial motor vehicles, and

» To identify data requirements, data collection procedures, reports, and any other measures that can be used to improve the ability of States and the U.S. Department of Transportation (DOT) Secretary to:

   o Evaluate future crashes involving CMVs.

   o Monitor crash trends and identify causes and contributing factors.

   o Develop effective safety improvement policies and programs.

The Consolidated Appropriations Act of 2021 provides $30,000,000 for FMCSA to begin the study of causal factors for crashes involving large trucks in the U.S. [13]. This amount roughly doubles the funding for the LTCCS (about $16,000,000, according to the TRB Motor Carrier Safety Research Analysis Committee Letter Report: March 13, 2017) with the intention of doubling the number of sampled fatal crashes.

A concept analysis of alternatives compared collecting and analyzing a new dataset of crash data with (a) no study, (b) database development only, or (c) utilization of only historical data (e.g., NHTSA's Fatality Analysis Reporting System (FARS), FMCSA's Motor Carrier Management Information System (MCMIS), and State police crash reports (PCRs)). The analysis scored alternatives on data usefulness, cost, length of time, and scalability/repeatability. The analysis concluded that collecting and analyzing a new dataset was superior primarily in terms of ensuring the data meets the requirements while remaining within the appropriation [14].

Time Restrictions

Definition: Data collection will be completed in two years with analysis targeted to be completed toward the end of 2028.

RATIONALE

Over the past year, the CCFP has completed the following in preparation for data collection:

» Create the sample design and analysis plan

» Distribute, collect, and analyze a survey of potential partner S/Js for interest and investigative capability

Prior to collecting data for the study, the CCFP needs to complete the following activities:

» Prepared detailed data collection and analysis plans for individual research questions

» Establish data collection agreements with partner S/Js

» Complete data collection onboarding & training

» Prepare IT infrastructure for data reception and analysis

» Pilot test the study

These prerequisite activities are estimated to require two years starting from early 2024. The data collection period follows these prerequisite activities and is also planned to span two years. Data collection is tentatively scheduled to begin in March 2026, with a goal of obtaining 1,000 samples annually. Data analysis is estimated to commence in March 2027 and conclude in September 2028, with the final report of the study currently targeted for finalization in March 2029.

By FMCSA's estimation, two years provides sufficient time to prepare for data collection. Fatal crashes involving heavy-duty trucks (the target sample of the study) average 3,500 crashes annually in the U.S. [10]; therefore, two years is estimated to provide enough time for the study to meet its 2,000-crash sample-size goal when allowing for restrictions on sampling units and resources. A longer data collection period would allow for a larger sample size. However, FMCSA is concerned that anything longer will not produce timely results that can be implemented in a meaningful way [14].

Sampling Data Restrictions

Vehicles: Heavy-Duty Trucks

Definition: The sample includes only crashes involving Class 7 and 8 trucks. These are trucks with a gross vehicle weight rating of more than 26,000 pounds, which are classified as "heavy-duty" trucks by the Federal Highway Administration [15].

RATIONALE

The CCFP seeks to identify causes of crashes involving CMVs of all sizes, but this study serves as an initial study sampling only crashes involving heavy-duty trucks. Follow-up studies will collect data on crashes involving smaller trucks and buses. The framework for this study is designed to be scalable and repeatable, allowing its application to be leveraged in future CCFP phases [2].

Over 70% of CMV crashes involve heavy-duty trucks [10]. By prioritizing Class 7 and 8 trucks for the sample, the study focuses on the group in which crashes are most frequent and severe, allowing a more targeted implementation of countermeasures to advance the goal of reducing fatal crashes.

Selecting heavy-duty trucks as the subject of the CCFP's initial data collection and analysis will also help identify crash factors that may be unique to trucks in this classification. The recent rise in crashes involving medium-duty trucks, distinct from trends in crashes involving heavy-duty trucks [16] implies these crashes may have different causes. NHTSA is completing a study specifically of medium-duty trucks, so government resources are more efficiently used at this time if there is no overlap [4]. By deferring the CCFP's study of medium-duty trucks to later years, the CCFP can build on NHTSA's research results, apply lessons learned from both NHTSA and the Heavy-Duty Truck Study, and design a new study that addresses any limitations of that research.

The likelihood of identifying causal factors unique to the group in a sample consisting solely of crashes involving heavy-duty trucks is higher than in an equally sized sample which may include crashes involving trucks belonging to other classifications. Including both heavy- and medium-duty trucks in the same sample often requires truck class to become a covariate in the statistical analyses, shrinking the marginals (or decreasing the degrees of freedom) and thus decreasing statistical power [17]. A sample of the same size comprising only crashes involving heavy-duty trucks skirts the need for the covariate and maximizes power. Effects unique to medium-duty trucks must be left to a separate study.

The LTCCS lost resources by investigating out-of-scope crashes involving light-duty trucks due to police officers on the scene confusing light-duty with medium-duty trucks and then calling the go-teams [4]. In contrast, heavy-duty trucks have proven to be easier to distinguish from smaller trucks [18]. Not many medium-duty trucks are subject to U.S. DOT and State regulations as these vehicles are not standard; thus, studies of these trucks require the personnel collecting the data to have more extensive training. Focusing on heavy-duty trucks allows the CCFP to build and test the organizational and technical infrastructure before applying it to classes that present more challenges to study. By the time the CCFP turns its attention to other CMVs, uncertainties will be resolved regarding data collection

content and quality, deployment of data collection resources, and the coordination of State vs. Federal enforcement.

Outcomes: Fatalities Only

Definition: The study will only sample crashes resulting in at least one fatality. The study may also collect a convenience sample of crashes with at least one incapacitating or non-incapacitating but evident injury.

RATIONALE

The study aims to help reduce fatal crashes involving heavy-duty trucks. Statistics generated with NHTSA's Fatality and Injury Reporting System Tool[1] reveal that the number of heavy-duty trucks involved in fatal crashes increased by 19% from 2020 to 2022. The study aims to yield information that will help FMCSA and the truck safety community identify activities and other measures likely to lead to significant reductions in the frequency and severity of these crashes. Fatal crashes are substantially less common than crashes with non-fatal injuries, but they represent the most severe crashes and thus warrant the most quickly implemented countermeasures.

The LTCCS sample failed to yield enough statistical power to analyze the subset of fatal crashes. For a given representative sample size of all crashes with injuries and fatalities, fatalities will constitute a small fraction due to fatal crashes being far less frequent than non-fatal injury crashes. The result is a small marginal for fatal crashes and low statistical power regarding fatal crashes, even though they are precisely the crashes with the greatest need for countermeasures. By collecting data on only fatal crashes, the marginal is maximized and thus power is maximized [17].

The inclusion of non-injury "towaway" crashes can be problematic. The LTCCS encountered difficulty identifying towaway crashes in a sufficiently timely fashion to get an investigator on scene. If a truck investigator is not on scene before the vehicle is towed, vital information about the crash can be lost (e.g. if the brakes were not "locked") [4]. Furthermore, towaway crashes are cleared from the highway more quickly than injury crashes, and law enforcement, in practice, is less likely to call in the go-team for towaway crashes because it does not want to wait on scene [4].

Fatal crashes, in contrast, typically receive a more thorough investigation, including full crash reconstruction. With the study relying on pre-existing crash investigation agencies to collect data, focusing on fatal crashes aligns the requirements of the study with the policies of the investigation agency, ensuring consistent collection of data.

---

[1] https://cdan.dot.gov/query

For entry into FARS, a crash must involve a motor vehicle traveling on a trafficway customarily open to the public and must result in the death of a vehicle occupant or a nonoccupant within 30 days of the crash. To remain consistent, the Heavy-Duty Truck Study will use the same definition of a fatal crash.

In addition to the sample of 2,000 fatal crashes, the study will accept data from full investigations of non-fatalities if available without imposing an excess burden or cost. Future studies may focus on different crash severities [5].

Sample Size: 2,000 Crashes

Definition: As a target, the study will collect full data on 2,000 events. These events will be described as "crashes" not as "accidents," consistent with the American National Standard Institute [19].

RATIONALE

A target of 2,000 crashes meaningfully improves on the sample size of the original LTCCS. For the potential crash factors that the Heavy-Duty Truck Study is expected to examine, a sample size of 2,000 allows researchers to detect factor conditions associated with as little as a 2.5% excess in crashes involving heavy-duty trucks [9].

Budget restrictions limit the sample size that may be collected and processed. With a $30 million budget, a rough order of magnitude analysis allows for approximately $17.2 million for actual data collection over two years ($8.6 million per year). Other annual costs total $6.8 million over two years, and one-time expenses (e.g., preparation and analyses) account for the remaining $6.0 million [20]. Based on original estimates using NHTSA's CISS infrastructure, $6.8 million is the annual cost for collecting data on 1,000 events per year for a total of 2,000 [21]. It is assumed that the data collection annual cost using existing investigation agencies is comparable. A meaningful increment in statistical power can be achieved by doubling the sample size to 4,000 [9]. However, that is estimated to cost $57.0 million [21].

The period of data collection also constrains the sample size [22]. There are approximately 3,500 fatal crashes involving heavy-duty trucks per year in the U.S. [10]. With a data collection period of two years, an exhaustive collection of all crashes would yield approximately 7,000 crashes. Given restrictions on the number of participating investigative agencies and the study's requirements for investigation team capabilities and dispositions, data format and variables, and record inclusion, the study most likely can collect data on only a fraction of the 7,000 crashes theoretically available in the period of data collection.

Representativeness: National, Stratified by Region

Definition: The study intends to ensure that its data is representative and documented in a way conducive to calculating national population crash rates while ensuring that all regions of the contiguous U.S. are represented [23]. However, it will not be possible to collect a sufficiently large sample for data to be aggregated by individual county or census tract, nor is it feasible to stratify crash locations by the level of population density.

RATIONALE

While the CCFP intends to obtain a nationally representative sample of crashes involving heavy-duty trucks, this study, like the LTCCS before it, seeks to ensure crashes are sampled from different regions of the U.S. For example, this study may follow the LTCCS's method of sampling an equal number of sampling units from northern, eastern, southern, and western U.S. As in the LTCCS, weights may be applied to crashes in each stratum to arrive at the national frequencies of crashes. Such stratification was used in the LTCCS, and the reviews of the LTCCS method commended it on its representativeness [6] [7].

However, within each stratum, there may be a limited number of suitable sampling units available, given requirements for investigation team capabilities, inclination to participate, data format and variables, and record inclusion. To obtain equal sampling units in all strata, it may be necessary to compromise on certain requirements, which may impact data analysis and interpretation.

The LTCCS stratified its sample by population density as well as region, selecting equal numbers of sampling units comprising central cities, well-populated counties, and sparsely populated counties. Such stratification is not necessary to achieve a nationally representative sample. Indeed, it can complicate the calculation of national rates [24]. However, such stratification would facilitate statistical comparisons across population levels and add to the study's face validity.

Whatever its merits, it is not feasible to stratify by population density for this study due to the following restrictions (detailed in the referenced sections):

» Data Collection Workforce: Pre-Existing, In-Place Crash Investigation Agencies (page 24)

» Jurisdiction Level of Investigation Agencies: Generally State Level (page 25)

» Sampling Units: Equivalent to Entire Crash Investigation Agency (page 28)

With sampling units composed of predominantly entire States, the CCFP cannot select sampling units that are strictly urban, suburban, or rural, since each State generally comprises a mix of population densities.

Relating Factors to Risk: Crash Analysis for Contributing Factors

Definition: The primary means to relate factors to crash risk will be through analysis of crashes to determine the key factors' contributions to the crash followed by statistical analysis of other factors on the contributing factors.

RATIONALE

Most of the study's research questions concern the impact of factors on crash risk. Typically, impact is assessed by comparing crash frequencies for a factor's conditions with exposure rates of those conditions. For example, the proportion of crashes involving novice CDL drivers could be compared to the proportion of driving done by novice CDL drivers. With data limited to that collected by multiple

crash investigation agencies across the country (see Jurisdiction Level of Investigation Agencies: Generally State Level , page 25), there is no feasible means to collect exposure data. For instance, the CCFP does not have the leverage to induce agencies to conduct roadside polls of CMVs passing crash locations to determine exposure of a factor. Such teams may not have the expertise, authority, or resources for such polling anyway.

Useful exposure information potentially could be sourced from other datasets for some factors. However, unbiased exposure data for many factors may not exist. For example, road-side inspection data is not truly random —not all trucks have an equal chance of inspection [25]. For other factors, exposure information may be available by employing an "induced exposure" method to the crash data by dividing crashes into those that are relevant to the factor hypothesized to impact crash risk, and those that are not. However, given that variables available for defining relevance are largely limited to standards for crash investigation, using the induced exposure technique may not be workable for many factors [7].

For some factors and research questions, it may be appropriate to compare the rates of factor conditions in fatal crashes to their rates in injurious crashes, using the "convenience" sample of injurious crashes involving heavy-duty trucks (see Outcomes: Fatalities Only, p20). In other words, one can assess the correlation of the factor with crash severity, rather than fatal crash probability. However, the data quality, confounds, and sampling biases of the convenience sample are uncontrolled and unknown, so such analyses may prove to be uninterpretable. Furthermore, while some research questions concern the impact of a factor on crash severity, most concern the impact on fatal crash probability, and an analysis for severity is not necessarily a valid substitute; a factor can increase crash probability but not impact the relative distribution of injurious and fatal crashes.

It appears inevitable that exposure data will not be available for most research questions. For these questions, the study will, where it is feasible, avoid the need for exposure data by employing validated crash analysis procedure to determine contributing factors. Leveraging the in-depth investigations of the crashes, this study can identify contributing factors using analysis similar to NTSB's determination of probable cause in its investigations and the LTCCS's determination of critical reason [27]. By determining if a factor helped *cause* a crash, rather than if it was merely *present* in a crash, these crash analyses do not require exposure data. Developing and validating objective rules for attributing a factor's contribution to a crash is labor intensive and cannot be feasibly completed for the dozens of factors of interest. However, if this is done for a relatively small number of key factors, the study can divide crashes into those with and without a particular contributing cause to examine what other variables correlate with the factor contributing to crashes. This can help determine the role of "upstream" regulatory, organizational, and cultural factors in eliciting a contributing factor.

Both acquiring exposure data and employing crash analysis may impact the variables that need to be collected on each crash.

Data Collection Workforce: Pre-Existing, In-Place Crash Investigation Agencies

Definition: In contrast to the original plan to use the CISS infrastructure of Primary Sampling Units (PSUs) to collect data on crashes (as was done for the LTCCS), the study will collect crash data through pre-existing investigation and crash reconstruction agencies, such as those of States and local jurisdictions.

RATIONALE

FMCSA conducted an alternative analysis comparing collecting data through CISS infrastructure versus through State and local agencies. The alternative analysis concluded that FMCSA should proceed with using State and local agencies because it has fewer risks and challenges; is more cost efficient; will strengthen partnerships with S/Js, fostering future research collaborations; and will likely produce better quality data [3]. In contrast, using the CISS infrastructure:

» Would likely require exceeding the $30 million budget.

» Uses PSUs that, being geared more to passenger car crashes, appear to have limited resources to support post-crash inspections on trucks.

» Has issues with consistent notification of go-teams, resulting in failure to collect data on some crashes, which the study can ill afford given the limited number of crashes involving heavy-duty trucks in CISS PSUs [28].

By partnering with State and local agencies that are already investigating the same crashes as a matter of practice, the CCFP avoids having to contract investigators specifically for the study, reducing costs to FMCSA and duplicate efforts.

Partnering with State and local agencies also reduces daytime bias in a cost-effective way. For the LTCCS, NHTSA crash investigators were available 24 hours per day. However, FMCSA truck inspectors were only on call during daytime. This appears to have resulted in the LTCCS oversampling daytime crashes. The Heavy-Duty Truck Study could resolve the oversampling issue by paying FMCSA truck inspectors to be available 24 hours per day. However, that increases the data collection costs or limits the number of inspectors on hand at any given time in a sampling unit [29], which leads to less data collected. By partnering with State and local agencies, the CCFP can predominantly select those that have inspectors available 24 hours per day. CCFP selection can also favor jurisdictions that legally hold crashed vehicles until they are inspected. These measures reduce the potential for bias towards daytime crashes, ensure more complete data collection, and avoid paying FMCSA inspectors overtime.

While discussing the possible use of CISS infrastructure with FMCSA, NHTSA expressed concern over competition for investigators between CISS and CCFP, with NHTSA declaring, "It (is) going to be a balance…. We will not disrupt CISS operations [which usually concern passenger cars] at the sites to

do the [CCFP's data collection]" [22]. By partnering with State and local agencies instead of using CISS infrastructure, the investigators may be better focused on crashes that are more likely to be in scope for the study.

Investigating crashes involving heavy-duty trucks requires special expertise and facilities. For example, engineers with knowledge of heavy-duty trucks may be needed for speed reconstructions or similar items, especially with trailer roll-overs. NHTSA developed the CISS with an emphasis on crashes involving passenger cars, and this is reflected in the expertise available, the form of the investigations, and the data collected. With CISS, truck experts may not be consistently available in some sampling units participating in the crash investigations. Foregoing the CISS infrastructure, selecting agencies with expertise in crashes involving heavy-duty trucks, and training investigators can address this issue.

Current budget estimates conservatively assume that the data collection costs when using State and local agencies is the same for the current sampling units as when using CISS PSUs (see Sample Size: 2,000 Crashes, page 21). However, more crash data most likely can be collected for the same cost by working with State and local agencies [30].

By leveraging data the State and local agencies are already collecting, FMCSA reduces duplicative data and its associated costs and burden to individuals [5].

Jurisdiction Level of Investigation Agencies: Generally State Level

Definition: The investigation teams conducting data collection will be almost completely associated with U.S. State jurisdictions, rather than local (e.g., county or municipal) jurisdictions.

RATIONALE

Pre-existing investigation agencies include those associated with State governments and those associated with local jurisdictions, such as country or municipal authorities. Crashes investigated by local jurisdictions are generally not investigated by their respective State. Approximately 30% of qualifying crashes nationwide are investigated by local jurisdictions [31].

Consideration was given to partnering with investigative agencies at both the State and local level, and even stratifying the sample by jurisdiction level, to enhance the representativeness of the sample. Local jurisdictions generally represent metropolitan areas and thus have proportionally more urban populations than States on average. Thus, stratifying on jurisdiction type could approximate stratification by population density.

However, further study reveals this to be ineffective. Few local jurisdictions expressed willingness to participate in the study. Thus, the study is very unlikely to recruit a substantial number of capable local jurisdictions. A local jurisdiction, by definition, covers relatively small areas, and thus each would investigate only a few crashes during the data collection period. Thus, partnering with local jurisdictions, even if they were willing, raises issues with efficient use of study resources and the capabilities of the investigating agency (see Crashes in Each S/J per Year: At Least Six, page 27). Where opportunity allows, the sample includes crashes investigated by local jurisdictions in addition to State jurisdictions. However, the reality is that only a few local jurisdictions can be expected to participate in the study.

Systematic recruitment of S/Js focuses on selecting States for participation, and the resulting sample will be dominated by crashes investigated by State agencies.

The dominance of State jurisdictions opens the possibility of bias entering the sample if such local jurisdictions tend to have different distributions of crash factors than States. State jurisdictions include some metropolitan areas, so bias associated with population density can be assessed and possibly corrected through comparisons within data collected by State jurisdictions. Other biases from excluding these local jurisdictions are controlled through comparing the sample data to nationally exhaustive fatal crash data (see page 55).

Investigation Team Capabilities: Best Data Collection Capability

Definition: The investigation teams conducting data collection within each S/J agency must have proven expertise and experience in heavy-duty truck crash reconstruction, as indicated by their background, products, and historical frequency of crash investigation.

RATIONALE

Crash reconstruction and other data collection regarding truck crashes requires special expertise. Expertise in crashes involving passenger cars is not sufficient because crashes involving trucks have physical, organizational, and regulatory characteristics absent from cars (e.g., much greater height and mass relative to barriers; greater potential for rollover, jackknifing, and other crash dynamics; different braking and engine technologies; fleet operations; roadside State inspections; specialized driver training; Federal hours-of-service regulations).

Given the reliance on S/J crash investigation agencies for data collection, FMCSA cannot simply select a random sample of States to include in the sample design. This would result in recruitment of S/Js that do not have the capability to supply the high-detail data necessary for the study. To use program resources efficiently, FMSCA will limit sample participation to S/Js that have the best capability to collect complete, high-quality data on qualifying crashes and to troubleshoot study-related issues as they arise [5] as determined by information collection surveys of State agencies (see Quantifying S/Js' Investigative Capability, p32). Therefore, agencies contributing to the sample must be those with trained and experienced full-time teams that routinely conduct detailed investigations (e.g., dedicated CMV investigation teams that respond to a minimum of six crashes involving heavy-duty trucks per year [32]) and have other indications of expertise [33].

Sampling the most capable S/Js raises the prospect of sampling bias due to the possibility of S/Js with lesser investigative capabilities tending to have different distributions of crash characteristics. However, such a potential bias occurs even if such S/Js were included (e.g., through random sampling). Low capability S/Js by definition cannot supply the data the study needs. Therefore, any systematic differences in the crashes of such S/Js will be unavoidably under-sampled anyway. This is an inevitable consequence of relying on pre-existing, in-place investigative agencies (see page 24). For efforts to detect and correct for bias, see Coverage Bias Measurement and Correction, page 55.

Crashes in Each S/J per Year: At Least Six Qualifying Crashes

Definition: The study recruits S/Js that historically have at least six qualifying crashes per year.

RATIONALE

The study has limited resources that impact the quantity and quality of data obtained, and thus must use its resources efficiently to get the most high-quality data as feasible. There are overhead costs associated with partnering with S/Js related to setting up agreements, accessing systems and data, and training S/J personnel. These overhead costs cannot be justified for S/Js that are likely to contribute relatively little data; that is, S/Js that have relatively few qualifying crashes per year. Furthermore, an S/J with relatively few crashes per year is less likely to acquire the experience and processes necessary to conduct high-quality investigations. This implies the study should exclude S/Js with historically few qualifying crashes. Based on the experience with the LTCCS, the Heavy-Duty Truck Study's threshold for inclusion was set to six qualifying crashes per year.

This exclusion opens the possibility of bias entering the sample if very small States tend to have different distributions of crash factors than other States. However, precisely because they are predicted to contribute only a small number of crashes to the sample, any potential bias is also very small. For example, examining 2020-2021 data of NHTSA's FARS, qualifying crashes from States with less than six per year sum to 0.46% of all qualifying crashes. Like all potential sampling biases, any such biases from excluding these States are controlled through the weighting of crash cases (see page 53) and comparing the sample data to nationally exhaustive fatal crash data (see page 55).

For a discussion of a related potential for bias and its control, see Jurisdiction Level of Investigation Agencies: Generally State Level, page 25.

Crash Investigation Agency Disposition: Full Participation

Definition: The sample S/Js must be disposed to cooperate fully with all the requirements of the study, including those for crash selection/sampling, data collection and sharing, data format and storage, and training requirements.

RATIONALE

Given the reliance on S/J crash investigation agencies for data collection, sample agencies must be disposed to cooperate fully with all the requirements of the study. They must be willing and legally allowed to participate, complying with sampling, data collection, and turnover requirements. FMSCA will need to recruit an appropriate mix of S/Js that are so disposed [5]. Negotiations with agencies may necessitate certain compromises that add additional restrictions on the sampling design. Regardless of agreements with the S/Js, events beyond the control of the CCFP or an S/J may result in an S/J not fully complying with the requirements for participation. For example, S/Js may be required to investigate to a certain level all qualifying crashes under their purview. However, a historically unusual rash of qualifying crashes during the data collection period may exceed the S/J's resources.

Sampling Units: Equivalent to Entire Crash Investigation Agency

Definition: Sampling units must correspond to entire S/J crash investigation agencies collecting the data.

RATIONALE

Data collection of crashes across the nation requires the definition of "sampling units," each being a unit of trained experts, facilities, and other resources that are prepared to collect data on crashes on short notice. Teams in sampling units need to collect data from the crash site before it is cleared, so each sampling unit is responsible for crash data collection within a limited geographic area that teams can drive across in a reasonable time.

Early plans for the study called for using CISS PSUs as sampling units for collecting data. Each PSU included a team responsible for collecting data for CISS. PSUs are carefully distributed around the contiguous U.S. to provide a nationally representative sample of vehicular crashes.

However, while the PSUs were successfully used as sampling units in the LTCCS, they present issues when collecting data for the Heavy-Duty Truck Study. With the scope limited to fatal crashes involving heavy-duty trucks (see Vehicles: Heavy-Duty Trucks, page 19, and Outcomes: Fatalities Only, page 20), historic data indicated that the PSUs cannot produce enough in-scope data within the two-year restriction for data collection (see Time Restrictions, page 18) [34]. Using the PSUs over two years, the LTCCS collected data on only 307 crashes that meet the qualifications for inclusion in the Heavy-Duty Truck Study, in contrast to the target sample size of 2,000 crashes (see Sample Size: 2,000 Crashes, page 21). The risk of getting less than 2,000 cases in two years remains regardless of the specific PSUs selected [22]. Furthermore, using the PSUs may not represent national frequencies for fatal crashes involving heavy-duty trucks [3]. The PSUs were designed primarily for passenger vehicle crash data collection [34]. For this and other reasons, the study will not use the CISS PSUs as sampling units, nor use the CISS infrastructure in general (see Data Collection Workforce: Pre-Existing, In-Place Crash Investigation Agencies, page 24). Instead, the study will use S/J crash investigation agencies.

Given the restriction on using the most capable S/J agencies (page 14), the study's sampling units must correspond to entire agencies. If the CCFP were to instead sample only some of the crashes within a highly capable S/J, then the study must sample additional crashes from a less capable S/J to obtain the same sample size. Given the national frequency of fatal crashes involving heavy-duty trucks, collecting data on only a subset of crashes from the most capable S/J may impede achieving a sample size of 2,000 crashes from S/Js with acceptable investigative capabilities.

While the selection process will seek to select representative S/Js as sampling units, the sampling units are thus not necessarily uniform in crash frequency, distribution of vehicle types or lines of business, environmental characteristics, and other factors.

Data Format and Variables: Standardized

Definition: Variables and formats in the dataset will comply with industry and/or national standards applicable for truck-involved crashes. Agencies that do not sufficiently follow the standards may be excluded from the sample.

RATIONALE

With data coming from multiple agencies across the country, the data format and variables must be consistent across agencies to allow the datasets to be merged. That is, all datasets must use practically the same variables with the same definitions and comprise the same data elements. For example, quantitative variables should all have the same units and normalization. This is partially accomplished through training S/Js to use consistent formats and variables. Also, in some cases, the CCFP can algorithmically translate idiosyncratic variables from an S/J's data into formats consistent with others (e.g., convert units, aggregate separate categories into a single standard category).

The CCFP has limited leverage to dictate the data format and variables of S/J investigation teams. The CCFP has the best chance of gaining data consistency by encouraging S/Js to use formats and variables defined in industry and national standards for crash data, such as the NHTSA Model Minimum Uniform Crash Criteria (MMUCC) or the standards of the Commercial Vehicle Safety Alliance (CVSA). Many S/Js follow these standards already, so the CCFP's adoption of such standards will result in the most S/Js having consistent variables and formats. Furthermore, for S/Js currently with idiosyncratic variables or formats, the CCFP is more likely convince S/Js to switch to a national standard than to formats and variables unique to the study.

Nonetheless, it is anticipated that some S/Js will not collect crash data in a compatible form. Agencies that do not sufficiently follow the standards may be excluded from the sample because the study is constrained by the quality of the data that the participating agencies supply. Even among selected S/Js that by policy follow standards and training, there may be some inconsistencies across agencies that may impose additional restrictions on data analysis and interpretation.

In addition to standards limiting data collection to certain variables, some variables may not be available because of privacy agreements. Collaboration with motor carriers may be necessary to ensure that the transfer of proprietary data to crash investigators does not infringe on existing data privacy agreements [35]. This may involve transferring obfuscated or aggregated data, which could impact the precision of the analyses.

Despite restrictions related to standards and proprietary information, the study needs to include additional novel variables to provide required data beyond what is already available in national databases such as FARS and MCMIS [14]. For example, additional variables may represent automated driving systems that operate within a limited operational design domain, such as a variable indicating whether the vehicle system was operating the vehicle at the time of the crash. Some research questions require that carriers supply data from onboard monitoring systems, telematics, or video recordings [1] [36]. Obtaining these data in a consistent format (or at all) is an ongoing challenge.

The sampling plan for the Heavy-Duty Truck Study is subject to the following restrictions:

» Vehicles: Heavy-Duty Trucks

» Outcomes: Fatalities Only

» Sample Size: 2,000 Crashes

» Representativeness: National, Stratified by Region

» Relating Factors to Risk: Crash Analysis for Contributing Factors

» Data Collection Workforce: Pre-Existing, In-Place Crash Investigation Agencies

» Jurisdiction Level of Investigation Agencies: Generally State Level

» Investigation Team Capabilities: Best Data Collection Capability

» Crashes in Each S/J per Year: At Least Six

» Crash Investigation Agency Disposition: Full Participation

» Sampling Units: Equivalent to Entire Crash Investigation Agency

» Data Format and Variables: Standardized

The next section details the study's sampling frame, defining the data to be sampled and how it is selected for inclusion. The sampling frame is constructed to operate within the restrictions above.

SAMPLING FRAME

This section defines the sampling frame for the Heavy-Duty Truck Study. That is, it defines the data to be sampled and how the data are selected for inclusion in the study. In general, the study sampling frame is intended to define a sample that:

» Provides high investigative quality.

» Is representative of fatal crashes involving Class 7/8 trucks in the U.S.

» Is practical to achieve.

Regarding the last objective, the sampling frame is subject to the constraints documented in page 17. These restrictions include the collection of data through a partnership between FMCSA and the crash investigative teams of S/Js, generally States. The crash investigative teams collect on-scene crash information, among other data.

Unit of Analysis

A crash is the unit of analysis. Additional data with different units may be fused with the dataset of crashes (e.g., to provide normalizing or exposure information), depending on the research question to address.

Scope of Data

Data collection includes two kinds of crashes: qualifying crashes and non-qualifying crashes. A qualifying crash is one that:

» Occurs in a State of the United States or in the District of Columbia.

» Involves at least one heavy-duty truck as ascertained by the on-scene law enforcement, where a heavy-duty truck is Class 7 or 8 truck. This is a truck with a gross vehicle weight rating (GVWR) over 26,000 lbs [15].

» Precipitates injuries in at least one individual that prove to be fatal within 30 days.

It is assumed that data collection includes an investigation and full reporting on all crashes with a *potential* for fatality within 30 days. Investigated crashes without a fatality within 30 days of the crash are treated as non-qualifying crashes.

Qualifying crashes are the primary data source for the Heavy-Duty Truck Study and will be used to answer the main research questions. The sampling frame is tailored to achieve the study objectives to the greatest extent possible with the available data from qualifying crashes.

The CCFP will include data, when available, on non-qualifying crashes investigated by S/J teams. As permitted by resources and agreements with the S/Js, these data are collected to form a "convenience sample" to answer certain research questions such as those concerning crash severity, rather than crash occurrence.

Quantifying S/Js' Investigative Capability

S/Js are recruited to participate based on their demonstrated capacity to conduct complete high-quality investigations of qualifying crashes without substantial sampling bias. The capability for a State to conduct investigations was determined through the responses to a series of four Information Collection (IC) surveys of State records managers, law enforcement command staff, coordinators of the Motor Carrier Safety Assistance Program (MCSAP), and other contacts knowledgeable of a State's reporting, investigations, inspections, and reconstructions of road vehicle crashes [5].

Scoring Criteria

To quantify investigative capability, responses to selected items from the IC surveys were assigned numerical values, and these values were summed to create an Investigation Capability Score (ICS) for each the State. The ICSs were used to guide State selection for Heavy-Duty Truck Study participation as specified in State Recruitment, page 40. The survey items and criteria for assigning numeric values were set by FMCSA subject matter experts to represent characteristics of a capable investigative team.

The following are survey items and the criteria for assigning numeric values to their responses:

1. **ABILITY TO PARTICIPATE**

   Score as **1** if the following item has the given response:

IC	Item	Response
2	1.1.1 Would there be any challenges in participating?	No challenges anticipated

2. **POLICY OR POTENTIAL TO NOTIFY INSPECTORS OF ALL OR NEARLY ALL QUALIFYING CRASHES**

   Score as **1** if the following item has the given response:

IC	Item	Response
3	1.1.2 Does your agency *currently* notify **post-crash inspectors** of all (or nearly all) fatal crashes involving a Class 7/8 truck?	Yes

   Score as **½** if all of the following items have the given response:

IC	Item	Response
3	1.1.2 Does your agency *currently* notify **post-crash inspectors** of all (or nearly all) fatal crashes involving a Class 7/8 truck?	NOT Yes
3	1.1.3 If your agency were to participate in the Heavy-Duty Truck Study, could your agency change or enhance its crash notification processes to notify **post-crash inspectors** of *all (or nearly all)* fatal crashes involving a Class 7/8 truck?	Yes

3. POLICY OR POTENTIAL TO NOTIFY RECONSTRUCTIONISTS OF NEARLY ALL QUALIFYING CRASHES

Score as **1** if the following item has the given response:

IC	Item	Response
3	1.1.8 Does your agency *currently* notify **post-crash investigators or reconstructionists** of *all (or nearly all)* fatal crashes involving a Class 7/8 truck?	Yes

Score as **½** if **all** of the following items have the given response:

IC	Item	Response
3	1.1.8 Does your agency *currently* notify **post-crash investigators or reconstructionists** of *all (or nearly all)* fatal crashes involving a Class 7/8 truck?	NOT Yes
3	1.1.9 If your agency were to participate in the Heavy-Duty Truck Study, could your agency change or enhance its crash notification processes to notify post-crash investigators or reconstructionists of *all (or nearly all)* fatal crashes involving a Class 7/8 truck?	Yes

4. POLICY OR POTENTIAL TO COLLECT ADDITIONAL POST-CRASH INSPECTION DATA

Score as **1** if the following item has the given response:

IC	Item	Response
3	1.3.2 Does your State collect additional **post-crash inspection data** for fatal crashes involving at least one Class 7/8 truck—beyond what is sent to the Motor Carrier Management Information System (MCMIS)?	Yes

Score as ½ if **all** of the following items have the given response:

IC	Item	Response
3	1.3.2 Does your State collect additional **post-crash inspection data** for fatal crashes involving at least one Class 7/8 truck—beyond what is sent to the Motor Carrier Management Information System (MCMIS)?	NOT Yes
3	1.3.3 If your State participates in the Heavy-Duty Truck Study, would your State be willing to collect additional **post-crash inspection data**—beyond what you already collect—for fatal crashes involving at least one Class 7/8 truck?	Yes

5. ABILITY TO ALTER METHOD/TRAINING

Score as ½ **for each** of the following that items have the given response:

IC	Item	Response
3	1.4.4 If your agency participates in the Heavy-Duty Truck Study, FMCSA will need to provide **study-specific post-crash inspection training**, to ensure post-crash inspectors 1) understand study data collection processes and 2) collect the required study data. FMCSA will most likely use a "train the trainer" approach for this training. Could your agency incorporate this **study-specific post-crash inspection training** into your existing training regimen?	Yes
3	1.4.9 If your agency participates in the Heavy-Duty Truck Study, FMCSA will need to provide **study-specific post-crash investigation training**, to ensure post-crash inspectors 1) understand study data collection processes and 2) collect the required study data. FMCSA will most likely use a "train the trainer" approach for this training. Could your agency incorporate this **study-specific post-crash investigation/reconstruction training** into your existing training regimen?	Yes
3	1.4.14 If your agency participates in the Heavy-Duty Truck Study, FMCSA will need to provide **study-specific crash reconstruction training**, to ensure post-crash inspectors 1) understand study data collection processes and 2) collect the required study data. FMCSA will most likely use a "train the trainer" approach for this training. Could your agency incorporate this **study-specific crash reconstruction training** into your existing training regimen?	Yes

6. DISPATCH OF POST-CRASH INSPECTORS OF CLASS 7/8 TRUCKS INVOLVED IN QUALIFYING CRASHES

Score as listed for each of the following responses to:

IC-4 1.1.11 To what percentage of **fatal CMV crashes** are **post-crash inspectors** dispatched?

Response	Score
75% or More	1
50% to 74%	½
25% to 49%	¼

7. **DISPATCH OF INVESTIGATORS TO QUALIFYING CRASHES WITH ON-SCENE INFORMATION COLLECTION.**

Score as listed for each of the following responses to:

IC-4 1.1.19 To what percentage of **fatal CMV crashes** are **post-crash investigators** dispatched?

Response	Score
75% or More	1
50% to 74%	½
25% to 49%	¼

8. **DISPATCH OF RECONSTRUCTIONISTS TO FATAL CRASHES**

Score as listed for each of the following responses to:

IC-4 1.1.27 To what percentage of **fatal CMV crashes** are **crash reconstructionists** dispatched?

Response	Score
75% or More	1
50% to 74%	½

Response	
25% to 49%	¼

9. EXPERIENCED CMV CRASH RECONSTRUCTION TEAM

Score as listed for each of the following responses to:

IC-4 1.1.35 How many years has your **crash reconstruction program** been conducting crash reconstructions for **qualifying crashes**?

Response	Score
10 or More Years	1
5 to 9 Years	½
1 to 4 Years	¼

10. COMPLETION OF INVESTIGATIONS OF QUALIFYING CRASHES

Score as listed for each of the following responses to:

IC-4 1.1.21 What percent of **qualifying crashes** in your agency are subject to a **post-crash investigation**?

Response	Score
75% or More	1
50% to 74%	½
25% to 49%	¼

11. STORES CRASH AND INVESTIGATION DATA IN ELECTRONIC FORMAT

Score as **0.2 for each** of the following that items have the given response:

IC	Item	Response
3	1.2.3 How does your State store **police crash report (PCR) data**?	Local server (electronic storage)
3	1.2.5 How does your State store **CMV supplemental data**?	Local server (electronic storage)
3	1.2.6 Does your State have any of the following mechanisms for sharing **crash data** with FMCSA?	Web service OR Data export
3	1.2.9 How does your agency store **post-crash investigation data**?	Electronic State crash data repository OR Local server (electronic storage)
3	1.2.13 Does your agency have any of the following mechanisms for sharing **crash reconstruction data** with FMCSA?	Web Service OR Data Export

With these criteria, the minimal possible ICS is 0 and the maximal is 11.5.

Survey Scoring Results

Descriptives statistics for the ICS are in Table 4.

Table 4. Statistics for ICS across all States

Mean	6.928
Standard deviation	2.838
Minimum	0.700
1st Quartile	4.900
Median	7.600
3rd Quartile	9.400

Maximum	10.600
Skewness	-0.664
Kurtosis	-0.655

States' ICSs vary over almost the entire range of possible values. The distribution of ICS is generally bell-shaped while somewhat skewed toward higher scores. The average intercorrelation among the criteria scores was fairly low (0.283), with values ranging from –0.104 to 0.744, consistent with ICS being a composite score of investigative capability on several dimensions. These statistics suggest that ICS will be effective in distinguishing S/Js on investigative capability.

ICS is not significantly correlated with the States' fatality crash rate of CMVs per vehicle-mile traveled (r = –0.228, Fisher z = –1.57, p = 0.116), indicating that recruiting States by ICS does not necessarily bias the sample towards S/Js with better or worse crash records. A plan for thorough assessment of the sample for bias is detailed in Coverage Bias Measurement and Correction, page 55.

Regional Stratification

To assist in obtaining a broad, nationally representative sample of crashes and to control for variability among regions of the country, the sample is stratified by geographic region. All States are grouped into one of the regions of the U.S. as defined in Table 5 and shown in Figure 1. Data from NHTSA's FARS for 2020 and 2021 indicates that these regions historically contain approximately the same number of either medium- or heavy-duty trucks involved in fatal crashes and are thus suitable for both the Heavy-Duty Truck Study and a proposed follow-on study for crashes involving medium-duty trucks (see Table 6).

To stratify the sample by region, States are recruited for Heavy-Duty Truck Study participation such that the sample of qualifying crashes from each region is predicted to be approximately equal. The predicted number of qualifying crashes during the data collection period is indicated by the FARS 2020-2021 data, so stratification is performed by recruiting States such that the total number of 2020-2021 qualifying crashes from each region is about the same. Recent data suggest the predicted number of crashes will be within a few percent of historic averages [10].

As seen in Table 5, the number of States per region varies considerably, with 18 and 19 States in the North and West regions, respectively, and only 6 or 7 in the South and Southeast regions, respectively. This is because the frequency of qualifying crashes across States varies considerably, with fewer crashes per State on average in the North and West than the South and Southeast. This implies that the number of States recruited per region will vary. For example, the West region may have more States recruited than the South region because proportionally more of the West States have a relatively low frequency of qualifying crashes, so more States must be recruited to have an approximately equal predicted number of qualifying crashes for the regions.

Table 5. States within each proposed region of stratification

North	South	Southeast	West
Connecticut	Alabama	Florida	Alaska
Delaware	Arkansas	Georgia	Arizona
District of Columbia	Louisiana	Kentucky	California
Illinois	Mississippi	North Carolina	Colorado
Indiana	Oklahoma	South Carolina	Hawaii
Maine	Texas	Tennessee	Idaho
Maryland		Virginia	Iowa
Massachusetts		West Virginia	Kansas
Michigan			Minnesota
Missouri			Montana
New Hampshire			Nebraska
New Jersey			Nevada
New York			New Mexico
Ohio			North Dakota
Pennsylvania			Oregon
Rhode Island			South Dakota
Vermont			Utah
Wisconsin			Washington
			Wyoming

Figure 1. Regions for stratification

Table 6. Average number of trucks involved in fatal crashes per region from NHTSA FARS 2020–2021[2]

Region	Number of States	2020-2021 Average Number of Trucks in Fatal Crashes Annually	
		Heavy-Duty Trucks	Medium-Duty Trucks
North	18	1,010.0	368.0
South	6	971.0	350.5
Southeast	8	1,002.0	401.0
West	19	1,019.0	409.0

State Recruitment

Initial Elimination

State recruitment is the process of selecting States for participation in the Heavy-Duty Truck Study. Two of the study's restrictions (see Summary, page 15) apply quantifiable criteria for eliminating States from consideration for recruitment:

» Investigation Team Capabilities: Best Data Collection Capability (page 26)

» Crashes in Each S/J per Year: At Least Six (page 27)

ICS indicates a State's investigative capabilities. The number of qualifying crashes in FARS 2020-2021 data indicates a State's historic number of crashes per year. These values are shown in Table 7 for the contiguous U.S. States and the District of Columbia. The table is sorted by region and then ICS. To

---

[2] Query submitted via Fatality and Injury Reporting System Tool (FIRST) (https://cdan.dot.gov/query, Vehicles tab) where Years = 2020–2021, States as rows and GVWR as columns. Strictly speaking, these are data for the number of *trucks* involved in fatal crashes. Since a crash may involve more than one truck of a class, the number of trucks involved in fatal crashes exceeds the number of fatal crashes involving trucks, as recorded in [10]. The statistics comprise all fatal crashes occurring in the States of each region; the form and thoroughness of investigation of any of these crashes is unknown. Table 5 also does not show the division of crashes between those investigated by a given State and those investigated by any local jurisdiction within the State.

apply the restriction regarding investigative capability, the study eliminates States in the bottom quartile of ICS scores (less than 4.90). This excludes the following States in Table 7:

» Wisconsin

» Illinois

» Michigan

» Arkansas

» Louisiana

» Tennessee

» Montana

» Colorado

» North Dakota

» Oregon

» Wyoming

» California

To apply the restriction regarding crashes per year, the study eliminates States with fewer than 12 qualifying crashes in 2020-2021 (six per year on average). This excludes the following additional States shown in Table 8, all from the North region:

» Vermont

» New Hampshire

» District of Columbia

» Rhode Island

Table 7. State ICS and frequency of qualifying crashes

Region	State	ICS	Qualifying Crashes	Eliminate
North	Vermont	10.40	7	X
North	Massachusetts	10.10	42	
North	New Hampshire	9.90	11	X
North	New York	9.65	182	
North	Missouri	9.60	181	
North	New Jersey	9.60	110	
North	Maine	8.90	26	
North	District of Columbia	8.60	4	X
North	Connecticut	8.20	36	
North	Maryland	7.70	75	
North	Pennsylvania	7.65	248	
North	Indiana	7.15	276	
North	Rhode Island	6.90	4	X
North	Ohio	6.70	263	
North	Delaware	6.10	16	

Region	State	ICS	Qualifying Crashes	Eliminate
North	Wisconsin	4.65	130	**X**
North	Illinois	3.60	276	**X**
North	Michigan	2.15	133	**X**
South	Mississippi	10.40	160	
South	Texas	10.40	1080	
South	Oklahoma	7.60	163	
South	Alabama	6.60	213	
South	Arkansas	1.35	157	**X**
South	Louisiana	0.70	169	**X**
Southeast	South Carolina	9.40	198	
Southeast	Kentucky	9.10	188	
Southeast	Virginia	8.60	181	
Southeast	North Carolina	7.60	238	
Southeast	Florida	7.45	506	
Southeast	Georgia	6.10	365	
Southeast	West Virginia	4.90	51	
Southeast	Tennessee	2.65	277	**X**

Region	State	ICS	Qualifying Crashes	Eliminate
West	Washington	10.60	111	
West	South Dakota	10.40	33	
West	Iowa	9.60	125	
West	Kansas	9.40	131	
West	Minnesota	8.90	113	
West	Utah	8.60	67	
West	Nevada	8.10	72	
West	Nebraska	7.40	85	
West	Idaho	6.85	70	
West	Arizona	6.40	170	
West	New Mexico	5.15	126	
West	Montana	4.70	39	X
West	Colorado	3.40	125	X
West	North Dakota	3.15	20	X
West	Oregon	2.80	101	X
West	Wyoming	2.40	43	X
West	California	1.20	591	X

Achieving a Sample of 2,000 Crashes

The goal of recruiting States is to produce a regionally stratified sample of States that will thoroughly investigate at least 2,000 qualifying crashes over two years (1,000 crashes per year). Not all qualifying crashes that occur *within* a State are necessarily investigated *by* the State. Some are investigated by local jurisdictions (see Data Collection Workforce: Pre-Existing, In-Place Crash Investigation Agencies), and these are expected to provide relatively little data to the sample (see Jurisdiction Level of Investigation Agencies: Generally State Level). To conservatively determine if the sample is predicted to yield at least 2,000 qualifying crashes, one considers only the historic number of *State-investigated* qualifying crashes per State.

Furthermore, experience from the LTCCS and responses from the IC surveys indicate that not all qualifying crashes in a selected area get sampled or thoroughly investigated. There is also the risk (methodologically speaking) that the national frequency of crashes may decrease by the start of the data collection period. To assure full data collection of a sample size of at least 2,000 qualifying crashes, enough States should be recruited such that the historic (2020-2021) number of State-investigated qualifying crashes is substantially over 2,000. Experience from LTCCS and responses to the IC surveys, plus a "safety margin," suggest a factor of 1.667 should be applied, and thus enough States should be recruited so that the number of State-investigated qualifying crashes from 2020-2021 is about 3,333. Stratifying by the four regions, the recruited States from each region should have investigated approximately 833 qualifying crashes collectively in 2020-2021.

Top-Ranked States by ICS

In recruiting States from each region, priority goes to States capable of higher-quality investigations above and beyond having an ICS over the first quartile (see Quantifying S/Js' Investigative Capability, page 32). To give priority to high-ICS States while simultaneously stratifying by region, the States are ranked by ICS within each region, as shown in Table 8. The higher-scoring States are then recruited from each region until the cumulative number of 2020-2021 crashes investigated by the States in each region is about 833, as indicated by the checkmarks in Table 8.

Table 8. States to target for recruitment based on ICS rank, stratified by region

Region	State	ICS	State-Investigated Qualifying Crashes	Cumulative Crashes	Target for Recruitment
North	Massachusetts	10.10	28	28	✓
North	New York	9.65	123	154	✓
North	Missouri	9.60	148	302	✓

Region	State	ICS	State-Investigated Qualifying Crashes	Cumulative Crashes	Target for Recruitment
North	New Jersey	9.60	41	343	✓
North	Maine	8.90	13	356	✓
North	Connecticut	8.20	32	388	✓
North	Maryland	7.70	70	458	✓
North	Pennsylvania	7.65	176 *	634	✓
North	Indiana	7.15	128	762	✓
North	Ohio	6.70	162	924	
North	Delaware	6.10	11 *	935	
South	Mississippi	10.40	90	90	✓
South	Texas	10.40	659	749	✓
South	Oklahoma	7.60	151	900	
South	Alabama	6.60	156	1056	
Southeast	South Carolina	9.40	176	176	✓
Southeast	Kentucky	9.10	88	264	✓
Southeast	Virginia	8.60	145	409	✓
Southeast	North Carolina	7.60	200	609	✓

U.S. Department of Transportation
Federal Motor Carrier Safety Administration

CCFP Crash Causal Factors Program

Region	State	ICS	State-Investigated Qualifying Crashes	Cumulative Crashes	Target for Recruitment
Southeast	Florida	7.45	399	1008	✓
Southeast	Georgia	6.10	259 *	1267	
Southeast	West Virginia	4.90	36 *	1303	
West	Washington	10.60	79	79	✓
West	South Dakota	10.40	31	110	✓
West	Iowa	9.60	94	204	✓
West	Kansas	9.40	124	328	✓
West	Minnesota	8.90	88	416	✓
West	Utah	8.60	54	470	✓
West	Nevada	8.10	55	525	✓
West	Nebraska	7.40	32	557	✓
West	Idaho	6.85	57	614	✓
West	Arizona	6.40	133	747	✓
West	New Mexico	5.15	67	814	✓

* Estimates based on national average proportion of State-investigated crashes.

Selective Reduced Sampling within Some States

The above recruitment procedure yields an approximately equal number of crashes per region to achieve regional stratification. However, the number of States is relatively small, and recruitment of a

State adds a wide possible range of predicted crashes to the region. To better balance the predicted number of crashes sampled from each region and to achieve a broader sample within each region, the sampling for States with very large numbers of crashes is reduced by randomly sampling a subset of investigative districts within the State.

Texas and Florida are associated with the largest number of State-investigated qualifying crashes. Included in its entirety, Texas dominates the data from the South region, leaving "room" for recruitment of only one other State, as shown in Table 8. To avoid over-representing Texas's crashes, and any potential associated biases, the study will sample investigative districts within Texas such that Texas's contribution to the sample is proportional to its contribution to all qualifying crashes in the nation. Using 2020-2021 data, about 14% of the nation's qualifying crashes occurred in Texas (1080/7998). Fourteen percent of 3333, the target number of historic crashes for the study, is 467. If Texas's selected districts are associated with 467 crashes, then that leaves a target of about 366 crashes (833 - 467) associated among the remaining recruited States of the South region. In addition to Mississippi (top ICS scorer of the South), this allows for the recruitment of Oklahoma and Alabama.

This brings the South region's total State-investigated qualifying crashes to 864, acceptably close to the target of 833.

Including all of Florida's crashes in the sample gives the Southeast region 1,008 historic crashes, well over the target of 833. This results in the Southeast region being disproportionate to other regions. Using 2020-2021 data, about 6.3% of the nation's qualifying crashes occurred in Florida (506/7998). Six-point-three percent of 3,333 is 211. If Florida's selected districts are associated with 211 crashes, then the Southeast region is associated with 820 crashes, acceptably close to the target of 833.

Qualitative Assessment and Exclusion

Details from the IC surveys and discussions with State officials may uncover attributes that make a State unsuitable for inclusion in the sample. Such is the case with Pennsylvania. State Police leadership has declined to allow interviews to be conducted while their investigation is under way. Without this essential data delivered in a timely manner, Pennsylvania must be excluded from the sampling.

Ohio takes the place of Pennsylvania. Of the two North region States not already targeted for recruitment (see Table 8), Ohio has the higher ICS (6.70). Historically, Ohio has a comparable number of State-investigated qualifying crashes (162) to Pennsylvania (176 estimated). Thus, replacing Pennsylvania with Ohio keeps the predicted State-investigated qualifying crashes from the North region close to 833.

Final List of States

Taking into account the selective sampling with Texas and Florida, Table 9 has the final list of States targeted for recruitment to achieve a regionally stratified sample that fits the restrictions of the study (see page 17).

Table 9. Recruitment targets

Region	State	ICS	State-Investigated Qualifying Crashes	Cumulative Crashes
North	Massachusetts	10.10	28	28
North	New York	9.65	123	154
North	Missouri	9.60	148	302
North	New Jersey	9.60	41	343
North	Maine	8.90	13	356
North	Connecticut	8.20	32	388
North	Maryland	7.70	70	458
North	Indiana	7.15	128	586
North	Ohio	6.70	162	748
South	Mississippi	10.40	90	90
South	Texas	10.40	659 **	749
South	Oklahoma	7.60	151	900
South	Alabama	6.60	156	1056
Southeast	South Carolina	9.40	176	176
Southeast	Kentucky	9.10	88	264
Southeast	Virginia	8.60	145	409

Region	State	ICS	State-Investigated Qualifying Crashes		Cumulative Crashes
Southeast	North Carolina	7.60	200		609
Southeast	Florida	7.45	399	**	1008
West	Washington	10.60	79		79
West	South Dakota	10.40	31		110
West	Iowa	9.60	94		204
West	Kansas	9.40	124		328
West	Minnesota	8.90	88		416
West	Utah	8.60	54		470
West	Nevada	8.10	55		525
West	Nebraska	7.40	32		557
West	Idaho	6.85	57		614
West	Arizona	6.40	133		747
West	New Mexico	5.15	67		814

*A sample of select investigative districts.

As seen in Table 9, in seeking a target of 3,333 crashes stratified by region, the study seeks partnership with 29 States. This represents 88% (29/33) of the of States that passed initial elimination. That is, the study seeks to target nearly all State-investigated qualifying crashes for States with a minimally acceptable investigative capability. The final list comprises 79% (3260/4108) of qualifying crashes from all States that passed initial elimination. The study seeks a nearly exhaustive sample to the extent possible within its restrictions.

Despite efforts to balance the sampling across regions, it is very likely the proportions in the sample associated with each region will deviate from the proportions of crashes in each region overall. Biases associated with these disproportions are controlled by weighting crash cases (see page 53) and comparing the sample data to nationally exhaustive fatal crash data (see Coverage Bias Measurement and Correction, page 55).

Alternates

If a State declines or withdraws from participation prior to the start of State training, the State with the highest ICS will be recruited from remaining unrecruited States of the same region (see Table 7). Table 10 lists the first alternate States for each region. Montana is the first alternative for the West, even though its ICS is marginally below the one quartile threshold; it is still the best alternate. For the South, the next-highest ranked State is Arkansas with an ICS of 1.35 (4th percentile). Rather than recruiting such a State with little chance of providing useful data, more districts from Texas will be sampled as an alternate.

Table 10. First alternates for recruitment

Region	State	ICS	State-Investigated Qualifying Crashes
North	Delaware	6.10	11 (est.)
South	More from Texas	10.40	(variable)
Southeast	Georgia	6.10	263
West	Montana	4.70	28

Any potential sampling biases from the recruitment or withdrawal of States are controlled by weighting crash cases (for details, see page 53) and by comparing the sample data to exhaustive crash data (see Coverage Bias Measurement and Correction, page 55).

* Wisconsin was added as a participating state as of April 2025.

** As of July 2025, New Mexico notified that they will not be participating in the CCFP Heavy-Duty Truck Study and will be replaced by the state of Montana.

Local Jurisdiction Recruitment

The recruitment of local jurisdictions is subject to the same conceptual criteria and restrictions as State jurisdictions. The result is that few local jurisdictions can be included primarily due to the following restrictions in Table 11, mostly stemming from the relatively small size of the jurisdictions (see also the restriction Jurisdiction Level of Investigation Agencies: Generally State Level, page 25).

Table 11. Study restrictions and local jurisdictions

Restriction	Page	Tendency among Local Jurisdictions
Investigation Team Capabilities: Best Data Collection Capability	26	Relatively low investigative capability
Crashes in Each S/J per Year: At Least Six	27	Few crashes per year
Crash Investigation Agency Disposition: Full Participation	27	Low willingness to participate.

With so few candidates, selection of local jurisdictions was performed without systematic rating, ranking, and balancing. Instead, if a local jurisdiction was found that passes these restrictions, it was recruited. Investigative capability was assessed qualitatively through discussions with the local jurisdiction officials, rather than through scoring responses to the IC surveys.

Potential partner local jurisdictions include some from Arizona, Indiana, Missouri, New York, and Texas. If one or more local jurisdictions from Texas elect to participate, then the sampling from the State jurisdiction of Texas may be reduced to maintain balance among the regions and reduce the work demands on the State.

Subsampling within S/Js

With the exception of Texas and Florida (see Selective Reduced Sampling within Some States, page 47), each S/J is expected to investigate and submit data on all qualifying crashes that occur under their purview over the data collection period. Among the ICS criteria is the historically demonstrated capability of a State to investigate nearly all qualifying crashes under its purview. However, there is always the possibility, either due to an area-wide precipitating event (e.g., due to weather or jurisdiction budget cuts) or by coincidence, that too many qualifying crashes occur for the S/J team to fully investigate all.

If it becomes necessary for the S/J to choose a crash for full investigation, then, for the purposes of research, the choice should ideally be random. However, there are practical limits to this approach. For example, a S/J may be bound by policy to investigate the more severe crashes when a choice is forced. Timing may also dictate the choice since the occurrence of a crash cannot be predicted. If the only available investigation team is already on scene for one crash, they cannot be fully available for a subsequent crash elsewhere. To avoid gaps in the variables, selecting one crash to fully investigate is preferable to partially investigating more than one crash.

These subsampling realities may introduce coverage bias in the sample. Such biases are controlled by weighting crash cases (for details, see page 53) and by comparing the sample data to exhaustive crash data (see Coverage Bias Measurement and Correction, page 55).

Weighting

For the purposes of generalizing the sample to the Nation, the study is subject to possible bias due to disproportionate sampling. Disproportionate sampling can occur when the proportion of crashes in a sample from a given place is not equal to the proportion of all actual crashes from that place. Bias in rates from disproportionate sampling can be corrected by weighting the crashes in the sample using the known proportions of actual crashes. The study has two potential sources of disproportionate sampling. The first is from stratification, related to the S/J recruitment process. The second is due to subsampling, related to the number of qualifying crashes exceeding an S/J's capability to investigate.

Disproportionate Stratified Sampling

Table 12 illustrates a hypothetical case of disproportionate stratified sampling and its correction, using only two of the four strata, or regions, (South and Southeast) for simplification. In this example, suppose the States in Table 9 ultimately participate in the study as planned.

Table 12. Hypothetical example of bias from disproportionate sampling and its correction

	South Stratum		Southeast Stratum		Overall	
Actual Crashes	State	Qualifying Crashes	State	Qualifying Crashes		
	MS	160	SC	200		
	TX	1080	KY	190		
	OK	160	VA	180		
	AL	210	NC	240		
	AR	160	FL	510		
	LA	170	GA	370		
			WV	50		
			TN	280		
Total		1940		2020	3960	
Proportion of Crashes		49%		51%		
Rate of Condition (an unknown)		60%		40%		
Frequency of Condition (an unknown)		1164		808	1972	49.80%
Sampled Crashes	Recruited State	Investigated Qualifying Crashes	Recruited State	Investigated Qualifying Crashes		
	MS	90	SC	90		
	TX	320	KY	60		
	OK	110	VA	70		
	AL	100	NC	100		
			FL	110		
Total		620		430	1050	
Proportion of Crashes		59%		41%		
Frequency of Condition		372		172	544	51.81%

		3.129	4.698		
Correction		3.129	4.698		
Corrected Frequency		1164	808	1972	49.80%

biased

At the end of the period of data collection, the actual number of qualifying crashes within both sampled and non-sampled States is determined through the classification of exhaustive crash data, such as from MCMIS. That is, MCMIS is filtered for all qualifying crashes for all States occurring within the data collection years, and the frequencies of such crashes are determined for each State. This constitutes the actual crashes. In this hypothetical example, there were 1,940 crashes for all States in the South region and 2,020 crashes for all States in the Southeast region, for a total of 3,960 crashes, with about equal proportions in each region. In the sample, which included only crashes investigated by recruited State agencies, the South region happens to have 90 + 320 + 110 + 100 = 620 crashes, while the Southeast region happens to have 90 + 60 + 70 + 100 + 110 = 430 crashes. The total crashes in the sample for these two regions is 620 + 430 = 1050 crashes, with a disproportionate 620/1050 = 59% of the crashes coming from the South region.

Suppose there is a variable, factor, or characteristic X of crashes that is part of the study's data collection but absent from the exhaustive crash data. The rate of X for all States in each region is unknown since not all States were recruited. However, suppose it occurs at a rate of 60% in the South region and 40% in the Southeast region. That would mean there were 60% × 1,940 = 1,164 crashes with X in the South region and 40% × 220 = 808 crashes in the Southeast region for a total of 1,972 crashes, or 1,972/3,960 = 49.80% of crashes overall have X. However, due to the disproportionate sampling of the South region in this hypothetical case, the sample indicates that X occurs in 544/1,050 = 51.81% of crashes overall, a bias of about 2 percentage points.

To correct the bias, one weights (multiplies) the crashes by the ratio of the actual number of qualifying crashes in the region over the sample number of crashes. For the South region, the correcting weight is 1,940/620 = 3.129, while for the Southeast region, the correcting weight is 2,020/430 = 4.698. Thus, the corrected frequency of X is 3.129 × 372 = 1,164 crashes for the South and 4.698 × 172 = 808 for the Southeast. As can be seen in Table 12, these are the correct frequencies for all crashes in the regions and yields the overall correct 49.80% rate for X.

To summarize, bias from disproportionate stratified sampling is corrected by multiplying sample crashes by the following stratification weight:

$$w_j = \frac{\acute{N}_j}{N_j}$$

Where:

> w_j = stratification weight for stratum (region) j
> $\acute{N}_j$ = number of all qualifying crashes in the stratum (region), based on exhaustive data
> N_j = number of crashes in the stratum (region) sample

Correcting for disproportionate stratified sampling with this stratification weight assumes that the sample is representative of the whole region. For example, if 40% of qualifying crashes in the region

have *X*, then 40% of the crashes sampled from the region have *X*. Systematic bias in the sampling from a region (e.g., if recruiting States by ICS somehow changes the rate of *X* from 40% to something else) is handled through Coverage Bias Measurement and Correction (page 55).

Disproportionate Subsampling

Subsampling can yield bias in much the same way as stratification sampling. With subsampling, if one S/J does not investigate all the qualified crashes under its purview, then it is relatively underrepresented. To correct for this underrepresentation and remove its bias, one weights the sample crashes in the S/J by the ratio of the actual number of qualifying crashes in the S/J over the number of investigated crashes. That is, crashes in each S/J are multiplied by the following:

$$w_{ij} = \frac{N_{ij}}{n_{ij}}$$

 Where:

w_{ij} = subsampling weight for S/J *i* of stratum (region) *j*
N_{ij} = number of all qualifying crashes in S/J *i* of stratum (region) *j*, based on exhaustive data or S/J records
n_{ij} = number of investigated crashes of S/J *i* of stratum (region) *j*

Like correcting for disproportionate stratified sampling, the correction for disproportionate subsampling assumes that the investigated crashes are representative of all qualifying crashes in the entire S/J. Systematic bias in subsampling (e.g., if a S/J prioritizes investigating more severe crashes) is handled through Coverage Bias Measurement and Correction (page 55).

Total Weight

For the purposes of generalizing the sample to the Nation, the corrections for disproportionate stratified sampling and disproportionate subsampling are combined by multiplying the stratification weight and subsampling weight together. That is, sample crashes from S/J *i* in stratum (region) *j* are weighted as follows:

$$W_{ij} = w_j\, w_{ij} = \frac{N_{ij}}{n_{ij}}\frac{\acute{N}_j}{N_j}$$

Coverage Bias Measurement and Correction

Coverage bias is a form of systematic bias that occurs in the sample when some cases in the population do not have an opportunity to be sampled. Because States are recruited based on ICS (rather than random sampling), the study may have coverage bias. Any crash variable that correlates with ICS will be biased. Similarly, any subsampling may also induce coverage bias on any variable that correlates with the subsampling criteria (e.g., crash severity).

To assess for bias in the sample of crashes, the sample data are compared with nationally exhaustive crash data in MCMIS. All variables in MCMIS are also in the HDTS sample data, although the sample data includes many more variables than MCMIS. MCMIS is filtered for qualifying crashes for all States (whether sampled or not) and the period of data collection. The study sample data are weighted as described on page 53. The mean and standard deviation for every numeric variable in MCMIS are compared to those for the same variable from the sample. The frequency distribution of every categorical variable in MCMIS is compared to that for the same variable from the sample. Declared discoveries of bias from significant differences between MCMIS and the sample are constrained to a false discovery rate (FDR) of 0.05% by evaluating the statistical test results of the comparisons between MCMIS and the sample with a procedure such as Benjamini-Yekutieli that does not assume independence of the statistical tests.

Any discoveries of bias are used to inform interpretations of the analysis results. For some analyses, especially those involving one or more variables shared between MCMIS and the study, the sample data or analysis results may be mathematically corrected to remove bias. For example, if a factor is discovered to be biased in the sample, then the factor's correct rate from another source may be used to estimate the true extent of hazard that the factor presents on the roadways.

Limitations and Issues

The study plans to collect data on 1,000 qualifying crashes per year for two years. Given there are only about 4,000 qualifying crashes per year nationwide, the study must sample a substantial portion of all qualifying crashes. When one excludes S/Js that lack investigative agencies or are otherwise unwilling, unable, or unsuitable for recruitment, the pool of crashes to sample becomes even smaller.

It is essential that the study collect high-quality investigative data. To enhance our understanding of the factors that contribute to crashes involving CMVs, variables beyond those currently collected must be collected, and the most useful variables come from high-quality investigations.

Given these goals, the study must prioritize recruiting S/Js with the best investigative capabilities, and, as much as possible, obtain the data from *all* qualifying crashes under the S/Js' purviews for the period of data collection. As such, random sampling of crashes or S/Js is not a viable option.

The lack of random sampling increases the risk of bias in the variables. The presence of bias can be assessed and corrected, at least for the univariate distributions of variables that are in both the study and exhaustive crash data, such as MCMIS. In cases where bias cannot be assessed, the results, specifically their magnitude, need to be reported as approximations.

Concern about bias should not be exaggerated. The vast majority of the study's research questions do not involve establishing national rates at which factors contribute to crashes. Rather, the research questions are concerned with establishing a *correlation* between a factor and crashes. Most research questions are of the form, "How does *X* affect crash risk?" Generally, the presence of a correlation is assumed to be less sensitive to bias than overall univariate rates. For example, if wet roads are

correlated with crash risk in a sample from some States, then one may expect wet roads to correlate with crash risk in all States, even if some States have wet roads more often than others.

Summary

The sampling frame for the Heavy-Duty Truck Study is intended to define a sample of crashes in the U.S. that:

» Provides high investigative quality.

» Is representative of fatal crashes involving Class 7/8 trucks in the U.S.

» Is practical to achieve given the restrictions of the study.

The crash data comes from investigations by priorly existing crash investigative agencies of S/Js, generally States. The steps for sampling the data are as follows:

1. Assign each State to a region.

2. Score the States on their investigating capabilities, assigning each an ICS.

3. Eliminate States with ICS in the bottom quartile or with historically fewer than six qualifying crashes per year.

4. Rank the remaining States on their investigating capabilities within each region using ICS.

5. Recruit the highest-ranked States from each region until the cumulative number of historic crashes investigated by the State agencies in each region is close to a targeted 833 crashes, which is estimated to be sufficient to achieve the target of 2,000 total crashes in the sample.

6. Selectively sample agencies and districts from Texas and Florida to avoid their over-representation.

7. Exclude any State the proves unsuitable or unwilling to be sampled in subsequent discussions and negotiations, and replace it with next-highest ranked State in the region.

8. Recruit local jurisdictions qualitatively assessed to be suitable for the sample using the same conceptual criteria as used for States.

Any bias due to disproportionate sampling will be corrected by weighting data. Other potential biases in the sampling are quantitatively assessed, and, where possible, corrected by comparing the distributions of all variables that are present in both the sample and MCMIS.

ANALYSIS PLAN

Introduction

The common aim of crash research is to understand how a set of factors influences the risk of crashes. This is accomplished by collecting and analyzing crash data in search of evidence that indicates the significance of each factor's impact towards crashes. The validity of such evidence rests on two factors: data quality and the method(s) chosen to interpret the data.

The objective of the data collection phase is to optimize sample data quality within the research's cost restrictions. There are multiple aspects to data quality, chief among which is accuracy, or the degree to which the information captured reflects the true attributes of a crash. Data volume and completeness are also key considerations, which concern achieving sufficient sample size and capturing information on each crash variable of interest from each sample. Finally, data collection aims also to maximize *representativeness*, or the degree to which the sample collected is representative of the population the research is interested in. The process to interpret this data falls within the analysis phase of the research.

Conclusions regarding each analyzed factor's influence on risk may be derived through the application of qualitative or quantitative inference techniques. The analysis aims to obtain evidence of the impact of each factor of interest towards crash risk in the population the sample is representing.

The Heavy-Duty Truck Study seeks to identify factors that increase crash risk among heavy-duty trucks in the U.S. To do so, it plans to collect a nationally representative sample of at least 2,000 fatal crashes involving U.S. heavy-duty trucks. The effort is directed at vastly expanding the factors of analysis to include variables not captured in FMCSA's Motor Carrier Management Information System (MCMIS) repository, which holds limited attributes on roughly three million crashes involving heavy-duty trucks in the U.S.

This section serves as a guide in analyzing future and existing data available to FMCSA in the Heavy-Duty Truck Study. It aims to support the validity of the study by providing:

1.) A catalog of accepted qualitative and quantitative frameworks employed in crash research today, and

2.) An assessment of the suitability of the study's data against these methods.

Analysis Types

Risk levels of an adverse event such as a motor vehicle crash are typically measured by the probability of its occurrence and the severity of its damages. A factor which increases the severity of crashes may not necessarily increase their likelihood, and vice versa. For example, lack of seat belt use may

increase severity of damages sustained by passengers in a crash, although it is not generally inferred to increase the crash's likelihood.

Crash analysis literature today employs both qualitative and quantitative approaches in assessing how various road factors contribute to the probability or severity of crashes. In **qualitative research**, the basis for conclusion is derived primarily from investigation of unstructured, descriptive data, such as crash narratives and reports. Here, the researcher(s) rely on domain knowledge to interpret sample data, discover themes from crashes, and conclude how various factors influenced the probability or severity of crashes in the population which the sample represents.

In **quantitative analysis**, the researcher seeks statistical evidence to assess how factors impact the probability or severity of crashes. The researcher transforms raw information into quantifiable data, often expressed as a table, where crash factors act as independent variables and a measure of crash frequency or severity serves as the dependent variable. This tabular, numeric data is then used to construct crash models that best *fit* the variables, and conclusions are derived by interpreting the parameters of these models.

Qualitative Research

The explanatory case study approach, used commonly to explain causal factors of a particular phenomenon, is a qualitative framework suitable for analysis of complex interaction between factors and causal relationships otherwise difficult or impractical to observe through statistical inference [37]. Applied in crash research, all sources of descriptive data for each crash case such as narratives and reports are used in assessing how factors interact and influence the occurrence or severity of crashes. The use of the entirety of descriptive data in- stead of quantitative, tabular data (which may have lost the complexities unique to each particular crash) is ideal in studies with relatively small sample size and which evaluate a large number of crash factors, such as the Heavy-Duty Truck Study.

The product of qualitative research is a conceptual model of how each factor contributes to the risk of crashes. The credibility of such models hinges directly on the researchers' domain knowledge in fields relating to road crashes (human psychology, physics, traffic regulation, automotive engineering, etc.), which serves as the lens used to transform and code raw information, discover themes in the data, and develop a model which details the interpretation of evidence and the researchers' findings on each factor's influence towards crash probability and severity. Naeem et al. [38] outline a systematic thematic analysis process for creating a conceptual model from raw data:

- » **STEP 1:** Transcription, familiarization with the data, and selection of quotations
- » **STEP 2:** Selection of keywords
- » **STEP 3:** Coding of the data
- » **STEP 4:** Development of themes
- » **STEP 5:** Conceptualization through interpretation of keywords
- » **STEP 6:** Development of a conceptual model

The variance in how crash reporting jurisdictions investigate and document their crashes makes qualitative analysis especially useful as it does not mandate alignment of variables in each case. Additionally, the conceptual model building process allows the research to construct evidence by using the combined expertise in each of the research collaborators' domains as lens from which to interpret descriptive crash data. Homes et al. [38] provide examples of the application of qualitative approaches in crash research today.
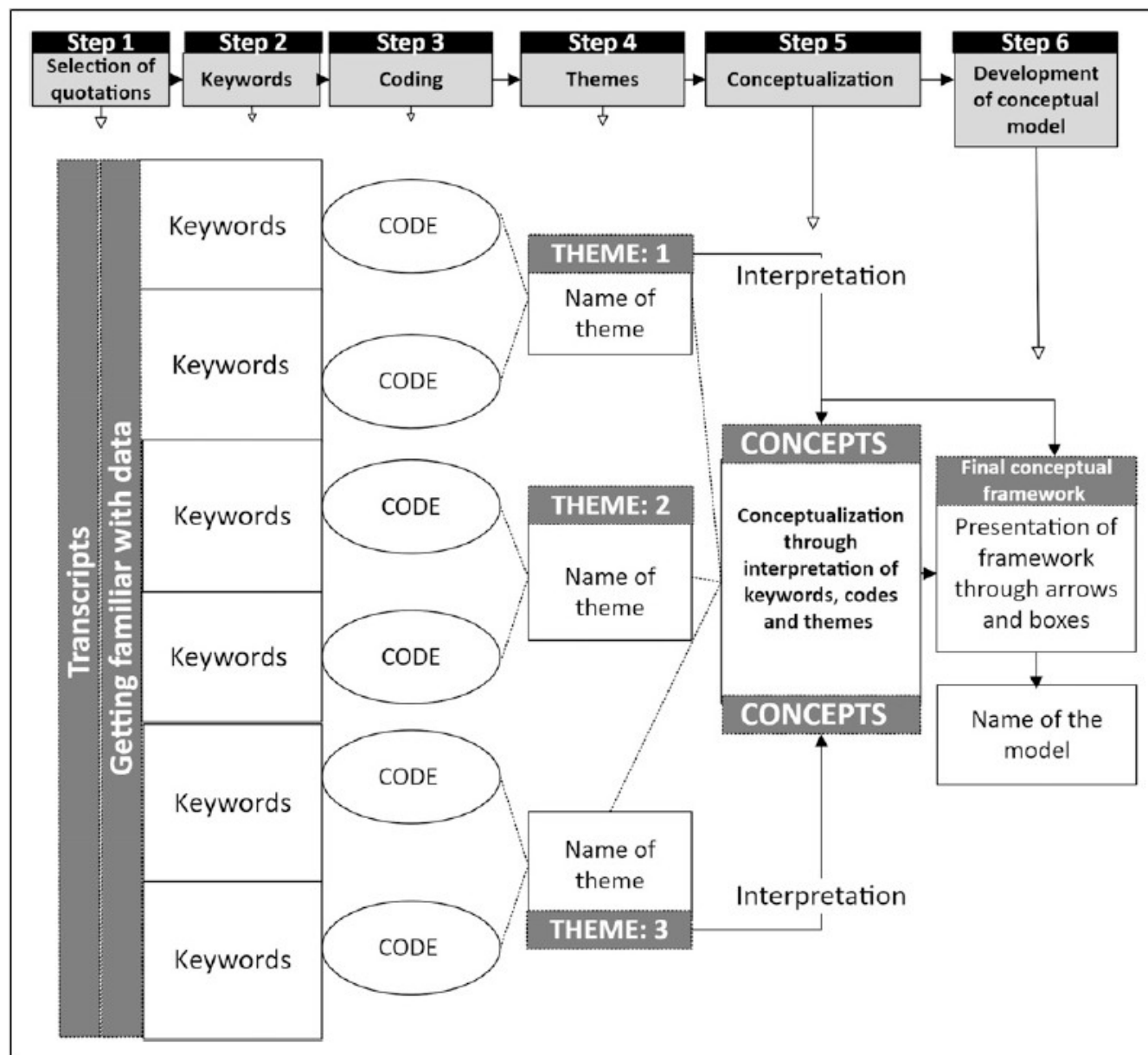

Figure 2. Thematic Analysis Framework for creating a conceptual model [39]

Quantitative Research

Quantitative research seeks statistical evidence of the impact of road factors towards crash risk in the population of interest. Whereas qualitative research uses primarily descriptive data, the application of quantitative approaches in crash research today relies on numeric data to construct mathematical models that describe how factors relate to a particular crash outcome. Whereas qualitative approaches use domain expertise as the primary basis for evidence of findings, quantitative analysis argues from a statistical angle the contribution of factors towards the probability or severity of crashes. This is performed most commonly through the interpretation of parameters of a model that best describes the relationship between the variables in the dataset, as indicated by its predictiveness in forecasting a crash outcome.

Quantitative crash modeling today falls mainly into two categories: frequency and severity modeling. The objective of each is to search for statistical evidence of the contribution of factors (singular or interaction) towards the probability or severity of crashes.

CRASH PROBABILITY MODELING

Crash probability modeling (also known as crash frequency modeling) evaluates how factors influence the frequency or probability of a crash. Data collection in frequency modeling is typically performed through monitoring a highway segment for a period of time, where the sampling unit is defined as an observation of factors and number of crashes for a particular length of time and distance.

For example, a researcher may monitor a 50-mile stretch of highway for 30 days and define an observation as the number of crashes per hour, per mile. Each day, therefore, will yield 1,200 observations (24 hours x 50 miles), which ultimately translates to 36,000 observations per 30-day month. Each observation is represented as one row in the final table of data, where the independent variables are crash factors of interest and the dependent variable is the number of crashes per mile, per hour.

Once the data are collected, the researcher then fits and tests models in search of one or more models that best describe the relationships among the independent and dependent variables of the collected dataset. Researchers use a model's predictiveness, measured through methods such as goodness-of-fit tests or cross validation procedures, to understand how well a model describes the relationship between factors and crash outcomes. The predictiveness forms the basis for the researcher to use the model's parameters as indicators of each factor's contribution toward the probability or frequency of a crash. For example, researchers may find the negative binomial regression as a strong performing predictive model in describing the relationship between factor variables and crash outcomes. Based on this finding, they may interpret the positive coefficient of a variable as indicating that an increase in the variable increases the likelihood of a crash. Dominique Lord and Fred Mannering [38] detail examples of frequency modeling frameworks and their respective characteristics (see Crash Frequency Models, page 68).

CRASH SEVERITY MODELING

Crash severity modeling evaluates how factors interact and influence the severity of a crash. The general process of constructing the model begins similarly to crash probability modeling by transforming raw information into tabular, quantifiable data consisting of independent and dependent variables. Here, however, the sampling unit is one crash occurrence instead of an observation of crashes on a road segment for a defined length of time. The independent variables remain factors of interest, whereas the dependent variable is now a measure of crash severity.

Similar to crash probability modeling, crash severity modeling is typically done by constructing a mathematical model that best describes the relationship between the independent variables (factors) and the dependent variable (a measure of severity, such as fatality) of the crash data. How much a factor contributes to the severity of crashes can then be interpreted from the parameters of the resulting model. For example, in crash severity models where fatality is the measure of severity (a binary outcome where a crash is classified as fatal or non-fatal) and logistic regression is applied, a positive coefficient for a factor may be interpreted as indicating that the factor increases crash severity. Savolainen et al. [41] and Santos et al. [42] provide a comprehensive list of crash severity models as applied in recent literature, and their respective advantages and disadvantages.

CCFP Heavy-Duty Truck Study Data

The suitability of each quantitative methodology depends on the research's objectives and the availability of data needed the approach. The objective of the Heavy-Duty Truck Study is to reduce the number of fatal heavy-duty truck crashes in the U.S. by identifying factors which increase their crash risk. Reducing fatalities can be accomplished by:

1. Preventing the crashes from occurring in the first place, and

2. Reducing their severity in the event of an occurrence.

This section will assess the suitability of each of the quantitative analysis approaches (probability and severity modeling) against CCFP Heavy-Duty Truck Study data.

Primary Data Source

The Heavy-Duty Truck Study seeks to collect a nationally representative sample of at least 2,000 heavy-duty truck crashes in the U.S. While prioritizing fatal crashes, the study will collect a convenience sample non-fatal injury crashes. The final dataset of the Heavy-Duty Truck Study is projected to be a sample of at least 2,000 fatal crashes involving heavy-duty trucks plus a convenience sample of injury crashes involving heavy-duty trucks, consisting of unstructured (primarily narratives and reports) and structured (tabular) data. The structured dataset variables will consist of at least one variable for each research question. Each of the research questions corresponds to at least one crash factor that the study seeks to evaluate.

The Suitability of Crash Frequency Models

To argue from a statistical perspective that a factor increases the probability of a crash requires exposure data. While the definition of the term varies across studies [43], exposure can be thought of as a measure of the opportunity for a crash to occur. An example best illustrates its importance in linking factors and crash frequency.

Suppose a year-long study involving a set number of driver-vehicle pairs is interested in evaluating the impact the time of day has toward the frequency of crashes, and that the year concludes with an equal number of crash occurrences during the day and at night. At a glance, this may imply that the time of day is not a significant contributing factor toward the probability of crashes. Yet suppose data are available which show that, throughout the year, the number of miles traveled during the day is twice that at night. The availability of this measure of exposure allows the researcher to create compelling evidence that driving at night increases the likelihood of crashes, as the nighttime crash rate is twice the rate of daytime crashes.

The traditional practice in crash frequency modeling of monitoring a fixed length of highway is designed to account for exposure. In this sample design, every vehicle which passes through this monitored area travels the same distance. This method allows the researcher to control for exposure while gathering the independent variables associated with each observation. Controlling the exposure variable is a critical basis in establishing the logical validity of crash frequency models.

Applying this data collection methodology to account for exposure and perform crash frequency analysis in the Heavy-Duty Truck Study may be conceivable, but it would be impractical to implement. The study aims to understand factors solely of crashes involving heavy-duty trucks. Applying the methodology in the context of Heavy-Duty Truck Study would mean analyzing factors against the number of crashes involving heavy-duty trucks that occur specifically in the geographical area of data collection. Such a procedure would not be cost-effective as crashes involving heavy-duty trucks are much rarer than crashes involving only passenger vehicles or lighter trucks (MCMIS records contain just over 3 million crashes involving trucks with GVW greater than 26,000 lbs in over three decades—about 100,000 of which were fatal. In comparison, NHTSA records contain over 6.7 million total motor vehicle crashes in 2019 alone).

The issue is amplified because the Heavy-Duty Truck Study seeks a nationally representative sample (which would necessitate monitoring multiple geographical areas and road types) and because it seeks to evaluate a vast number of variables. It would not be feasible to collect this many variables by simply observing during the time it takes for the trucks to pass through the fixed test area (for example, collecting variables about driver condition would necessitate stopping each vehicle that passes through the segment).

For the reasons above, crash frequency modeling is not suitable for the Heavy-Duty Truck Study. The relative rarity of crashes involving heavy-duty truck, the size of number of factors the study seeks to analyze, and data collection restrictions renders the methodology unsuitable for the study.

Suitability of Crash Severity Models

As the CCFP plans to collect a convenience sample of non-fatal crashes (in addition to the 2,000 fatal ones), crash severity modeling approaches may be applied to assess how factors increase the severity of crashes given the availability of sufficient control data available from the convenience sample. Given the dichotomous nature of the measure of severity used as a dependent variable (fatal vs. non-fatal), binary-outcome severity models can be constructed and tested (Assessment of Various Severity Model Performances, page [67] lists all usable frameworks) provided that there is sufficient convenient samples of control crashes such as non-fatal crashes. Parameters of models that best describe the relationship between factors and crash severity can be used as indicators of the significance of each factor's contribution toward the severity of the crashes. Any significant indications can be used in conjunction with results from qualitative analysis to build a holistic model that addresses each of the Heavy-Duty Truck Study research questions.

Analysis Plan

The Heavy-Duty Truck Study seeks to use all available data to answer its research questions, each of which corresponds to a crash factor it seeks to evaluate. The scope of data available includes unstructured and structured data from sources such as partner jurisdictions, existing FMCSA internal databases (such as MCMIS), and accessible external databases.

Frameworks outlined by Yin [37] and Naeem et al. [38] will be used to build qualitative evidence through the interpretation of primarily unstructured data such as crash narratives and reports. By using a qualitative framework, study can capture complex interactions between factors and causal relationships that are impossible to observe through statistical inference. The study can also leverage the collective lens of field expertise to search for themes and understand how factors influence the probability and severity of crashes involving heavy-duty trucks.

Using MCMIS as a primary data source, the study will construct and evaluate linear and non-linear crash severity models and interpret the parameters of best performers in search of statistical evidence of the contribution of factors toward crash severity. Crash Severity Modeling Framework (page 65) illustrates the process of building and evaluating models. Situmorang [44] models the current MCMIS data on crashes involving heavy-duty trucks using an interpretable machine-learning framework and offers indications of how factors currently captured in the database increase the severity of those crashes. Model indications will be used to construct a unified model that details the research's findings on how each analyzed factor contributes to crash risk (see Figure 4 for the overall Heavy-Duty Truck Study plan).
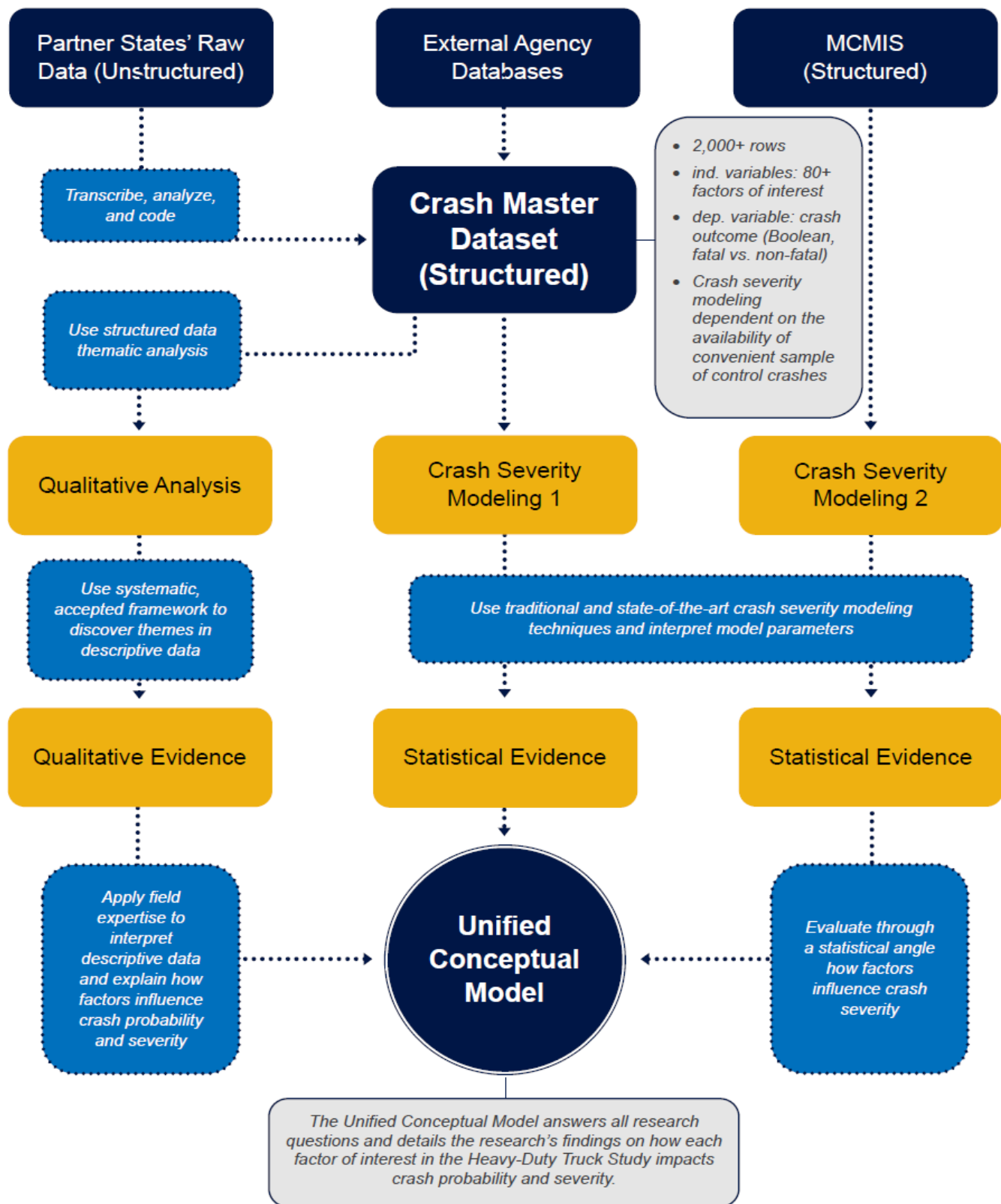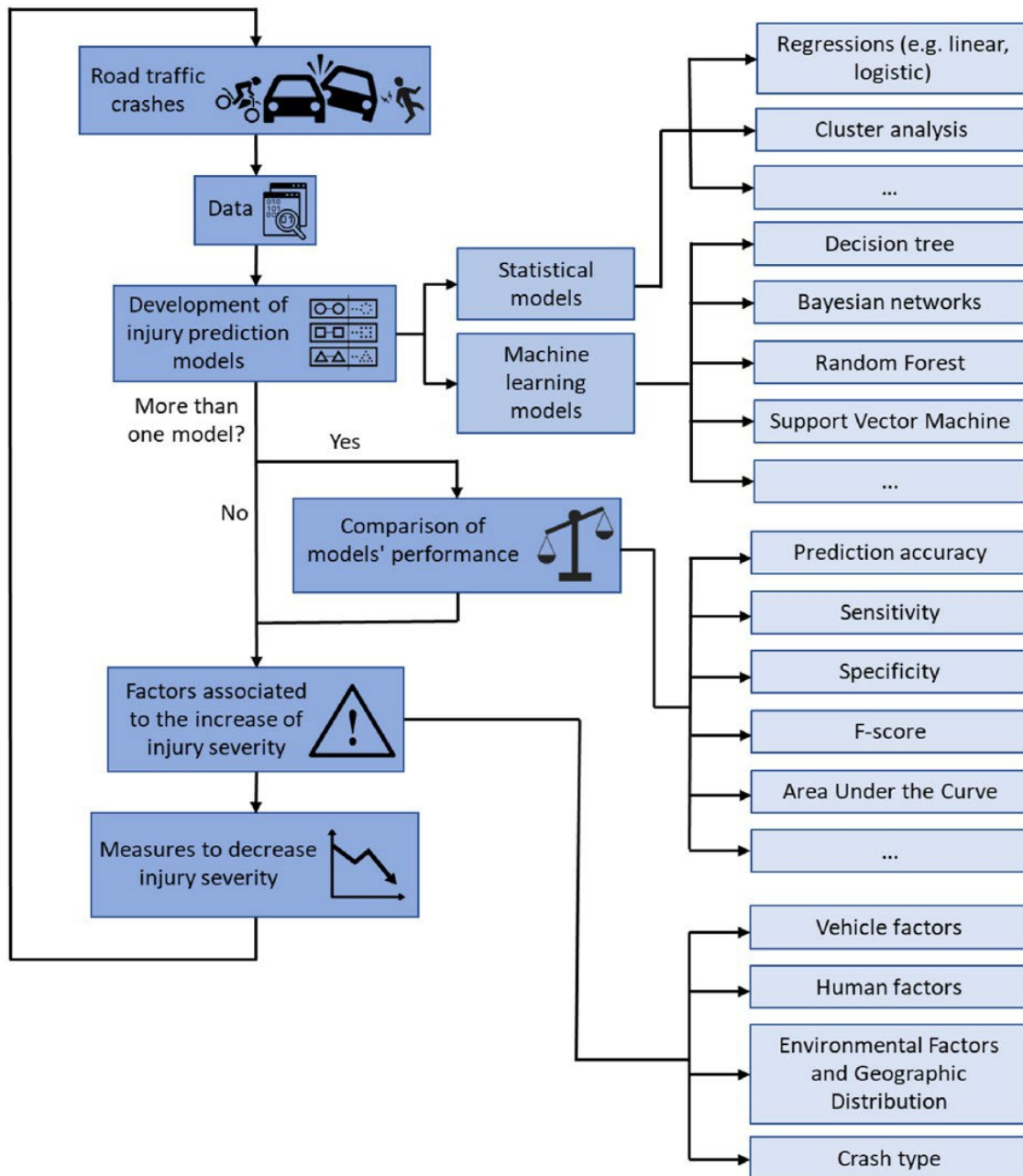
Figure 3. CCFP Heavy-Duty Truck Study Plan

CRASH SEVERITY MODELING FRAMEWORK

Figure 5: Santos et al. [42]

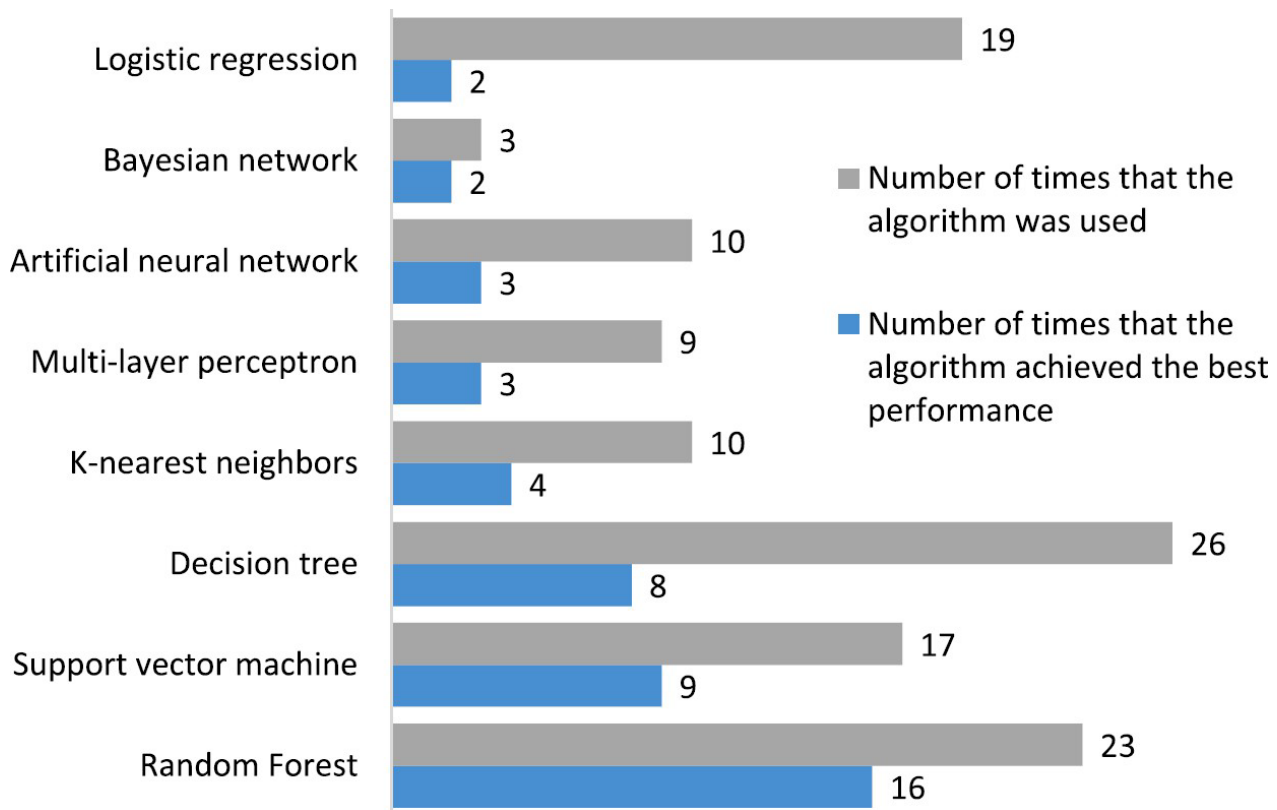## ASSESSMENT OF VARIOUS SEVERITY MODEL PERFORMANCE

Figure 6: Santos et al. analyzed studies which tested various models in the severity model building and selection process [42]. Their analysis indicates that models based on Random Forest, Support Vector Machine, and Decision Tree perform best in predicting crash severity outcomes. Until recently, non-linear methods such as these were considered black-boxes with parameters that are not interpretable. Recent advancements of wrapper functions, such as LIME and SHAP explainers, and glassbox models such as InterpretML's Explainable Boosting Machine, offer ways to interpret state-of-the-art models and how each factor contributes to the severity prediction of the models. These indications can be used as a method to rank how factors influence crash severity.

CRASH FREQUENCY MODELS

Table 12: Lord and Mannering… [40]Table 13. Crash frequency model advantages and disadvantages

Model	Advantages	Disadvantages
Poisson	Most basic model; easy to estimate	May need to determine or evaluate the type of temporal correlation a priori; results sensitive to missing values
Negative Binomial/ Poisson- gamma	Easy to estimate can account for over-dispersion	Relatively complex to implement; may not be easily transferable to other datasets
Poisson- lognormal	More flexible than the Poisson-gamma to handle overdispersion	May not be easily transfer- able to other datasets
Zero- inflated Poisson and negative binomial	Handles datasets that have a large number of zero-crash observations	Cannot handle under- dispersion; can be adversely influenced by the low sample-mean and small sample size bias
Conway– Maxwell– Poisson	Can handle under- and over-dispersion or combination of both using a variable dispersion (scaling) parameter	Complex estimation process; may not be easily transferable to other datasets
Gamma	Can handle under- dispersed data	Complex estimation process; requires formulation of correlation matrix
Generalized estimating equation	Can handle temporal correlation	May need to determine or evaluate the type of temporal correlation a priori; results sensitive to missing values
Generalized additive	More flexible than the traditional generalized estimating equation models; allows non-linear variable interactions	Relatively complex to implement; may not be easily transferable to other datasets
Random- effects	Handles temporal and spatial correlation	May not be easily transfer- able to other datasets
Negative multinomial	Can account for over- dispersion and serial correlation; panel count data	Cannot handle under- dispersion; can be adversely influenced by the low sample-mean and small sample size bias

Model	Advantages	Disadvantages
Random-Parameters	More flexible than the traditional fixed parameter models in accounting for unobserved heterogeneity	Complex estimation process; may not be easily transferable to other datasets
Bivariate/multivariate	Can model different crash types simultaneously; more flexible functional form than the generalized estimating equation models (can use non-linear functions)	Complex estimation process; requires formulation of correlation matrix
Finite mixture/Markov switching	Can be used for analyzing sources of dispersion in the data	Complex estimation process; may not be easily transferable to other datasets
Duration	By considering the time between crashes (as opposed to crash frequency directly), allows for a very in-depth analysis of data and duration effects	Requires more detailed data than traditional crash frequency models; time-varying explanatory variables are difficult to handle
Hierarchical/multilevel	Can handle temporal, spatial and other correlations among groups of observations	May not be easily transfer- able to other datasets; correlation results can be difficult to interpret
Neural network, Bayesian neural net- work, and support vector machine	Non-parametric approach does not require an assumption about distribution of data; flexible functional form; usually pro- vides better statistical fit than traditional parametric models	Complex estimation process; may not be transferable to other datasets; work as black-boxes; may not have interpretable parameters

ABBREVIATIONS

AEB................Automatic Emergency Braking

ANSI ....American National Standard Institute

CCFP ....Crash Causal Factors Program

CDC ....Center for Disease Control and Prevention

CISS ....Crash Investigation Sampling System

CMV...............Commercial Motor Vehicle

CVSA ....Commercial Vehicle Safety Alliance

DOT ....Department of Transportation

EDT................Electronic Data Transfer

FARS ....Fatality Analysis Reporting System

FDR ....False Discovery Rate

FMCSA ....Federal Motor Carrier Safety Administration

GVWR............Gross Vehicle Weight Rating

ICS.................Investigation Capability Score

IT....................Information Technology

LTCCS ....Large Truck Crash Causation Study

MCMIS...........Motor Carrier Management Information System

MCSAP ....Motor Carrier Safety Assistance Program

MMUCC ....Model Minimum Uniform Crash Criteria

NHTSA...........National Highway Traffic Safety Administration

NTSB ....National Transportation Safety Board

PCR ....Police Crash Report

PSU ....Primary Sampling Units

ROM ....Rough Order of Magnitude

BIBLIOGRAPHY

[1] K. Stowe, "Draft Research Questions for the Large Truck Crash Causal Factors Study (LTCCFS)," Volpe Center, project notes, Cambridge, MA, 2023.

[2] J. Guarino, "FMCSA Studies on Crash Causal Factors: Phased Approach Cost Estimate," FMCSA, project notes, Washington, DC, 2022.

[3] K. Stowe and J. Gurarino, "Large Truck Crash Causal Factors Study: Study Approach Alternatives Analysis," FMCSA, draft internal report, Washington, DC, 2021.

[4] J. Guarino, "LTCCFS Scope: Options," FMSCA, project notes, Washington, DC, 2021.

[5] FMCSA, "Supporting Statement Crash Causal Factors Program: Knowledge of Systems and Processes," FMCSA, draft submission for review under Paperwork Reductionc Act, Washington, DC, 2022.

[6] J. L. Gerberding, "Evaluation on Large Truck Crash Causation Study Design," Center for Disease Control and Prevention, Atlanta, GA, 2005.

[7] J. Hedlund, "Statistical Analyses of Large Truck Crash Causation Study Data," Transportation Research Board, Washington, DC, 2003.

[8] D. Blower and K. L. Campbell, "The Large Truck Crash Causation Study," University of Michigan, Ann Arbor, MI, 2002.

[9] M. Zuschlag, "Power Analysis for Large Truck Crash Causal Factors Study (LTCCFS)," Volpe Center, internal report, Cambridge, MA, 2022.

[10] FMCSA, "Large Truck and Bus Crash Facts 2020 (Vehicle Table 5)," FMCSA, Washington, DC, 2020.

[11] M. H. Belzer, "Work-stress factors associated with truck crashes: An exploratory analysis," *The Economic and Labor Relations Review,* vol. 29, no. 3, pp. 289-307, 2018.

[12] D. Blower, "Comparison of Certain Variables in the Large Truck Crash Causation Study With Variables Coded in a Review of the Trucks Involved In Fatal Accidents Case Materials," University of Michigan, Ann Arbor, MI, 2007.

[13] Consolidated Appropriations Act, *Public Law 116 - 260,* 2021.

[14] N. Michel and J. Guarino, "Concept Analysis of Alternatives Large Truck Crash Causal Factors Study (LTCCFS)," FMCSA internal report, Washington, DC, 2022.

[15] Federal Highway Administration, "Law Enforcement Vehicle Identification Guide," 29 October 2020. [Online]. Available: https://ops.fhwa.dot.gov/publications/fhwahop10014/long_f21.htm. [Accessed 1 February 2024].

[16] A.-M. Chouinard, "Medium-duty Vehicle Crash Scenario Framework and Crash Avoidance Applications," Advanced Vehicle Division, Volpe Center, internal report, Cambridge, MA, 2021.

[17] J. Guarino, "LTCCFS Requirements Analysis," FMCSA, project notes, Washington,DC, 2021.

[18] J. Guarino, "Large Truck Crash Causal Factors Study (LTCCFS): Scope Proposal," FMCSA, project notes, Washington, DC, 2021.

[19] American National Standard Institute, "Manual on Classification of Motor Vehicle Traffic Crashes (8th ed.)," American National Standard (ANSI D.16-2017), 2017.

[20] K. Stowe and N. Michel, "ROM Estimates of Study Cost by Scope," FMCSA, internal spreadsheet, Washington, DC, 2021.

[21] J. Michael, C. Chidester and K. Stowe, "Draft LTCCS Update Estimates," Volpe Center, project notes, Cambridge, MA, 2021.

[22] T. Morgan, "Questions from FMCSA," NHTSA, internal correspondence, Washington, DC, 2021.

[23] D. Meyer, "Project Charter Large Truck Crash Causal Factors Study (LTCCFS), CCFP Workgroup Charter," Office of Research and Registration Analysis Division, FMCSA draft internal document, Washington, DC, 2021.

[24] N. Bondy, "Methodology for the Large Truck Crash Causation Study," FMCSA, Washington, DC, 2004.

[25] K. Stowe and T. Lagos, *Meeting notes (remote),* Cambridge, MA, 2022.

[26] M. Zuschlag, "High Level Requirements Deconstruction, Sample Requirements Desconstruct," Volpe Center, internal spreadsheet, Cambridge, MA, 2022.

[27] D. Blower and K. L. Campbell, "Methodology of the Large Truck Crash Causation Study," Federal Motor Carrier Safety Administration, Washington, DC, 2005.

[28] S. Upadhyaya and T. Morgan, "Crash counts for Class 7/8 Large Trucks for 32 PSUs," Volpe Center, MCMIS snapshot spreadsheet, Cambridge, MA, 2021.

[29] J. Guarino and F. Zhang, "Preliminaries of LTCCFS Sample Design," FMCSA, project notes, Washington, DC, 2021.

[30] K. Stowe and T. Lagos, *Meeting notes (remote),* Cambridge, MA, 2022.

[31] M. Zuschlag and K. Conroy, "CCFP Team Working Copy of CCFP State Survey Responses - ICS Scoring 01," Volpe Center, internal spreadsheet, Cambridge, MA, 2024.

[32] D. Meyer, *LTCCFS Sample Design Resources,* Cambridge, MA, 2023.

[33] K. Stowe, *Meeting notes (remote),* Cambridge, MA, 2023.

[34] F. Zhang and T. Morgan, "LTCCFS PSU Sample Selection Options," NHTSA, internal correspondence, Washington, DC, 2021.

[35] FMCSA, "Recommended Reading for the Large Truck Crash Causal Factors Study Requirements Workshop," FMCSA, internal guidance, Washington, DC, 2021.

[36] B. Routhier, *ECM and ELD Information,* Washington, DC, 2021.

[37] R. K. Yin, Case Study Research and Applications: Design and Methods, Sixth ed., Thousand Oaks, CA: SAGE Publications, Inc, 2018.

[38] M. Naeem, W. Ozuem, K. Howell and S. Ranfagni, "A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research," *International Journal of Qualitative Methods,* vol. 22, 2023.

[39] B. D. Holmes, K. Haglund, K. M. M. Beyer and L. D. Cassidy, "Qualitative methods of road traffic crash research in low- and middle-income countries: a review," *International Journal of Injury Control and Safety Promotion ,* no. 2, pp. 194-199, 2019.

[40] D. Lord and F. Mannering, "The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives," *Transportation Research Part A: Policy and Practice,* vol. 44, no. 5, pp. 291-305, 2010.

[41] P. T. Savolainen, F. L. Mannering, D. Lord and M. A. Quddus, "The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives," *Accident Analysis & Prevention,* vol. 43, no. 5, pp. 1666-1676, 2011.

[42] K. Santos, J. P. Dias and C. Amado, "A literature review of machine learning algorithms for crash injury severity prediction," *Journal of Safety Research,* vol. 80, pp. 254-269, 2022.

[43] R. Chapman, "The concept of exposure," *Accident Analysis & Prevention,* vol. 5, no. 2, pp. 95-110, 1973.

[44] P. Situmorang, *Analysis of factors contributing to fatality in U.S. heavy-duty truck crashes,* Unpublished, 2024.

[45] Infrastructure Investment and Jobs Act, *Public Law No: 117-58,* 2021.

[46] D. Meyer, *Meeting notes (remote),* Cambridge, MA, 2022.

[47] J. Guarino, "Large Truck Crash Causal Factors Study (LTCCFS): Scope Proposal," FMCSA, project notes, Washington, DC, 2021.