

Information Collection Request Supporting Statements: Part B
Investigation-Based Crash Data Studies
OMB Control No. 2127-0706

Abstract:¹

The National Highway Traffic Safety Administration (NHTSA) is seeking approval from OMB of this information collection request (ICR) for extension with modification of its currently approved information collection for Investigation-Based Crash Data Studies. NHTSA is authorized by 49 U.S.C. § 30182 and 23 U.S.C. § 403 to collect data on motor vehicle traffic crashes to aid in the identification of issues and the development, implementation, and evaluation of motor vehicle and highway safety countermeasures. The information collected serves to identify and develop safety countermeasures that will reduce the severity of injury and property damage caused by motor vehicle crashes. These Investigation-Based Crash Data Studies, Crash Investigation Sampling System (CISS), Special Crash Investigations (SCI), and Special Studies, involve voluntary information collections through which NHTSA collects detailed data on real world motor vehicle crashes. Specifically, these systems collect data on vehicle safety system performance, occupant injury information including their kinematic interaction with interior components and scene geometry, marking and traffic controls.

Respondents are police agencies that collection information on police-reported motor vehicle crashes, employees of tow yards where crashed vehicles are stored, people involved in these crashes, and hospitals with medical records for the people injured in the crash.

For the standard investigation-based crash data studies acquisition process, once a crash has been selected for investigation, crash technicians or investigators locate, visit, measure, and photograph the crash scene; locate, inspect, and photograph vehicles; conduct a telephone or personal interview with the involved individuals or surrogate; and obtain and record crash injury information received from various medical data sources.

These information collections support NHTSA's mission to save lives and prevent injuries due to traffic crashes. The data collected from these systems are used to describe and analyze circumstances, mechanisms, and consequences of serious motor vehicle crashes in the United States. Additionally, these data are used by NHTSA to identify the primary factors related to the source of crashes and their injury outcomes, develop and evaluate effective safety countermeasures, the establishment and enforcement of motor vehicle regulations, that reduce the severity of injury and property damage caused by motor vehicle crashes.

The previous request for CISS (2017) indicated 5,605 burden hours, this request increases the burden to 8,170. The request for the collection of information is revised due to a) Increasing the number of crashes investigated by Crash Technicians for 2021 and future years, b) adding Special Study crashes into this package, c) adding tow yards, hospitals and law enforcement burden and d) adding Special Crash Investigation (SCI) crashes into this package. The combined impact is an increase of 2,565 burden hours to NHTSA's overall total.

Although this ICR covers NHTSA's three Investigation-Based Crash Data Studies, not all of these systems use a statistical sample. SCI and some Special Studies are not based on a static sample and will not be discussed in the rest of this document. CISS and Special Studies that do utilize a statistical sample and will be explained further in this document.

¹ The Abstract must include the following information: (1) whether responding to the collection is mandatory, voluntary, or required to obtain or retain a benefit; (2) a description of the entities who must respond; (3) whether the collection is reporting (indicate if a survey), recordkeeping, and/or disclosure; (4) the frequency of the collection (e.g., bi-annual, annual, monthly, weekly, as needed); (5) a description of the information that would be reported, maintained in records, or disclosed; (6) a description of who would receive the information; (7) the purpose of the collection; and (8) if a revision, a description of the revision and the change in burden.

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Describe the potential respondent universe and any sampling or other respondent selection methods to be used.

NHTSA's CISS and some Special Studies are investigation-based crash data studies which use statistical sampling methods to select cases. They are classified as two types according to their crash selection methods:

- CISS: A probability sample of in-depth investigation of crashes involving passenger vehicles towed from the scene in the US. CISS collects data at the crash level through scene analysis, vehicle level through vehicle damage assessment, and occupant level through interviews together with injury sources. This is NHTSA's primary crash investigation study.
- Special Studies that use NHTSA's sampling systems FARS, CRSS, or CISS: A probability sub-sample is selected from one of the NHTSA's existing probability crash samples. Three major NHTSA surveys can be used to select sub-samples for further investigation-based studies: The Fatality Analysis Reporting System (FARS), the CISS, and the Crash Report Sampling System (CRSS).

In the following, we explain each of the two selection methods in detail.

(1) CISS

CISS is NHTSA's primary investigation-based crash data collection survey. It is a standalone probability sampling system. CISS sample data supports a wide range of analysis needs.

The CISS universe is the set of police-reported motor vehicle crashes on a traffic way involving a passenger vehicle towed from the scene resulting in a Police Accident Report (PAR)². The estimated CISS population size is about 2.7-2.8 million a year. The CISS sample is selected through a stratified multi-stage cluster scheme as follows.

Cluster

Divide the country into geographic units called Primary Sampling Units (PSUs). A PSU is a county or group of counties. PSUs were formed as groups of adjacent counties with end-to-end distance of no more than 65 miles for urban areas and 130 miles for rural areas to ensure efficient data collection. PSUs were also formed in such a way that there is a 90% of chance that there are at least 5 fatal crashes involving a passenger vehicle within the PSU boundary each year. PSU formation respects the Census region and urbanicity boundaries. Some outlying areas of Alaska and small islands of Hawaii were excluded. There is a total of 1,784 PSUs in the PSU frame.

PSU Strata

A composite measure of size (MOS) is assigned to each PSU in the frame. PSU MOS is the combination of the estimated crash counts of seven types of crashes (fatal crashes, incapacitated recent model year crashes, non-incapacitated injury recent model year crashes, no injury recent model year crashes, incapacitated middle model year crashes, non-incapacitated injury middle model year crashes, incapacitated older model year crashes) defined by injury severity and vehicle model year that were assembled based on data needs.

The PSU frame is stratified into 8 primary PSU strata by two variables – region (Northeast, West, South, and Midwest) and urbanicity (urban and rural). Within each primary stratum, PSUs are further stratified by other secondary stratification variables such as total road miles by road type, total expected number of crashes, and vehicle miles traveled. PSUs with similar characteristics were grouped into secondary strata by attempting to make the stratum MOS approximately equal and minimizing the between-PSU variance within stratum. As the result, a total of 24 PSU strata were formed.

² A Police Accident Report is also known as a Police Crash Report (PCR) in some jurisdictions.

First Stage (PSU Sampling)

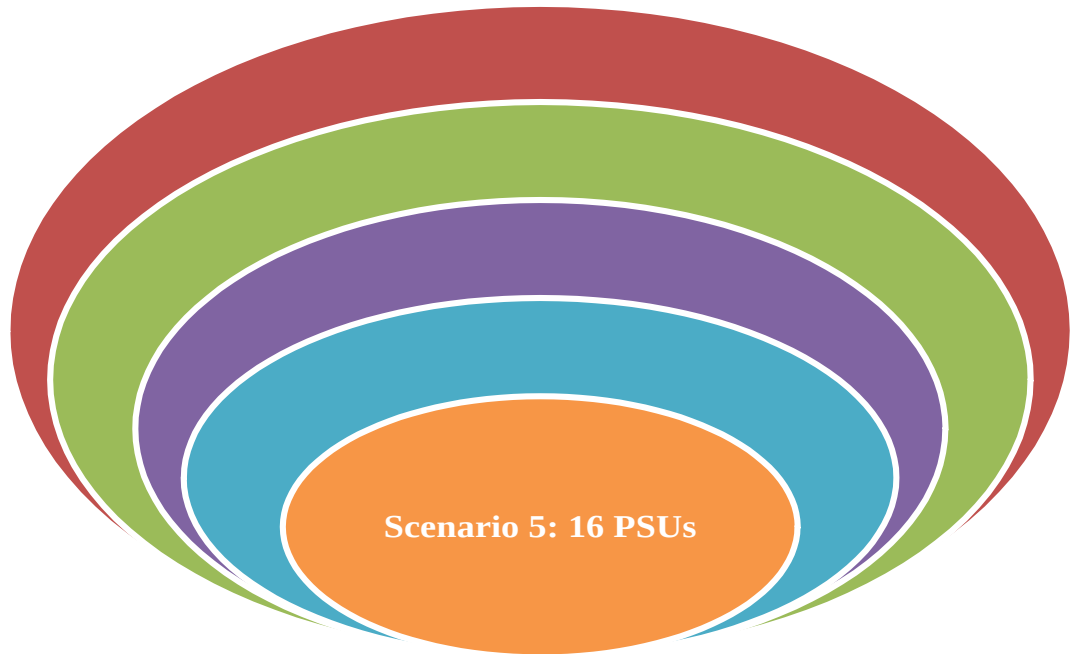
From each of the 24 PSU strata, 2 PSUs were selected by a probability proportional-to-size (PPS) sampling method with the PSU MOS. In addition, one large PSU was selected with certainty. This resulted in a total of 49 PSUs as the first scenario PSU sample.

To cope with the uncertainty of future budgetary fluctuations, a sequence of nested PSU sub-samples were selected from the first scenario PSU sample. For example, the second scenario PSU sample has 40 PSUs and 20 PSU strata. To select the second scenario PSU sample, 4 PSU strata in the first scenario sample selection were collapsed with other strata to form total 20 strata. Then 2 PSUs were selected from the first scenario PSU sample in each of the 20 strata. A total of 5 scenarios of PSU samples were developed with sample size ranging from 16 (scenario 5) to 49 (scenario 1) – each is a sub-sample of the previous scenario PSU sample (see figure and table below). Each PSU sample in 5 scenarios or a PSU sample with any sample size between the scenarios can be used as a PSU sample and the resulting PSU selection probability is PPS. This approach produces nested PSU samples, and PSU sample size can be changed without reselecting the sample. Typically, non-certainty PSU selection probability in a stratum is:

$$\pi_i = \frac{2 MOS_i}{\sum_i MOS_i}$$

Here i is an index for a non-certainty PSU in a stratum, MOS_i is PSU MOS, summation is over the non-certainty PSUs in the stratum.

Figure 1. Nested PSU Samples for CISS



Scenario	Number of PSU Strata	Number of Certainty PSU	Total Number of Sampled PSUs	Sampled Non-Certainty PSU per Stratum
1	24	1	49	2
2	20	0	40	2
3	16	0	32	2
4	12	0	24	2
5	8	0	16	2

Second Stage (PJ Sampling)

The Secondary Sampling Units (SSU) are police jurisdictions (PJs). A PJ frame is developed for each PSU and is the totality of PJs generating PARs in the sampled PSU. A coarse MOS, a combination of crash counts in six crash categories of interest (total crashes, fatal crashes, injury crashes, pedestrian crashes, motorcycle crashes, and commercial motor vehicle crashes), is assigned to each PJ and is called the coarse PJ MOS. Within each PSU, PJs are stratified into three PJ strata by their coarse MOS: certainty stratum, large MOS stratum, and small MOS stratum. Then, a finer composite MOS is assigned to each PJ. To define the finer and the final PJ MOS, first, the ten CISS PAR domain counts (see Table 1) were estimated. The SSU MOS is then defined as follows:

$$MOS_{j \setminus i} = \sum_{s=1}^{10} \frac{n_{++s}}{n} \frac{N_{ijs}}{N_{++s}}$$

where i = an index for PSU
 j = an index for PJ
 s = an index for PAR domain ($s = 1, \dots, 10$)
 n = the desired total sample size of crashes
 n_{++s} = the desired sample size of crashes in PAR domain s N_{++s} = the estimated population of crashes in PAR domain s
 N_{ijs} = the estimated population number of crashes in PAR domain s , PJ j and PSU i

Table 1. CISS PAR Domains, Descriptions, Allocations and Population Sizes

CISS PAR Domain	Description	Target Percent of Sample Allocation	Estimated Population	Population Percent
1	<ul style="list-style-type: none"> At least one occupant of towed passenger vehicle is killed 	5%	9,576	0.51%
2	<ul style="list-style-type: none"> Crashes not in Stratum 1 involving: A recent model year passenger vehicle in which at least one occupant is incapacitated 	10%	17,304	0.93%
3	<ul style="list-style-type: none"> Crashes not in Stratum 1 or 2 involving: A recent model year passenger vehicle in which at least one occupant is non-incapacitated, possibly injured or injured but severity is unknown. 	20%	162,037	8.71%
4	<ul style="list-style-type: none"> Crashes not in Stratum 1-3 involving: A recent model year passenger vehicle in which all occupants are not injured 	15%	325,332	17.48%
5	<ul style="list-style-type: none"> Crashes not in Stratum 1-4 involving: A mid-model year passenger vehicle in which at least one occupant is incapacitated 	6%	23,739	1.28%
6	<ul style="list-style-type: none"> Crashes not in Stratum 1-5 involving: A mid-model year passenger vehicle in which at least one occupant is non-incapacitated, possibly injured or injured but severity is unknown 	12%	210,407	11.31%
7	<ul style="list-style-type: none"> Crashes not in Stratum 1-6 involving: A mid-model year passenger vehicle in which all occupants are not injured 	10%	418,702	22.51%
8	<ul style="list-style-type: none"> Crashes not in Stratum 1-7 involving: A older model year passenger vehicle in which at least one occupant is incapacitated 	6%	28,690	1.54%

9	<ul style="list-style-type: none"> Crashes not in Stratum 1-8 involving: A older model year passenger vehicle in which at least one occupant is non-incapacitated, possibly injured or injured but severity is unknown. 	10%	220,815	11.87%
10	<ul style="list-style-type: none"> Crashes not in Stratum 1-9 involving: A older model year passenger vehicle in which all occupants are not injured 	6%	443,151	23.83%
Total		100%	1,859,752	100%

Source: Estimated from 2011 CDS data.

Note: This table uses the following definitions:

Recent model year: 4 years old or newer

Mid-model year: 5-9 years old

Older model year: 10 years old or older

Therefore, PJs with larger high interest crash compositions (as defined by the oversampled domains such as domains 1, 2, and 3 in Table 1) are selected with larger probabilities. A PJ sample is selected using Pareto sampling with the finer PJ MOS.

Pareto sampling method produces an approximate PPS sample and handles the PJ frame changes by minimizing changes to the existing sample. Pareto sampling method is applied to the PJ sample selection for each of the two non-certainty PJ strata (i.e., large MOS stratum and small MOS stratum) within the sampled PSU i , as following:

Step 1: Generate a permanent uniform random number $r_{ihj} \sim U(0,1)$ for each PJ j in the PJ stratum h of PSU i .

Step 2: Identify certainty PJs by the condition for each PJ stratum h :

$$\frac{m_{ih} MOS_{ihj}}{\sum_{j=1}^{M_{ih}} MOS_{ihj}} \geq 1$$

Here m_{ih} is the PJ sample size of stratum h determined by the sample allocation study described below and the budget constraint and M_{ih} is the PJ frame size for PJ stratum h within PSU i . MOS_{ihj} is the finer and the final PJ MOS.

Step 3: Identified certainty PJs are set aside, and PJ sample size is reduced by the number of certainty PJs. This process is repeated to the remaining PJs based on the reduced PJ sample size until there is no more certainty PJs. Let the total number of certainty PJs be m_{ch} . For the remaining $M_{ih} - m_{ch}$ non-certainty PJs in the frame, calculate PPS inclusion probability with non-certainty sample size $m_{ih} - m_{ch}$:

$$p_{ihj} = \frac{(m_{ih} - m_{ch}) MOS_{ihj}}{\sum_{j=1}^{M_{ih} - m_{ch}} MOS_{ihj}}$$

Step 4: Calculate the transformed random numbers:

i

Step 5: Sort the transformed random number from the smallest to the largest. The m_c certainty PJs plus the first $m_{ih} - m_{ch}$ non-certainty PJs on the above sorted list are the PJ sample for PJ stratum h within PSU i .

Pareto sampling is approximately PPS. The PJ selection probability of non-certainty PJs is:

$$\pi_{j \vee i h} \approx p_{i h j}$$

Third Stage (PAR Sampling)

The tertiary sampling units (TSUs) are PARs. PAR sample is selected from each sampled PJ (i.e., SSU) in each sampled PSU on a weekly basis. Each week in each selected PSU, technicians visit the selected PJs and list PARs recorded at that jurisdiction since the last visit. During the PAR listing process, PARs are listed into 10 PAR domains (defined in the Table 1). For some large PJs with large number of PARs, only a systematic sample of PARs is listed. This process is referred as PAR sub-listing. If one of every L PARs is listed in PJ j of PSU i , the sub-listing factor is L and the sub-listing probability for sub-listed PAR l is:

$$\pi_{l \vee i j} = \frac{1}{L}$$

For PJs with small number of PARs, there is no sub-listing (i.e., $L=1$). After PARs are listed from all sampled PJs, all listed PARs in the same PSU are pooled together. A PAR MOS is assigned to every listed PAR. PAR MOS is the multiplication of a PAR MOS factor, the PJ weight, and the sub-listing factor. PAR MOS factors are determined by simulation to ensure the predetermined PAR sample allocation (see Table 1) by the 10 PAR domains can be achieved.

A PAR sample of average 2 crashes per technician in each PSU every week is selected using Pareto sampling with the PAR MOS. The Pareto sampling produces a scalable sample, so it is able to replace the non-responding cases. A sampled case is determined as a non-responding case if a case vehicle which defines the case's PAR domain is not available for data collection because it is repaired or removed etc. Under the Pareto sample selection, the PAR selection probability of a non-certainty PAR k is:

$$\pi_{k \vee i j v l} = \frac{m_{i v}^{\square} MOS_{i j v l k}}{\sum_k MOS_{i j v l k}}$$

Here $m_{i v}^{\square}$ is the non-certainty PAR sample size for PSU i in week v , $MOS_{i j v l k}$ is the PAR MOS. The summation is over all listed non-certainty PARs of the week in the PSU.

The overall selection probability is:

$$\pi_{i j v l k} = \pi_i * \pi_{j \vee i} * \pi_{l \vee i j} * \pi_{k \vee i j v l}$$

The design weight is the inverse of $\pi_{i j v l k}$.

Sample Allocation Study

PSU, PJ and PAR sample sizes were determined using the optimization model that minimizes the overall variance given cost and variance constraints, thereby ensuring that the sample design for CISS will be at least as precise as the Crashworthiness Data System (CDS) for certain key estimates that were identified. Seven ($G = 7$) key variables were identified to be considered in the objective function. These key variables are:

- Crash level variables: Rear-end, Head-on, Angle,
- Vehicle level variable: Roll over, and
- Occupant level variables: Fatality, Incapacitating injury, Non-incapacitating injury

In the optimization model, a three-stage simple random sampling without replacement was used for simplicity. The optimization model consists of the objective function, cost constraint, and variance constraints as following:

$$\text{Minimize: } \sum_{g=1}^G V_{CISS}(\bar{\bar{y}}_g) = \sum_{g=1}^G \left\{ \frac{S_{1,g}^2}{n} \left(1 - \frac{n}{N}\right) + \frac{S_{2,g}^2}{nm} \left(1 - \frac{m}{M}\right) + \frac{S_{3,g}^2}{nmk} \left(1 - \frac{k}{K}\right) \right\}$$

$$\text{Subject to: } C = C_0 + nC_1 + nmC_2 + nmkC_3,$$

$$V_{CISS}(\bar{\bar{y}}_g) = \frac{S_{1,g}^2}{n} \left(1 - \frac{n}{N}\right) + \frac{S_{2,g}^2}{nm} \left(1 - \frac{m}{M}\right) + \frac{S_{3,g}^2}{nmk} \left(1 - \frac{k}{K}\right)$$

$$\leq V_{CDS}(\bar{\bar{y}}_g), \text{ for } g=1, \dots, G.$$

$$mk \geq l.$$

where

g : Subscript of the identified key estimate, $g=1, \dots, G$. Here $G=7$.

$\bar{\bar{y}}_g$: Identified key proportion estimate.

n, m, k : Optimal sample sizes of PSUs, PJs, and cases (PARs) to be determined.

N : Population size of PSUs

M : Average population size of PJs.

K : Average population size of PARs

$V_{CISS}(\bar{\bar{y}}_g)$: Variance of the identified key estimate $\bar{\bar{y}}_g$ in CISS?

$S_{1,g}^2, S_{2,g}^2, S_{3,g}^2$: Variance components at PSU-, PJ-, and PAR level.

C, C_0, C_1, C_2, C_3 : Total, fixed, PSU-, PJ-, and PAR level cost coefficients.

$V_{CDS}(\bar{\bar{y}}_g)$: Variance of the identified key estimate $\bar{\bar{y}}_g$ in the current system (NASS CDS).

l : known case load, i.e., the weekly PAR sample size for data collection for each sampled PSU.

The optimum PSU sample size varies depending on the budget. With 2 technicians per PSU, optimum PJ sample size is 7 PJs per PSU.

In the 2019 CISS, 233 PJs were sampled from the 32 sampled PSUs (scenario 3 in Figure 1). Among the sampled PJs, 228 PJs cooperated and 5 PJs did not cooperate. CISS selected 3,090 PARs from the 76,365 PARs listed in the 228 PJs. Among the selected cases, 2,781 cases were investigated for data collection while 301 cases were non-responding and 8 cases were identified as out-of-scope. 301 replacement cases were selected for 301 the non-responding cases. This dramatically increases the effective sample size and data quality. There were 293 PJs that are in the sampled PSUs but were not selected into the PJ sample. NHTSA collects crash counts by the PAR domains from these non-sampled PJs to improve the accuracy of the national estimates. The following table summarizes the sample sizes and the national estimate of CISS-applicable police reported crashes for the calendar year 2019.

Sampled PSUs	32
Sampled PJs	233
Cooperating PJs	228
Listed PARs	76,365
Sampled PARs (including non-responding cases and replacement cases)	3,090
Investigated cases	2,781
National Estimate of CISS Applicable Police-Reported Motor Vehicle Crashes	2,736,257

The CISS weights are created in the following steps:

Step 1: Calculate the base weights (the inverse of selection probabilities) at all three stages (PSU, PJ, and PAR).

Step 2: Adjust the base weights for PJ non-response using the following adjustment factor. If there is any non-

responding PJs for week v in PSU i , calculate a PJ non-response adjustment factor:

$$adj_{iv} = \frac{\sum_{j \in s_{iv}} w_{jvi} MOS_{ij}}{\sum_{j \in r_{iv}} w_{jvi} MOS_{ij}}$$

Here w_{jvi} is the PJ base weight, s_{iv} is the set of all sampled PJs and r_{iv} is the set of responding PJs in PSU i for week v . MOS_{ij} is the PJ MOS used in the PJ sample selection.

Step 3: Adjust the PAR non-response. PAR non-response adjustment cells were formed by 3 PAR domain groups defined in Table 1 (Domain 1, 3, 5, 6, and 8; Domain 2, 7, and 9; Domain 4 and 10) within PSU. If a cell has less than 6 responding PARs and at least one non-responding PAR, PAR domain groups are combined so that the cell has more responding PARs. The inverse of the weighted PAR response rate of the cell is used as the non-response adjustment factor for all the responded PARs in the cell.

Step 4: Calibrate the PJ and the PAR weights using the total number of PARs by PAR domain for each sampled PSU to further correct for potential Pareto weighting error, nonresponse and coverage biases, and increase the precision of the estimates.

Step 5: Calibrate PSU weights to capture population shift.

Step 6: Truncate the large case weights. Case weights larger than 3 percent of the PAR domain weight total are truncated to 3 percent of the PAR domain (defined in Table 1) weight total and the excessive weights are redistributed to other untruncated weights in the same PAR domain (a form of additional calibration).

Step 7: Adjusted Jackknife replicate weights are created. These adjusted Jackknife replicate weights are subjected to the same weighting process above therefore they capture the impacts of weight adjustments in the variance estimation.

Because of the low PSU sampling rate (number of sampled PSU divided by the total number of PSUs in the population: $32/1784=1.8\%$), we choose to use the Jackknife variance estimation method for the CISS variance estimation method. Final weight and the adjusted Jackknife replicate weights are provided with the CISS analysis file.

(2) Special Studies that use NHTSA’s existing sampling systems

Selecting and implementing a standalone probability sample is costly and time consuming. For many Special Studies it is infeasible to select a standalone sample due to the budget and time constraints. Instead, it is sensible for NHTSA to treat one of the NHTSA’s existing probability samples (FARS, CISS, and CRSS) as a first phase sample and select a second-phase probability sub-sample to conduct in-depth crash investigations.

Three major NHTSA sampling systems can be used to select second-phase samples for further investigation-based studies: The Fatality Analysis Reporting System (FARS), the CISS, and the Crash Report Sampling System (CRSS). NHTSA plans to use the PSU, PJ, and PAR samples of these surveys to select sub-samples for Special Studies.

FARS

FARS is an annual national census of fatal motor vehicle traffic crashes. FARS annual fatal traffic crash records can be used to select a sub-sample of fatal crashes for further in-depth investigation. For example, a random sample of fatal crashes resulting in a pedestrian’s death or a random sample of fatal crashes involving a medium duty truck can be selected from the previous year’s FARS records followed by further in-depth investigation.

CISS

CISS, as described above, is a three-stage probability sample of crashes involving passenger vehicles. It has established cooperation with the local law enforcement and has hired and trained NHTSA’s contracting crash investigators. Establishing the CISS data collection system, including the PSU and PJ sample selection, soliciting cooperation from the local law enforcement, hiring and training crash investigators etc., is costly and time

consuming. Once the CISS has been set-up, it is sensible to use the CISS’s infrastructure to conduct Special Studies. A Special Study may conduct cost effective data collection by utilizing the CISS resources and the cooperation with the regional agencies. NHTSA plans to use the CISS sample design and data collection technicians to conduct Special Studies to collect crash data to answer questions on specific issues related to vehicle and highway safety.

The extent of use of the CISS sample design and/or infrastructure can be different depending on the focus and budget of the Special Study. Special Studies can utilize the CISS infrastructure in two different ways by either (1): using the CISS trained crash investigators or (2) using both the CISS sample design and crash investigators. These options are discussed below in more detail.

First, some Special Studies do not use the CISS sample design, instead may use NHTSA’s other system such as FARS or CRSS to select samples. However, special studies utilize the CISS trained crash investigators and contractor to collect, interpret, and edit the crash information. For example, NHTSA will conduct Special Studies about medium trucks and pedestrians. These studies will use the FARS for the sample selection but utilize the CISS resources for the data collection. Specifically, four hundred in-scope (crashes involving a medium truck for the medium truck study and crashes involving pedestrians for the pedestrian study) fatal crashes will be selected from NHTSA’s Fatality Analysis Reporting System (FARS) file. The sampled fatal crashes will then be sent to the PJs that produced the crash reports for more detailed crash information. Upon receiving the requested information from the PJs, the CISS crash investigator will review the case and perform data collection for the selected cases. For these Special Studies, the sample design and weighting procedures are independent from the CISS, but the CISS’s crash investigators’ expertise are essential for the data collection.

Second, NHTSA Special Study may use both the CISS sample design and its crash investigators. In these studies, PSU, PJ, or PAR sub-sample will be selected from the CISS PSU, PJ, or PAR samples and the CISS trained crash investigators will collect the data.

CRSS

CRSS is a major record-based crash data collection system.³ CRSS is a multi-stage complex survey of police crash reports. Like the CISS, the CRSS sample is comprised of PSU, PJ, and PAR samples. NHTSA plans to select sub-samples from the CRSS samples for Special Studies. A sub-sample of PSUs, PJs, or PARs can be selected from the current CRSS PSU, PJ, or PJ samples to conduct investigation of the crashes.

In the following, we describe in detail the CRSS population and how the CRSS PSU, PJ, and PAR samples were selected.

The purpose of the CRSS is to provide annual, nationally representative estimates of police-reported motor vehicle crashes as well as characteristics of these motor vehicle crashes. The police accident report (PAR) is the sole source of data for the CRSS. CRSS population is the set of police-reported motor vehicle crashes on a traffic-way (strata 2 – 10 of Table 2).

Table 2. CRSS PAR Strata, Target Sample Allocation, and Population Sizes

CRSS PAR Stratum	Description	Target Percent of Sample Allocation	Estimated Population (GES 2011) ****	Population Percent
1	An in-scope Not-in-Traffic Surveillance (NTS) crash (take all)*			
2	Crashes not in Stratum 1 in which: Involves a killed or injured (includes injury severity unknown) non-motorist	9%	119,579	2.2%

³ Additional details about CRSS and how NHTSA collects this information are available in the supporting statements for the ICR with OMB Control No. 2127-0714.

CRSS PAR Stratum	Description	Target Percent of Sample Allocation	Estimated Population (GES 2011) ****	Population Percent
3	Crashes not in Stratum 1 or 2 in which: Involves a killed or injured (includes injury severity unknown) motorcycle or moped rider	6%	76,513	1.4%
4	Crashes not in Stratum 1-3 in which: At least one occupant of a late model year passenger vehicle** is killed or incapacitated	4%	22,272	0.42%
5	Crashes not in Stratum 1-4 in which: At least one occupant of an older passenger vehicle*** is killed or incapacitated	7%	84,659	1.6%
6	Crashes not in Stratum 1-5 in which: at least one occupant of a late model year passenger vehicle** is injured (including injury severity unknown)	14%	330,619	6.2%
7	Crashes not in Stratum 1-6 in which: involved at least one medium or heavy truck or bus (includes school bus, transit bus, and motor coach) with GVWR 10,000 lbs. or more	6%	302,781	5.7%
8	Crashes not in Stratum 1-7 in which: at least one occupant of an older passenger vehicle*** is injured (including injury severity unknown)	12%	800,390	15.0%
9.	Crashes not in Stratum 1-8 in which: involved at least one late model year passenger vehicle**, AND No person in the crash is killed or injured	22%	1,511,371	28.4%
10	Crashes not in Stratum 1-9: This includes mostly PDO crashes involving a non-motorist, motorcycle, moped, and passenger vehicles that are not late model year** and any crashes not classified in strata 1-9.	20%	2,078,263	39.0%

*: NTS cases are not in the scope of CRSS. They are set aside for NTS analysis.

** : Late model year passenger vehicle: passenger vehicle that are 4 years old or newer.

***: Older passenger vehicle: passenger vehicle that are 5 years old or older.

****: 2011 GES estimates were the most recent estimates at the time of the CRSS sample design.

The estimated CRSS population size (strata 2–10 of Table 2) was about 6.7 million in 2019. CRSS selects a sample from the population through a stratified multi-stage cluster scheme as follows:

First Stage (PSU Sampling)

The country is divided into geographic units called Primary Sampling Units (PSUs). A PSU is a county or group of counties and serves as a cluster. PSUs were formed as groups of adjacent counties subject to a minimum measure of size (MOS) condition to ensure enough cases will be sampled from each PSU and weights are approximately equal within each PAR stratum defined in Table 2. The CRSS PSU MOS was defined as:

$$MOS_i = \sum_{s=2}^{10} \frac{n_{++s}}{n} \frac{N_{i+s}}{N_{++s}}$$

where

s = the PAR stratum defined in Table 2.

n = the desired total sample size of crashes

n_{++s} = the desired sample size of crashes in the PAR stratum s

N_{++s} = the estimated population count of crashes in the PAR stratum s

N_{i+s} = the estimated population count of crashes in the PAR stratum s and PSU i .

In the formula, n_{++s}/n is the desired PAR strata sample allocation (the “Target Percent of Sample Allocation” column in Table 2), and N_{i+s}/N_{++s} is the relative estimated population counts of PSU i for PAR stratum s . In this way, a PSU with a larger high interest (as defined by the oversampled PAR strata defined in Table 2) combination of estimated population counts of all PAR strata has a larger MOS.

PSU formation respects US Census region and urbanicity boundaries. While 23 outlying counties in Alaska and three counties in Hawaii were excluded, the rest of the country is included in the PSU frame. There are 707 CRSS PSUs in the PSU frame.

The PSU frame was then stratified into eight primary PSU strata by two variables – region (Northeast, West, South, and Midwest) and urbanicity (urban and rural). Within each primary stratum, PSUs were further stratified by secondary stratification variables such as vehicle miles traveled, crash rate, truck miles traveled, and crash rate by road type. PSUs with similar characteristics were grouped into secondary strata with approximately equal MOS sizes. Secondary strata groupings were also based on minimizing the between-PSU variance within a stratum. As the result, 50 PSU strata were formed as indicated in Table 3.

Table 3. CRSS PSU Strata, PSU Population Counts, and Sample Size

PRIMARY STRATA	STRAT ID*	VMT_RATE_IMP**		TOT_CRASH_RATE**		TRK_MI_RATE**		ROAD_TYPE_RATE**		Number of PSUs	PSU Sample Size
		Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower		
1	101	1801	0					359	0	5	2
1	102	4064	1801					359	0	5	2
1	103	7159	4064					359	0	8	2
1	104	5791	0	0.028	0	153756	0	2175	359	6	2
1	105	8040	5791	0.028	0	153756	0	2175	359	7	2
1	106			0.028	0	249918	153756	2175	359	7	2
1	107			0.028	0	591241	249918	2175	359	7	2
1	108			0.039	0.028			2175	359	11	2
2	201					236701	0			22	2
2	202					1027526	236701			22	2
3	301	4135	0			45709	0			3	2
3	302	7465	4135			45709	0			8	2
3	303	9898	7465			45709	0			10	2
3	304					102554	45709			11	2
3	305	4444	0			339758	102554			13	2
3	306	6003	4444			339758	102554			11	2
3	307	11618	6003			339758	102554			10	2
4	401					66171	0	4345	0	28	2
4	402	6045	0			565025	66171	4345	0	27	2
4	403	11623	6045			565025	66171	4345	0	25	2
4	404							17641	4345	30	2
5	501	3620	0	0.048	0	125590	0			5	2
5	502	4530	3620	0.048	0	125590	0			8	2
5	03	4951	4530	0.048	0	125590	0			6	2
5	504	5016	4951	0.048	0	125590	0			3	2
5	505	5277	5016	0.048	0	125590	0			5	2
5	506	5746	5277	0.048	0	125590	0			6	2
5	507	6399	5746	0.048	0	125590	0			5	2

5	508	12826	6399	0.048	0	125590	0			8	2
5	509	5641	0	0.048	0	210430	125590			6	2
5	510	8348	5641	0.048	0	210430	125590			7	2
5	511	13892	8348	0.048	0	210430	125590			10	2
5	512			0.048	0	358684	210430			8	2
5	513			0.048	0	877546	358684			13	2
5	514			0.085	0.048					17	2
6	601					49854	0			35	2
6	602	6353	0			162415	49854			34	2
6	603	14415	6353			162415	49854			35	2
6	604					250190	162415			33	2
6	605	5693	0			1156242	250190			35	2
6	606	16271	5693			1156242	250190			35	2
7	700									1	1
7	701	6477	0	0.027	0	104522	0			7	2
7	702	6921	6477	0.027	0	104522	0			4	2
7	703	7861	6921	0.027	0	104522	0			5	2
7	704	5137	0	0.027	0	249358	104522			3	2
7	705	8070	5137	0.027	0	249358	104522			10	2
7	706			0.048	0.027	92716	0			9	2
7	707			0.048	0.027	186409	92716			7	2
8	801							3938	0	30	2
8	802							18292	3938	41	2

*: STRATID: Secondary PSU ID.

***: VMT_RATE_IMP = imputed vehicle miles traveled / (PSU MOS×1,000,000).

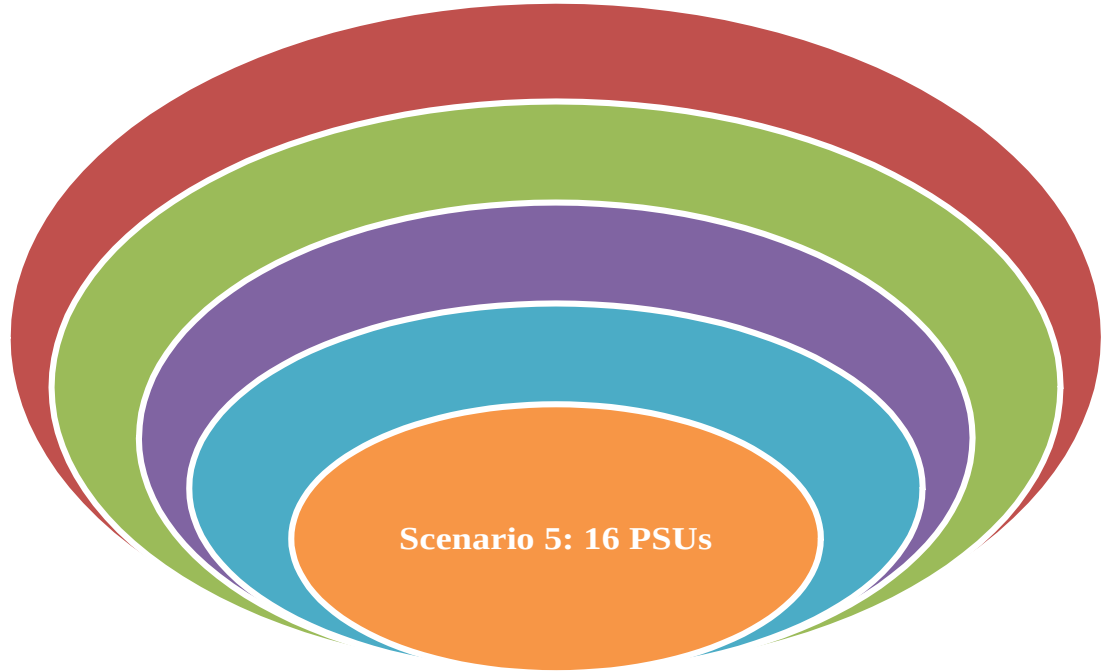
TOT_CRASH_RATE = (imputed 2008 injury crashes + imputed 2008 PDO crashes + 2007-2011 average fatal crashes) / (PSU MOS×1,000,000).

TRK_MI_RATE = Total truck miles / (PSU MOS×1,000,000).

ROAD_TYPE_RATE = (primary road miles + secondary road miles) / (PSU MOS×1,000,000).

A major challenge of the CRSS sample design is the uncertainty of the future operational budget. Due to unknown future funding levels and the need for a stable PSU sample, NHTSA implemented a scalable PSU sample, which allows for the PSU sample size to be decreased or increased with minimum impact to the existing PSU sample and for the selection probabilities to be tracked. To this end, a multi-phase sampling method was used to select the CRSS PSU sample by selecting a sequence of nested PSU samples. In this method, a PSU sample larger than what is actually needed is selected during the first phase of the PSU sample. From the first phase of the PSU sample, a smaller subset of the PSU sample is selected as the second phase of the PSU sample. From the second phase of the PSU sample, another smaller third phase of the PSU sample is selected. This process is continued until the PSU sample size reaches unacceptable levels. In this way, a sequence of nested PSU samples is obtained. Each of these PSU samples is a probability sample and can be used for data collection (see Figure 2). According to the prevailing budget level, a sample with the appropriate sample size is picked from the nested sequence. This allows us to easily track the selection probabilities and minimizes changes to the existing PSU sample.

Figure 2. Nested PSU Samples for CRSS



For the CRSS, five PSU samples were selected under the five scenarios. Table 4 summarizes the number of PSU strata and sampled PSUs for the CRSS PSU sample scenarios.

Table 4. CRSS PSU Sample Scenarios: Number of Strata and Sample Size

Scenario	Number of PSU Strata	Number of Sampled Non-certainty PSUs	Number of Sampled Certainty PSUs	Total Number of Sampled PSUs
1	50	97	4	101
2	37	74	1	75
3	25	50	1	51
4	12	24	0	24
5	8	16	0	16

For scenario 1, with a sample size of 100 and without stratification, one PSU was identified as a certainty PSU by the condition:

$$\pi_i = \frac{100 * MOS_i}{\sum_{i=1}^N MOS_i} \geq 1$$

Let N be the total number of PSUs in the PSU frame and i be an index for a PSU. The certainty PSU was selected with certainty⁴ and set aside. Then two PSUs were selected using proportional to size (PPS) sampling from each of the 50 scenario-1 strata. With a sample size of two for each PSU stratum, three PSUs were identified as certainty PSUs from three of the 50 scenario-1 strata by the condition:

⁴ In the probability proportional to size (PPS) sampling, a certainty PSU is identified when the selection probability is equal to or greater than one. If a PSU is identified as certainty, it must be in the sample and its selection probability is set to one. A non-certainty PSU is selected with its selection probability that is greater than 0 and less than 1. If a PSU has a selection probability closer to one, it has more chance to be in the sample. On the other hand, if a PSU has a selection probability closer to zero, it has less chance to be in the sample. For more details, please see Pages 13-28 in the published Technical Report <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812706>

$$\pi_{hi} = \frac{2 * MOS_{hi}}{\sum_{i=1}^{N_h} MOS_{hi}} \geq 1$$

Let N_h be the total number of PSUs in stratum h . The certainty PSUs were selected with certainty and set aside. The corresponding stratum PSU sample size was reduced by one. Then a PPS sample of non-certainty PSUs was selected using the revised PSU stratum sample size.

Scenario-1 sample has 101 PSUs. For a non-certainty PSU, the selection probability is:

$$\pi_{hi} = \frac{n_{hi} * MOS_{hi}}{\sum_{i=1}^{N_h} MOS_{hi}}$$

Let n_h be the non-certainty PSU sample size for PSU stratum h .

For scenario-2, with a sample size of 74 and without stratification, one PSU was identified as a certainty PSU and was set aside. Then 13 of the scenario-1 strata were collapsed with other strata to form the 37 scenario-2 PSU strata. The collapsing of strata follows the following rules:

- Only the secondary strata in the same primary stratum can be collapsed;
- Only the contiguous secondary strata can be collapsed;
- The resulting strata has a similar stratum total MOS within each primary stratum.

In each of the scenario-2 stratum, the sampled scenario-1 PSUs were treated as the sampling frame. Each PSU was assigned a new MOS equal to its scenario-1 stratum total MOS. Then two PSUs were selected from each scenario-2 stratum using PPS sampling based on the new MOS. In this way, the resulting selection probability of the scenario-2 PSU is still PPS selection probability. Other scenario samples were selected in a similar way.

The current CRSS PSU sample size is 61 (between scenarios 2 and 3) with 60 responding PSUs and one non-responding PSU.

Second Stage (PJ Sampling)

The secondary sampling units (SSU) of CRSS are police jurisdictions. Within each PSU, PARs are stratified by the police jurisdictions (PJ) where PARs are available and PJs become the second stage sampling units. A composite MOS is assigned to each PJ in the selected PSUs. Similar to the PSU MOS definition, it is sensible to assign larger selection probabilities to PJs with more high interest crashes as defined by the oversampled strata in Table 2. For each PJ in the selected PSUs, crash counts from the 9 PAR strata in Table 2 (Stratum 2-10) were estimated from the information collected from the PJs in the selected PSUs. For PJ j in the PJ frame within the sampled PSU i , the composite SSU MOS is defined as the following:

$$MOS_{j \vee i} = \sum_{k=2}^{10} \frac{n_{++s}}{n} \frac{N_{ijs}}{N_{++s}}$$

where

- s = the PAR stratum defined in Table 2.
- n = the desired total sample size of crashes
- n_{++s} = the desired sample size of crashes in the PAR stratum s
- N_{++s} = the estimated population count of crashes in PAR stratum s

N_{ijk} = the estimated population count of crashes in PAR stratum s , PJ j and PSU i

PJs are then stratified into two PJ strata by their MOS (large MOS stratum [largest 50%] and small MOS stratum [the rest]) in addition to certainty PJs. A PJ sample is then selected from each PJ stratum using Pareto sampling. The Pareto sampling method produces an approximate PPS sample, handles the frame changes and minimizes the changes to the existing sample at the same time. Pareto sampling method was applied to the PJ sample selection for each of the non-certainty PJ strata (large MOS and small MOS stratum) within the sampled PSU i , as the following:

Step 1: Generate a permanent uniform random number $r_{ihj} \sim U(0,1)$ for each PJ j in the PJ stratum h of PSU i .

Step 2: Identify certainty PJs by the condition:

$$\frac{m_{ih} * MOS_{ihj}}{\sum_{j=1}^{M_{ih}} MOS_{ihj}} \geq 1$$

Let m_{ih} be the PJ sample size and M_{ih} be the PJ frame size for PJ stratum h within PSU i . MOS_{ihj} is the PJ MOS.

Step 3: The identified certainty PJs are set aside. This process is repeated for the remaining PJs based on the reduced PJ sample size until there are no more certainty PJs. Let the total number of certainty PJs be m_{ch} . For the remaining $M_{ih} - m_{ch}$ non-certainty PJs in the frame, calculate the PPS inclusion probability with the non-certainty PJ sample size $(m_{ih} - m_{ch})$:

$$p_{ihj} = \frac{(m_{ih} - m_{ch}) MOS_{ihj}}{\sum_{j=1}^{M_{ih} - m_{ch}} MOS_{ihj}}$$

Step 4: Calculate the transformed random numbers and sort the transformed random numbers from the smallest to the largest as following:

\hat{r}_{ihj}

Step 5: The m_{ch} certainty PJs plus the first $m_{ih} - m_{ch}$ non-certainty PJs from the above list are the PJ sample for PJ stratum h within PSU i .

Pareto sampling is approximately PPS, and the PJ selection probability is:

$$\pi_{j \vee ih} \approx p_{ihj}$$

The 2017 CRSS PJ sample will be used for this program. The 2017 CRSS PJ sample size is 398.

Third Stage (PAR Sampling)

The tertiary sampling units (TSU) of CRSS are PARs. The CRSS PAR sample is selected by stratified systematic sampling. For each selected SSU (PJ), PARs are periodically obtained by either a technician's visit to the PJ or electronic transmission. All the PARs are listed in the order they become available and are stratified by the PAR strata identified in Table 2. Through this listing process, the PAR sampling frame in each selected PJ is prepared for PAR sample selection.

For a large PJ with too many PARs to be listed, PARs are sub-listed by systematic sampling. For example, only PARs with a PAR number ending in 0 through 4 may be listed if the sub-listing factor is 2 (i.e., 5 PARs among 10

PARs are listed). Or only PARs with a PAR number ending in 0 or 1 are listed if the sub-listing factor is 5 (i.e., 2 PARs among 10 PARs are listed). If L PARs among 10 PARs are sub-listed in PJ j in PSU i , the sub-listing probability for all sub-listed PARs are:

$$\pi_{l \vee ij} = \frac{L}{10}$$

After PARs are listed, a PAR sample is selected by systematic sampling from the listed (or sub-listed) PARs by PAR stratum within each selected PJ. PAR k selection probability within a PAR stratum is:

$$\pi_{k \vee ij} = \frac{n_{ijl}}{N_{ijl}}$$

Let n_{ijl} be the number of selected PARs and N_{ijl} be the number of listed PARs from each PAR stratum in PJ j of PSU i .

The overall selection probability is:

$$\pi_{ijk} = \pi_i * \pi_{j \vee i} * \pi_{l \vee ij} * \pi_{k \vee ij}$$

The design weight is the inverse of π_{ijk} .

Sample Allocation

For a three-stage sample design as this program, the PSU, SSU and TSU sample sizes can be estimated using optimization by minimizing the variance subject to cost assuming a three-stage simple random sampling without replacement.

The optimization model consists of the objective function, cost constraint, and variance constrains as the following:

$$\text{Minimize: } \sum_{g=1}^G V(\bar{\bar{y}}_g) = \sum_{g=1}^G \left\{ \frac{S_{1,g}^2}{n} \left(1 - \frac{n}{N}\right) + \frac{S_{2,g}^2}{nm} \left(1 - \frac{m}{M}\right) + \frac{S_{3,g}^2}{nmk} \left(1 - \frac{k}{K}\right) \right\}$$

$$\text{Subject to: } C = C_0 + n C_1 + nm C_2 + nmk C_3,$$

$$V_{CRSS}(\bar{\bar{y}}_g) = \frac{S_{1,g}^2}{n} \left(1 - \frac{n}{N}\right) + \frac{S_{2,g}^2}{nm} \left(1 - \frac{m}{M}\right) + \frac{S_{3,g}^2}{nmk} \left(1 - \frac{k}{K}\right) \leq V_{GES}(\bar{\bar{y}}_g),$$

for $g=1, \dots, G$.

where

g : Subscript of the identified key estimate, $g=1, \dots, G$.

$\bar{\bar{y}}_g$: Identified key proportion estimate.

n, m, k : Optimal sample sizes of PSUs, SSUs per PSU, and TSUs per SSU to be determined.

N : Population size of PSUs

M : Average population size of SSUs.

K : Average population size of TSUs.

V_{CRSS} : Variance of the identified key estimate $\bar{\bar{y}}_g$.

$S_{1,g}^2, S_{2,g}^2, S_{3,g}^2$: Variance component at PSU-, SSU-, and TSU-level.

C, C_0, C_1, C_2, C_3 : Total, fixed, PSU-, SSU-, and TSU-level cost coefficients.

$V_{GES}(\bar{\bar{y}}_g)$: Variance of the identified key estimate $\bar{\bar{y}}_g$ in General Estimates System (GES)⁵.

⁵ <https://www.nhtsa.gov/national-automotive-sampling-system/nass-general-estimates-system>

Notice because this program utilizes the CRSS infrastructure and collecting data through mails and emails, the cost for establishing PSU and SSU is negligible. Therefore, it becomes obvious the optimum sample size allocation is to first maximize the PSU sample size then maximize the SSU size. For this reason, NHTSA decided to use all 60 responding CRSS PSUs and all 398 CRSS SSUs for this program.

1. Describe collection of information procedures.

A CISS and SCI crash investigation starts with a Police Accident Report (PAR), the form law enforcement agencies use to document a motor vehicle crash. NHTSA collects PARs from cooperating police jurisdictions and custodial agencies in each State. In addition to data derived from the PAR, NHTSA obtains additional information to further the understanding of a crash, its causal factors, or outcomes. For CISS, once a crash has been selected for investigation, crash technicians or investigators locate, visit, measure, and photograph the crash scene; locate, inspect, and photograph vehicles; conduct a telephone or personal interview with the involved individuals or surrogate (another person who can provide occupant or crash information,, parent s for minor, parent or spouse for deceased); and obtain and record crash injury information received from various medical data sources. These data are used to describe and analyze circumstances, mechanisms, and consequences of serious motor vehicle crashes in the United States. During each activity the Crash Technicians record information on the crash, vehicle, and occupant forms as appropriate.

For the investigation-based Special Studies the data is targeted to a specific issue (i.e. child occupant protection, crash causation factors, etc.) as opposed to an entire investigation. Investigators will contact the investigating agency to obtain case materials (e.g. PAR, photos or other pertinent information about the crashes). The crash investigators will obtain and review the case material from local authorities to collect and code detailed crash reconstruction data on the cases. The source of the data will be the case materials the local authorities gathered while investigating these fatal cases. The investigators will not travel to the scene or inspect the vehicles. They will rely on the case material obtained from local authorities. While each issue-based special study has specific requirements (i.e., types of crashes and/or data collected) , the gathering of crash reports is expected to take similar times regardless of the targeted specific issue.

The data collected from the CISS, SCI and special studies are entered into NHTSA's Crash Data Acquisition Network (CDAN). The CDAN is an integrated, web-based information technology system that provides a single, central IT platform that maintains the data NHTSA collects from its CISS, SCI and Special Studies programs. CDAN has an adjudicated Privacy Impact Assessment (PIA) that ensure compliance with laws and regulation governing privacy of personal information.⁶

2. Describe methods to maximize response rates and to deal with issues of non-response.

We use the existing CISS as an example to explain the response rates at each sampling stage and the methods we used to increase the response rates.

CISS has a three stage sample design. The first stage sampling units, which are primary sampling units (PSUs), are counties or groups of counties. A PSU becomes a non-responding PSU only if all selected police jurisdictions (PJs) within the PSU are non-responding PJs. Since PJ samples are selected using Pareto sampling method, every non-selected PJ in the PJ frame can be selected as a replacement. Therefore, a PSU becomes non-responding PSU only if all PJs in the frame are non-responding PJs. In 2019 CISS, all 32 sampled PSUs were responding PSUs.

The second stage sampling units are PJs. A sampled PJ becomes non-responding PJ if it refuses to cooperate. To improve PJ cooperation rate, NHTSA visits each selected PJ and meet with local law enforcement officers to gain cooperation. In 2019 CISS, among the 233 sampled PJs in 32 PSUs, only 5 PJs were non-responding.

⁶ <https://www.transportation.gov/individuals/privacy/crash-data-acquisition-network-cdan-pia>

At the third stage, all PARs in the selected PJs are listed. Then a PAR sample is selected using Pareto sampling method. For a sampled case, if the case vehicle which defines the case's PAR domain is not available for data collection because it is repaired or removed etc., then the case becomes a non-responding case. For each non-responding case, a replacement case is selected. All responding cases and replacement cases are investigated. In 2019, among the 3,090 sampled PARs, 2,781 were investigated, resulting in a 90% PAR response rate.

The item response rate of CISS varies. For example, the item response rates of the crash file variables range from 72% to 100%; the item response rates of the general vehicle file variables range from 54% to 100%; the item response rates of the occupant file variables range from 13% to 100%. The Crash Investigation Sampling System (CISS) quality control system is designed to produce the most accurate, reliable, and complete database possible within the limits of available resources. All data will be automated and edited by a complex algorithm which checks for inconsistencies and questionable items. A subsample of all sampled crashes will be given a thorough review by an experienced researcher at a Data Quality Control Zone Center. Zone Center personnel will visit each PSU regularly to observe CISS crash technicians and investigators perform their investigation activities and to discuss systematic problems revealed in edit and Zone Center reviews of the team's cases.

Since the interview is a vital part of data collection, CISS investigators make special efforts to complete an interview when at all possible. Occupants will be contacted by telephone. CISS investigators will call at varying hours (often in evenings or on weekends) until they have located the person sought. When the person is unavailable, other passengers or witnesses are contacted. If the person sought cannot be located by telephone, investigators use personal visits or mail questionnaires. Each CISS investigator will be given special training in interviewing, which increases the likelihood persons will cooperate once they have been located and contacted. As a result of these procedures, which were also used in our legacy program (NASS), it is anticipated that CISS investigators will complete more than three-quarters of all occupant interviews. Accordingly, the interview item missing rate is about 25%.

As a final check on CISS data, approximately 5% of those interviewed will be re-contacted by Zone Center personnel to establish that they had in fact been interviewed and to verify some of their responses. This type of interview takes approximately 5 minutes.

4. Describe any tests of procedures or methods to be undertaken.

There is no tests procedure to be undertaken at this time.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Ms. Chou-Lin Chen, National Center for Statistics and Analysis, NHTSA, 202-366-1048 is responsible for CISS survey design and special studies.

NHTSA and Westat (contract DTNH22-12-F-00389) jointly developed the new Crash Investigation Sampling System (CISS) survey design. The CISS data collector and quality control contractors are KLD Associates, Inc. (contract 693JJ921C000011) and Calspan Corporation (contract693JJ921C000007), respectively.