Study and Sample Design Plan Review for Federal Compliance Programs: NHTSA's State Seat Belt Use Study

Martha J. Rozsi[1], James L. Green[1]
[1]Westat, 1600 Research Blvd, Rockville, MD 20850

Abstract

Traffic crashes are a leading cause of death in the United States, but seat belt use has been shown to save lives and prevent injuries. The National Highway Traffic Safety Administration (NHTSA) requires every state and territory to conduct a survey to monitor seat belt use. NHTSA issued the new Uniform Criteria for State Observational Surveys of Seat Belt Use on April 1, 2011 which the states are required to follow for their study design. These criteria require the states to update survey methods and select observation sites that are representative of the state's road segments in a cost-effective manner. Westat reviewed state study design plans for compliance with the operational, GIS, and statistical requirements of the criteria. We present the tools and process used for the reviews, as well as some observations and future recommendations on the statistical requirements.

Key Words: Federal Compliance, Traffic Safety, Protocol Review, Geographical Sampling, Traffic Observational Studies

1. Introduction

The National Highway Traffic Safety Administration (NHTSA), a part of the Department of Transportation, requires an annual state seat belt use study from all states and territories to estimate the proportion of drivers and front seat passengers who use seat belts. While each state's study must meet certain requirements defined by NHTSA in the Uniform Criteria for State Observational Surveys of Seat Belt Use each individual state has considerable flexibility in conducting its seat belt study.

Typically, to conduct a seat belt use study, road segments are sampled. A data collector goes to the sampled segments, finds a place on the road where traffic slows, such as on a curve or an intersection, and observes vehicles as they drive by. During the observation period, the data collector looks into the windows of the passing vehicles and notes if there is a front seat occupant, and if the driver and passengers are wearing seat belts. However, if a car passes by and the data collector was unable to tell if seat belts were being worn, then the vehicle is considered nonresponding.

Many decisions need to be made for designing such a study: what types of road segments should be sampled and how many of each, how sampling should be conducted, what time of day should observations be made, and what steps should be taken to minimize nonresponse. Decisions also need to be made about how to analyze the data from the

study: what estimator(s) to use, how to account for nonresponse, and what software to use.

NHTSA contracted with Westat to perform the geographic information systems (GIS), operational and statistical review of the submitted plans per the compliance requirements in the Uniform Criteria. This paper reviews the requirements given in the Uniform Criteria for seat belt study. It then examines the review tools that Westat developed and utilized to check the state and territory submitted plans for compliance with the Uniform Criteria the entire review process for checking compliance, overall observations of statistical requirements in compliant plans, and further issues noticed in plans that make compliance difficult.

2. Study Requirements and the Uniform Criteria

2.1 Rationale and Protocol for the Announcement

The Uniform Criteria were announced in the Federal Register on April 1st, 2011. It was published in the Federal Register Vol. 76 No. 63 April 1, 2011, Rules and Regulations, pp. 18042—18059. It provided the rationale and protocol for the seat belt use study that is required of all 50 states, the District of Columbia, and all United States territories.

The announcement allowed for a public comment period and a question and answer period to discuss changes from the former rule. One of the bigger changes in the Uniform Criteria was switching from a population-based requirement for county inclusion (where counties or county equivalents included in the sample needed to account for 85% of the population of the state/territory) to a fatality-based requirement for county inclusion. Now the states and territories needed to use at least three years of data from the Fatality Analysis Reporting System (FARS) to account for counties making up at least 85% of the fatalities in the state/territory.

2.2 Study Requirements

According to the Uniform Criteria, states and territories needed to be compliant on 19 different requirements. These requirements included one geographic information systems (GIS) requirement, five operational requirements, and thirteen statistical requirements. A plan was not compliant unless all 19 requirements were met. The review process focused on compliance with the Uniform Criteria, rather than the efficiency or relative efficiency of proposed study design plans.

2.2.1 GIS Requirement

The GIS requirement was that the source for the sample frame of road segments needed to be complete. Each state/territory was offered the use of a Topologically Integrated Geographic Encoding and Referencing (TIGER) dataset from the Census Bureau that gave a complete list of road segments, the segment lengths, latitude and longitude of the segment, the segment's road type, and county in which the segment resides. If a state/territory used the TIGER frame, the state/territory was compliant. However, the TIGER dataset has limited vehicle or mileage usage information, and some states/territories chose to use their own road segment databases that included additional information such as Vehicle Miles Traveled (VMT), Daily Vehicle Miles Traveled (DVMT), and Annual Average Daily Traffic (AADT). If a state/territory wanted to use a sampling frame that was not the TIGER dataset, then the state/territory needed to provide proof that it was complete and NHTSA needed to approve the database.

2.2.2 Operational Requirements

There were five operational requirements in which plans needed to be compliant. These requirements included assigning sampled sites to time periods, properly defining an observation site, procedures for rescheduling and substituting sites, proper data collection procedures, and the number of data collectors and quality control monitors. These five requirements are listed below.

- Sampled sites needed to be assigned to time periods. The eligible time periods needed to include all daylight hours between 7 a.m. and 6 p.m. for all days of the week. Observation sites could be grouped based on close geographic proximity to reduce data collection burden, but the first time period assignment needed to be randomly selected.

- The procedures for rescheduling and substituting sites needed to be complaint with the Uniform Criteria. This meant that if a site was temporarily unavailable, the site would be rescheduled for another time that was the same day of the week and the same time period. If a site was permanently unavailable, then the site could be replaced by a new site that is in the same county and of the same road type. This substituted site must be visited at another time that is the same day of the week and the same time period. Although this is an operational requirement, statisticians paid close attention to how the substituted sites were sampled as under some implementations this changed the probabilities of selection for the original sites.

- An observation site needed to be properly defined. An observation site is the physical location where the survey data is to be collected. This does not necessarily need to be a road segment, but all road segments need to be eligible for sampling. The set of potential observational sites therefore cannot be limited to roads having a stop sign or stop light, or to state-maintained roads.

- Data collection procedures needed to be compliant with the Uniform Criteria. Data collection needed to be completed within the calendar year for which the statewide seat belt use rate will be reported. If an intersection is used as the observation site, the data must be collected from the sampled road segment and not the intersecting segment(s). If the observation site is located on a road with two-way traffic, either both directions must be observed or only one direction chosen at random is to be observed. Data must be collected for all passenger motor vehicles. Data needs to be collected for all drivers and right front seat passengers, including those in booster seats, but excluding those in child seats. The data collected must include seat belt status of the driver, the presence of a right front passenger, and seat belt status of right front passenger if applicable. Finally, the data collectors are not allowed to be in law enforcement uniforms, have police vehicles or persons in law enforcement uniforms at the site, nor have signage stating a seat belt use survey is in progress.

- The number of data collectors and quality control monitors needed to be compliant with the Uniform Criteria based on the number of observation sites sampled. Quality control monitors needed to conduct random, unannounced visits to at least 5% of observation sites. The quality control monitor could not also be the data collector for a site.

2.2.3 Statistical Requirements

Plans needed to be compliant on 19 statistical requirements. These requirements addressed many statistical survey design features such as stratification, sampling methods, sample sizes, estimators and variance estimators, and standard errors. These requirements also included some items that are not typically considered statistical, but affect the statistical requirements, such as allowable exclusions to the sampling frame and the qualified state statistician.

The remainder of this paper will focus on these 19 statistical requirements, ignoring the GIS and operational requirements. These 19 statistical requirements will be found in section 5 with the overall observations of the plans.

3. Review Process and Tools

3.1 Review Process

Typically, each submitted plan would go through the same review process for determining compliance. There were three main reviewers (one GIS reviewer, one operational reviewer, and one statistical reviewer), as well as a statistical senior reviewer, and a senior operational reviewer. Once a plan was submitted, there would be a receipt of the submission and a notification was sent to each of the three main reviewers. Each reviewer had a separate checklist that would be completed for each plan review. Once the statistical review was completed by the main reviewer, the statistical senior reviewer would do a peer review with the statistical reviewer. There were three statistical reviewers, two senior statistical reviewers, one GIS reviewer, and four operational/senior operational reviewers.

The statistical review team held weekly status meetings. In these meetings, the previous week's activity was reviewed and discussed, incoming plans were assigned based on staff availability, and any complicated or interesting issues were shared with the entire team. Once a state/territory was assigned to reviewers, those same people would review all plans submitted for that state/territory.

After the reviewers were complete and any complicated issues were discussed, a summary and report were sent to the states. Plan revisions were made, and the process was repeated until the state/territory submitted a compliant plan.

Occasionally, email correspondence and conference calls with states and territories were required to help with any misunderstandings and complications. These emails and conference calls addressed statistical requirements such as correctly calculating the probabilities of selection for sites, correct estimators for the sample designs, and use of a two-phase approach.

3.2 Review Tools

Several key review tools were used to assist with the review process for compliance, especially with resubmissions from states/territories. These review tools included checklists, example plans and references, use of a SharePoint site, and tracked changes in submitted plans.

The first review tool was a checklist of all requirements, as well as checklists for each main reviewer. These checklists mapped each requirement to the appropriate Uniform

Criteria reference. The checklists also provided sample design features and examples. All the checklists mentioned the requirement type (GIS, Operational, or Statistical) for each requirement.

Example plans and reference documents were key in helping plans become compliant. The Model Plan was created by the review team, and gave an example of a compliant plan for a fictitious state. Although states and territories did not need to follow the example laid out by the Model Plan, it provided a plan outline that many state plans followed, and also provided reference material to show an example of what a compliant requirement looked like. For example, when plans did not present a nonresponse adjustment section, the reviews sent to the states/territories would mention that a nonresponse adjustment section needed to be provided and indicated the section in the Model Plan to consult for reference. Additional reference documents were provided on the SharePoint site that gave proofs for various estimators, including a hybrid estimator that involved knowing extra information (such as VMT) for some segment types, while not knowing extra information for other segment types.

A SharePoint site was accessible by Westat, NHTSA, and the states/territories. On the SharePoint site, states/territories would submit plans, which would automatically send notification that a plan had been submitted. Then reviewers would post the plan reviews on the SharePoint site (this part was not accessible by states), giving notification of the review being complete. The final review for a plan was edited and submitted on the SharePoint site for all parties to observe. The SharePoint site also provided the TIGER databases for all states/territories to download, as well as any references and example plans.

The use of tracked changes made comparing previous plan submissions to new submissions easier. A revised plan might often be noncompliant because of one requirement, and therefore should be compliant in the next submission. Instead of needing to reread the entire plan for determining compliance, the compare documents feature in Microsoft Word was used to pinpoint the changed text. Document comparison software was also used for plans submitted in PDF. This type of software expedited the review of resubmissions.

4. Statistical Requirements and Observations

The statistical requirements and observations involved design features/design variations, stratification, sampling methods, probabilities of selection, sample size calculations, response rate requirements, estimators, variance estimators, imputation, weight calculations, nonresponse adjustments, and standard error requirements. The statistical requirements also included mentioning exclusions taken to the sampling frame, as well as mentioning the state statistician and his/her qualifications. This section will define each of the 19 statistical requirements, and ranks of frequency of use for a particular feature by compliant plans, where 1 is the most common, 2 is the second most common, etc. Ranks are used instead of frequencies to protect confidentiality.

4.1 Requirement 1 – Sampling Units and Measure of Size

Requirement 1 involved the sampling units and their measure of size (MOS) if doing a probability proportional to size (PPS) sample. The Uniform Criteria stated that plans

must clearly state what the sampling unit is. All compliant plans used road segments as the final sampling unit.

The Uniform Criteria only stated a probability sample must be used, and therefore did not require a PPS sample. However, when a PPS sample was used, the measure of size needed to be stated. Some plans sampled counties and then road segments, while other plans only sampled road segments within the eligible list of counties. Table 1 gives the ranks of the MOS's used for road segments, while Table 2 gives the ranks of the MOS's used for counties.

Table 1: Measure of Size for Road Segments for Compliant Plans

Measure of Size	*Rank*
Length	1
VMT	2
AADT	5
VMT adjusted by length	5
MSA Size	5

The majority of states that used a PPS sample used length as the MOS, as shown in Table 1. This was partially because it was available on the Census TIGER databases provided to states/territories. It was also the MOS used in the Model Plan, but it was not evident if plans used it as the MOS only because it was used in the Model Plan.

Table 2: Measure of Size for Counties for Compliant Plans

Measure of Size	*Rank*
VMT	1
Population	3
Paved Miles	3
Vehicle Registration	6
Fatalities from FARS	6
Drivers Licenses	6

4.2 Requirement 3 – Exclusions to Sampling Frame

Requirement 3 involved exclusions that could be taken to the sampling frames. The counties that were eligible for sampling in the state/territory needed to account for 85% of the fatalities by using at least 3 years' worth of FARS data. Table 3 shows if plans only kept counties making up the 85% criteria, or if plans kept all counties.

Table 3: Kept Only Counties Meeting 85% Fatalities or Kept All Counties in Compliant Plans

85% FARS Criteria	*Rank*
Kept only counties that made 85% fatalities	1
Kept all counties	2

The majority of plans only kept what they needed to in order to be compliant. This helped with operational costs. Some plans that did keep all counties did so because they were combining the seat belt use survey with other traffic surveys, and therefore needed a larger sample.

Another eligible exclusion was in regards to road types. Certain road types could be excluded and the plan would still be compliant. The road types that could be excluded were rural local roads not in Metropolitan Statistical Areas (MSAs), non-public roads, unnamed roads, unpaved roads, vehicular trails, access ramps, cul-de-sacs, traffic circles, or service drives. Table 4 shows if plans excluded some or all road types.

Table 4: Exclusion of Possible Road Segment Types in Compliant Plans

Road Segment Exclusions	*Rank*
Excluded all eligible road types	1
Excluded some of the eligible road types, but not all	2
Did not exclude any roads	3

4.3 Requirement 4 – Stratification and Stage of Selection

Requirement 4 referred to the stratification and stage of selection in plans. Road segments were always used as a stage of selection in compliant plans, though some plans also used counties and census tracts as additional stages of selection, as shown in Table 5. Road segments were typically stratified, and when they were stratified, it was always done by road type (Table 6). Allocation of sampling road segments within counties, regions, or Census tracts was typically done based on VMT (Table 7). The majority of plans sampled more than two segments per stratum, but there were plans that sampled fewer than two segments per stratum, which had implications in variance calculations and nonresponse adjustment (Table 8).

Table 5: Stratification and Stages of Selection for Compliant Plans

Unit	*Stratification Rank*	*Stage of Selection Rank*
Road Segments	1	1
Regions	2	
Counties	3	2
Census Tracts	4	3

Table 6: Stratification of Road Segments in Compliant Plans

Stratification	*Rank*
Type of Road	1
None	2

Table 7: Allocation of Sampling Road Segments in Compliant Plans

Measure for Allocation	*Rank*
VMT	1
Evenly	2
Length	3

Table 8: Number of Road Segments per Stratum for Compliant Plans

Number of Units per Stratum	*Rank*
More than 2	1
Less than or equal to 2	2

4.4 Requirement 5 – Method Used for Sampling

The Uniform Criteria stated that a probability sample must be used. Requirement 5 was the type of sampling done to ensure a probability sample was being used. The majority of compliant plans used a PPS sample, but some other plans used a simple random sample (SRS). These were the only two types of probability samples used in compliant plans.

4.5 Requirement 6 – List of Observational Sites, Reserve Sample, and Probabilities of Selection

A compliant plan needed to include the list of sampled sites and their probabilities of selection, properly calculated based on the sampling method. A compliant plan also needed to include a reserve sample if the original site was permanently unavailable. Although this was an operational requirement, the statistical reviewer took interest in this requirement because it could affect the sample of the original sites and the probabilities of selection. Numerous plans used a two phase approach, where original sites and reserve sites were sampled at the same time, affecting the probabilities of selection. The Uniform Criteria did not state a requirement for the size of the reserve sample, but it did need to be large enough so that if a site was permanently unavailable, there was an eligible alternate site. The majority of plans took a small reserve sample of at most one alternate per sampled segment, but many plans followed the Model Plan and took 2 alternates per original segment. Table 9 shows what compliant plans used for the reserve sample.

Table 9: Reserve Sample Sizes for Compliant Plans

Size or Approach for Reserve Sample	*Rank*
Less than 2 Alternates per Segment	1
2 Alternates per Segment	2
Two phase approach	4
Continue Driving on Segment Until Available Site	4
More than 2 Alternates per Segment	5

4.6 Requirement 7 – Expected Sample Sizes

Plans were required to state an expected sample size. The sample sizes could refer to either the number of vehicles expected to sample, or the number of observations (drivers and front seat passengers). The majority of the expected sample sizes provided was based on previous seat belt use studies. Some compliant plans based sample sizes on optimizing cost, known AADT, or were estimated. Table 10 shows how the expected sample sizes were determined.

Table 10: Expected Sample Sizes for Compliant Plans

How Sample Size was Determined	*Rank*
Based on previous studies	1
Based on Cost Optimization and Requires Standard Error	2
Estimated	3
Based on VMT or AADT	4

The sample sizes ranged from 1,250 to 175,000, with an average sample size around 40,000. These sample sizes differed greatly because the sample sizes could refer to either vehicles or observations. Many of the larger expected sample sizes were also due to the fact that multiple traffic studies were going to be done at once, or region, county, or road type estimates were also desired, and therefore a larger sample was required.

4.7 Requirement 9 – State Statistician and Qualifications

Requirement 9 was that the plan must name the project statistician and provide credentials, affiliation, or resume. The Uniform Criteria specified that the survey statistician must be an individual knowledgeable of the design of probability-based multi-stage samples, statistical estimators for such designs, and the variance estimation of such estimators. Many states had their own statistician for this survey, but some states contracted out for a statistician.

4.8 Requirement 12 – Response Rate

The seat belt use survey requires a 90% response rate, meaning that the data collector needed to be able to view and determine 90% of the seat belt use for drivers and front seat passengers. If the overall response rate was less than 90%, additional observations must be taken to ensure an overall response rate of 90%. The majority of plans followed the wording in the Model Plan and stated that "if any site exceeded a rate of 10% unknown, potentially leading to an overall nonresponse rate of more than 10%, then an additional observation period would be collected from the site." The other compliant plans stated that additional data would be collected only if the overall unknowns exceed 10%.

4.9 Requirement 15 – Estimator

Compliant plans needed to include an overall state level seat belt use rate estimator. The majority of plans provided a simple ratio estimator like the one in the Model Plan, as this estimator did not require knowledge of VMT or AADT. Some plans provided a poststratified ratio estimator, the majority of which used VMT, though others used length. Some plans used a hybrid estimator that was provided on the SharePoint site. The hybrid estimators took advantage of both VMT and length, as available by the multiple road segment databases used by a plan. The ranks for the types of estimators used are in Table 11.

Table 11: Estimators for Compliant Plans

Type of Estimator	*Rank*
Simple Ratio Estimator	1
Poststratified Ratio Estimator	2
Hybrid Estimator	3

Although only a state level estimator was required, many plans included additional estimators. These additional estimators were either at the region level, county level, or road segment type level. When multiple stages of estimators were presented, it was important to verify that each estimator was correct for the sample design, and also that each estimator correctly rolled up to the next level (if that was the intended approach). Table 12 gives the ranks for what additional estimators were common in compliant plans.

Table 12: When Multiple Estimators Were Present in Compliant Plans

Level of Estimator	*Rank*
State	1
County	2
Road Segment Type	3
Region	4

4.10 Requirement 16 – Variance

Compliant plans needed to include methods for variance estimation, but did not necessarily need to provide a variance estimation equation. If a plan mentioned a program (e.g. SUDAAN) and approach (e.g. jackknife) to be used for variance estimation, the plan was compliant on this requirement. Table 13 ranks the frequency with which statistical software packages were used for variance estimation.

Table 13: Variance Calculations for Compliant Plans

Program	*Rank*
SUDAAN	1
SAS	2
Microsoft Excel	3
SPSS	4
R	5
PC CARP	7
No mention of software, but compliant equation	7

4.11 Requirement 17 – Imputation

If a state/territory was going to use imputation on missing data, the imputation plans must be laid out to be compliant with the Uniform Criteria. If imputation was not going to be done, the plan must state so. No plans stated that they were going to use imputation.

4.12 Requirement 18 – Weighting and Nonresponse Adjustment

Requirement 18 was regarding weighting and nonresponse adjustment of the data. For a plan to be compliant, it must provide correct base weight calculations and nonresponse adjustment plans. A site was nonresponding if one or more vehicles drove by but no observations of seat belt use could be made. A site with no vehicles driving by during the observation period is not nonresponding.

Finely stratified designs led to plans having strata with two or fewer sampled segments (see Table 8). This could potentially cause an entire stratum to be nonresponding. If a plan had two or fewer segments in a stratum, the plan needed to mention a strategy for dealing with nonresponse adjustment for a nonresponding stratum. The majority of these plans stated that collapsing of strata will be done if a stratum is nonresponding, but some plans mentioned another option of revisiting nonresponding sites for an additional observation period to ensure no site, and therefore no stratum, is nonresponding.

4.13 Requirement 19 – Standard Error

The final statistical requirement was that the standard error could not exceed 2.5 percentage points. Compliant plans needed to state procedures to reduce the standard error if it exceeded 2.5 percentage points. All compliant plans stated that in the event the precision objective is not met, additional observations will be taken.

5. Further Issues

A large range of issues and complications in study design plans made determining compliance difficult. These issues included the sampling frames being used, the approach used for sampling segments, simulation being required to estimate/calculate site probabilities, the use of maps instead of road segment databases, combining multiple

traffic studies, finely stratified designs, tinted windows, and clarification between stratification and stage of selection.

5.1 Sampling Frames

Census TIGER databases of road segments were made available to all states and territories. However, these databases did not contain information like VMT or AADT. Also, sometimes these databases misclassified road type for some road segments. As a result, some states and territories chose to use their own road segment databases. Other states and territories would try to supplement their own databases with the Census TIGER databases provided, causing issues with merging various databases and using multiple frames. Some databases attempted to be used were not a complete sampling frame, but instead were samples of the complete sampling frame. When this was the case, the probabilities of the sampled segments on the database needed to be provided to adequately provide the observation sites' probabilities of selection. Final databases being used as the sampling frame required NHTSA approval.

5.2 Two Phase Approach

Some plans used a two phase approach where both original segments and reserve segments were sampled at the same time. This complicated calculating the probabilities of selection for observation sites, and had implications for calculating base weights. A two phase approach was also sometimes used for the sample of original segments because the sample frames included ineligible road types, and the sample of road segments needed to be large enough to have the necessary road segment sample size.

5.3 Simulation

Some states and territories proposed to use road segments sampled from the previous designs for the seat belt survey to make longitudinal assessments. It was difficult for these states to calculate the selection probabilities because additional road segments were built after the previous sample was drawn, and the criteria for county exclusion changed from population to fatalities. To estimate the selection probabilities for these road segments, some states proposed simulation. However, the simulation had to accurately depict the actual sampling methods used for both the previous design and the current design, which required documentation of the previous design. If precise documentation could not be produced for the previous design, the plan was noncompliant.

5.4 Maps

Some states and territories proposed to use maps in place of a road segment database. The maps of the state/territory would be overlaid with a grid, and quadrants sampled. However, this approach was noncompliant since it was not clear either how the road segments within quadrants were being sampled, or if there was a complete list of eligible road segments for each quadrant.

5.5 Multiple Traffic Studies

Some states and territories wanted to combine multiple traffic studies into one. This caused issues with the sampling design since some areas would be excluded for certain studies, but eligible for others. This also caused issues in large sample sizes being required, the correct probabilities being calculated, and estimates correctly reflecting the overall sample design. Multiple studies being built in together also could cause operational issues, such as not being able to include all vehicles eligible for sampling because too much information would need to be sampled from each one. However, the

Uniform Criteria did not state that every vehicle needed to be sampled, but a vehicle sampling rate and probability of selection needed to be known. Therefore, it could be compliant for a plan to sample every three vehicles that passed as to be able to collect all information needed for multiple studies.

5.6 Finely Stratified Designs

Finely stratified designs were fairly common. Such plans stratified road segments within counties into up to eleven different road types, allowing only two segments to be sampled per stratum to keep overall sample sizes reasonable. This design feature caused complications in nonresponse adjustment and variance calculations under road segment nonresponse, especially with plans with less than one reserve per original road segment, and therefore plans would need to address these complications and the appropriate countermeasures.

5.7 Tinted Windows

Some states and territories expressed their concerns regarding tinted windows. With tinted windows, a data collector would have difficulties observing seat belt use, causing a nonresponding observation. These nonresponding observations could lead to an overall nonresponse rate of over 10%. One solution to this was to sample vehicles passing, to focus observation resources on a smaller number of vehicles. Another solution was to have better training for data collectors on how to see seat belt use through tinted windows.

5.8 Stratification versus Stage of Selection

Clarification between the terms stratification and stage of selection was often required in plans, as many plans used these terms interchangeably. Many plans would say that counties were used as a stage of selection, but all eligible counties were taken with certainty. In actuality, counties were not stages of selection, but were used as strata for sampling road segments. As long as it was clear through the description of the methodology used for sampling, a plan was compliant when these terms were clarified.

6. Conclusion

When reviewing study and sample design plans for federal compliance, it is important to acknowledge that each plan will have its own design, but still must follow the published rule. The review tools created and utilized in this process worked well, allowed each review to be as uniform as possible, and enabled every plan submitted to ultimately become compliant. It was also important to distinguish between compliance and efficiency when reviewing study design plans as the latter can be more difficult and time consuming to review, and subject to interpretation.

Although this review process pertained to a seat belt use study, the process and tools utilized can work for most study and sample designs requiring review for federal compliance. All studies will have specific requirements that need to be followed, so checklists should be used. Model documents provide an example layout that helps make submissions uniform and review easier. Most studies will require resubmissions for federal compliance, making SharePoint sites helpful with keeping track of submissions and reviews. When resubmissions occur for minor edits, tracked changes software allows the changes in versions to be pinpointed.

Acknowledgements

The review team for the seat belt use study for federal compliance project consisted of Fran Bents, James Green, Annie Lo, Adele Polson, Rui Jiao, Jing Kang, Martha Rozsi, Matthew Airola, Emily Burkhardt, Amy Benedick, Doreen Deleonardis, and Shova KC.

We would also like to thank the NHTSA staff in charge of the seat belt use study, with special thanks to Timothy Pickrell and Fan Zhang.

References

"Uniform Criteria for State Observational Surveys of Seat Belt Use." Federal Register 76 (1 April 2011): 18042-18059. Print.