University of CINCINNATI | CENTER FOR POLICE RESEARCH AND POLICY

2021 Pennsylvania State Police Traffic Stop Study
January 1 – December 31, 2021

September 2022

Robin S. Engel, Ph.D.
Jennifer Calnon Cherkauskas, Ph.D.

Center for Police Research and Policy
University of Cincinnati

Please direct all correspondence regarding this report to RA-PSPCOMM@pa.gov.

I. INTRODUCTION

On January 1, 2021, the Pennsylvania State Police (PSP) recommenced the department-wide collection of data during all member-initiated traffic stops of members of the public. This data collection effort was based on the foundation of PSP traffic stop studies conducted from 2002 to 2010 in partnership with a research team from the University of Cincinnati (UC). The UC team analyzed the PSP-collected data, produced publicly available reports regarding the findings, including the extent to which racial and ethnic disparities were evident, and provided actionable recommendations for the PSP to address issues related to policy, training, supervision, and data collection. Informed by this previous work and recent developments in best practices in the law enforcement field, a committee of PSP officials and their UC research partners reconvened in September 2020 to design the current data collection process. This renewed data collection effort has several goals including to: 1) identify patterns and trends in the initiation of traffic stops and stop outcomes, including whether racial/ethnic disparities exist, 2) help make traffic stops more equitable and safer for motorists and Troopers, 3) build trust with the public by being transparent with traffic stop data and related findings, and 4) identify opportunities for improvement in policy, training, and supervisory oversight.

Historical Context of PSP Traffic Stop Data Collection

In January 2002, the *Police/Citizen Contact Policy Committee*, composed of PSP administrators and the Principal Investigator (Engel), developed the original Contact Data Report (CDR), a paper-based Scantron form completed by PSP Troopers during all member-initiated traffic stops. After pilot-testing and modifications, the department-wide data collection process began in May 2002. The information collected included the: (1) stop –e.g., date/time, location, duration, roadway type, and reasons for the stop; (2) driver – e.g., gender, age, race/ethnicity, and residency; (3) vehicle – e.g., state of registration, number of passengers; (4) stop outcome – e.g., citation, written warning, arrest, search, property seized during the search; and (5) the Troopers' assigned station and employee identification number.

Initially, the completed CDRs were collected at the station level and mailed weekly to the UC team. These forms were scanned by project personnel using the Scantron machine purchased by the PSP. Once scanned, the forms were stored securely until the electronic datasets were collated, audited, and considered ready for analysis, at which time the actual scan forms were destroyed through shredding.

In addition to analyses of the PSP stop data, the research team also conducted independent observations of roadway usage and speeding behaviors to provide alternative benchmark comparisons for the stop data. Three quarterly reports and one final year report based on these data were delivered to PSP administrators in January 2004 for the first year of data collection (May 1, 2002 – April 30, 2003). The data collection was extended for an additional year (May 1, 2003 – April 30, 2004), and a final report for Year 2 was issued in March 2005. Each of these reports documented trends in PSP-initiated traffic stops, along with post-stop outcomes, including warnings, citations, searches, and arrests.

The research team collected data for a third year (May 1, 2004 – April 30, 2005); a final report was not delivered, however, due to inaccuracies in the data collected that were initially discovered during focus groups with Troopers and confirmed through an internal data audit. Corrections to the data collection process were implemented in September 2005. The Year 3 data was compared to data collected in the fourth year (May 1, 2005 – April 30, 2006) to determine the level of inaccuracy. Based on these findings, a report combining Year 3 and Year 4 data was issued in 2006, representing data collected during calendar years 2004 and 2005.

A new contractual relationship in 2006 and an extension in 2009 allowed for the collection and analysis of five additional years of data (2006 – 2010). The PSP developed and implemented a new electronic data collection system, the CDR-Xpress, in 2006. This allowed for the data to be transmitted electronically to the UC team. Reports documenting the existence of any racial/ethnic disparities in post-stop outcomes during these five years were provided to PSP officials annually, with the final report of data collected in 2010 was issued in 2011. After 2010, the collection of data during PSP member-initiated traffic stops was discontinued.

The main findings of the data collection from 2002-2010 can be summarized as follows:

- Initial Traffic Stop
  - There was no consistent evidence to suggest that PSP Troopers disproportionately stopped minority motorists.
  - Although large racial/ethnic disparities existed between stops and Census-based benchmarks when stop data was compared to benchmarks that better capture roadway usage and driving behavior, these reported disparities were significantly reduced and, in some cases, eliminated.
- Post-Stop Outcomes
  - The reason for stop and other legally relevant characteristics were, substantively, the strongest predictors of **ALL** post-stop outcomes (e.g., warnings, citations, arrests, searches).
  - There were no statistically significant differences in warnings or citations for Black, Hispanic, or drivers of other races when multiple explanatory factors were simultaneously considered.
  - Black drivers were significantly more likely to be arrested only in Year 1; no racial/ethnic differences in arrests were found to be statistically significant in subsequent reports.
  - Data fields that were added in 2010 (e.g., criminal history, impairment) strongly predicted arrests and searches during traffic stops.
  - Hispanic and Black motorists were significantly more likely to be searched for discretionary reasons compared to Whites but less likely to have contraband seized during searches.
    - Racial/ethnic differences in searches and seizures persisted even after additional training, increased supervision, and improvements in data collection. These patterns were consistent with other state police/highway patrol agencies across the

country. This suggests there are likely more complex explanations for racial/ethnic disparities in searches than individual police officer bias alone.

Report Organization

This report is based on data collected from January 1, 2021 to December 31, 2021. This is the first year of PSP traffic stop data collected in over a decade. This report describes the implementation process for the rebooted data collection effort and documents findings from a comprehensive data audit. However, due to the issues documented in the data audit, no substantive analyses of the traffic stop data were conducted. The report is organized into four sections: 1) introduction, 2) development and modifications of the data collection protocol, 3) data audit, and 4) summary and recommendations.

Section 2: Development of the Data Collection Protocol

Section 2 includes a description of the PSP's development of the initial data collection instrument in 2020. It also summarizes and explains updates to the study's data collection protocol implemented throughout 2021.

Section 3: Data Audit

Section 3 first describes the components of data integrity, threats to data integrity, and common data collection errors that can impact the reliability and validity of collected data. It then documents the UC team's audit of the 2021 PSP traffic stop data based on these standards. The methods and results of the two-phase data audit of 2021 stop data are presented.

Section 4: Report Summary

Section 4 summarizes the information presented and the research team's plans for analyses of the data collected by PSP in 2022.

II.   IMPLEMENTATION OF CDR DATA COLLECTION

To ensure that traffic stop studies are based on reliable and valid data, it is imperative that police departments initiate data collection efforts that incorporate considerable forethought and planning.[1] This section describes the PSP's process of developing and revising the Contact Data Report (CDR) data collection protocol, beginning in 2020 and continuing throughout 2021.

Development of Data Collection Protocol

When implementing a new data collection effort, the following factors are among the most important to consider: 1) selecting the mechanism for data collection, 2) developing the data collection instrument, 3) conducting a pilot test, 4) reinforcing the importance of data collection through training, supervisory oversight, and appropriate policy changes, and 5) developing a data auditing system. Each of the steps undertaken by the PSP to address these factors are described below.

Selecting the Mechanism for Data Collection

In September 2020, the PSP and the UC team reestablished their research partnership to collaboratively develop a new data collection process based on the foundation of the work that had been completed between 2002 and 2010 and incorporating more recent best practices. The PSP established a committee that included a designated project manager, representatives from executive staff, the Bureau of Communications and Information Services, and technological and database experts. It was determined that the most efficient and reliable method for Troopers to collect CDR data is to enter data electronically through their mobile data terminals (MDTs) in a software application already used by the PSP called TraCS (Traffic and Criminal Software). Electronic data collection is regarded as the most efficient and accurate method of capturing information related to traffic stops.[2]

Developing the Data Collection Instrument

The specific data collection instrument utilized by PSP Troopers was developed over the course of several meetings by the PSP internal committee. The committee was guided in their initial decisions by the department's previous data collection effort and the UC team's technical assistance based on their experiences with other agencies and a review of recent developments in

---

[1] Deborah Ramirez, Jack McDevitt, & Amy Farrell, (2000), "A Resource Guide on Racial Profiling Data Collection Systems: Promising Practices and Lessons Learned," US Department of Justice, Washington, DC., (2000); Tillyer, Rob, Robin S. Engel, and Jennifer Calnon Cherkauskas. "Best practices in vehicle stop data collection and analysis." *Policing: An International Journal of Police Strategies & Management, 33*(1), 69-92, (2010).
[2] Lorie Fridell, *By the Numbers: A Guide for Analyzing Race Data from Vehicle Stops,* Washington, D.C.: Police Executive Research Forum, (2004); Tillyer et al., 2010.

the field.[3] One of the primary goals in designing the PSP data collection system was to minimize the impact on Troopers' efficiency in the field and avoid redundancy in data entry.

To facilitate the auto-population of data fields in the CDR from other forms, members of the PSP's internal committee coordinated the linkage of the CDR form to other TraCS forms (e.g., Traffic Citations, Police Written Warnings), other templates in PSP's Records Management System (RMS), and PSP personnel data. The need for manual entry of information, which increases the chance of errors, was minimized by the use of drop-down menus whenever possible. The specific elements included on the form evolved as described in detail below. The final count of data fields as of February 2022 was 60; 22 are auto-populated, 26 employ drop-down menus with single options, five use drop-down menus with multiple selection options, and another seven may be auto-populated or manually entered.

Conducting a Pilot Test

Data collection pilot tests are simply a "dry run" for the data collection effort. They ensure that the research design is feasible, and the data collected is both reliable and valid. Pilot tests are typically conducted by a selected group of officers in a more limited geographic area. Based on findings from the pilot test, the data collection instrument is changed, and officer training is modified (if needed). The initial CDR form was pilot tested in two PSP Troops in late 2020. The UC team provided feedback on the initial form, resulting in minor improvements being implemented. This pilot test also allowed for a test of the operational procedures developed to transfer the CDR data from the PSP to the UC team in a manner that assured its security and confidentiality. Thereafter, PSP automated a weekly secure data transfer to the UC team.

Training, Supervisory Oversight, and Policy Updates Related to the Data Collection

To reinforce the changes associated with the new data collection process, the PSP created Field Regulation (FR) 6-18 "Contact Data Reporting" [4] and updated several regulations affected by the initiation of the CDR data collection. Field Regulation 6-18 describes the purpose of the CDR, instructions for when members are and are not required to complete CDR forms, and the responsibilities of the submitting PSP member and their reviewing supervisor. Specifically, it states, "Contact Data Reporting (CDR) enables the Department to collect and analyze data associated with traffic stops and arrests." FR 6-18 requires that CDR be completed for every member-initiated traffic stop regardless of enforcement outcome or whether the member prepares a separate department report. The CDR form is not required when PSP members respond to disabled motorists or vehicle accidents; stops that result from traffic safety checkpoints or other non-discretionary enforcement details are also excluded.

---

[3] Marie Pryor, Philip Goff, Farhang Heydari, & Barry Friedman, 2020). "Collecting, Analyzing, and Responding to Stop Data: A Guidebook for Law Enforcement Agencies, Government, and Communities," (2020), https://policingequity.org/images/pdfs-doc/COPS-Guidebook_Final_Release_Version_2-compressed.pdf.

[4] https://www.psp.pa.gov/contact/RTKL%20DOCUMENTS/FR%206-18.pdf

The initial training on this regulation and the data collection requirements was delivered via PSP's internal e-library system, where members are required to review all new regulations. PSP members were instructed to "complete CDR when required by this regulation and/or other Department directives, in accordance with the instructions in the TraCS online documentation, Help File, and internal Validation Function." The TraCS software through which a CDR is completed permits Troopers to click on any data field in the form and access documentation that provides further description of the field and what is required. PSP has developed additional CDR training based on preliminary results from 2022 that will be completed by PSP members during the fourth quarter of 2022.

Other departmental regulations related to enforcement actions were updated to include the requirement for CDR when any self-initiated traffic stop results in the completion of a:

- Traffic Citation (*RM 6-1; Traffic Citation/TraCS Traffic Citation*)
- Police Warning Notice (*RM 6-2; Police Warning Notice/TraCS Police Warning Notice*)
- Non-traffic citation (*RM 6-3; Non-Traffic Citation/ TraCS Non-Traffic Citation*)
- TraCS Criminal Complaint, or whenever any other county-specific criminal complaint form is used instead of the TraCS Criminal Complaint. (*RM 6-4; TraCS Criminal Complaint*)

The online training also included a video recorded by Dr. Engel, the lead researcher from the UC team, which was also available on the PSP intranet. This video explained the renewal of the CDR program, including a summary of the previous data collection effort. It explained the role of the research team, the types of analyses that would be conducted, and what the UC team can and cannot do with the provided data. For example, Dr. Engel emphasized that the team would use the data to identify any patterns and trends in racial/ethnic disparities but will not conduct any analyses to identify individual PSP members. It noted the methodological limitations of comparing traffic stop data to residential population statistics and clearly stated that this type of data cannot measure individual officer bias or attribute racial/ethnic disparities in stops or stop outcomes to individual officer bias.

In addition to the changes related to data collection, it is important to note that PSP policy, in effect since April 28, 2004, strictly prohibits bias-based profiling, defined as "any traffic stop, field contact, vehicle search, asset seizure/forfeiture, or enforcement action based on a common trait of a group. Common traits include, but are not limited to, race, ethnic background, gender, sexual orientation, religion, economic status, age, or cultural group." The policy states that "Traffic stops, field contacts, vehicle searches, asset seizures/forfeitures, and enforcement actions shall be conducted in accordance with existing law, Department directives, and regulations."[5]

---

[5] PSP AR 4-37 *Bias-Based Profiling Review*.

Data Auditing

Regardless of the sophistication of the statistical analyses and benchmark comparisons utilized by researchers, the research study is virtually meaningless if the traffic stop data itself is not valid. In addition to maintaining data integrity (discussed in Section 3), a data auditing system can also help ensure officer compliance with the data collection protocol. Officers will likely be more diligent in their data collection if they know it is being reviewed for comprehensiveness and quality.[6]

Supervisors are responsible for reviewing, approving, and filing Trooper submissions, as well as ensuring that any corresponding general offense reports in the Records Management System have been flagged as related to the CDR study. The software also sends an email notification anytime a member initiates a TraCS form that requires a CDR but was not completed.

The PSP has been actively working on the development of its internal TraCS CDR Audit Report. This process involves automating the comparison of the number of completed CDR forms with the number of TraCS warning and citation forms as well as computer-aided dispatch (CAD) data. It provides a mechanism for PSP supervisors to review cases that may have required a CDR, but for which one was not completed. The PSP anticipates that the audit dashboard will be fully operational for supervisors department-wide in late 2022.

Initiation of Data Collection

Following these preparatory months, Colonel Evanchick issued *PSP Special Order 2020-55* on December 30, 2020, which required PSP members to complete a CDR for all member-initiated traffic stops beginning January 1, 2021. This notification explained the purpose of the study, emphasized the importance of collecting this type of data as a method of providing transparency to the public, provided a summary of the regulations affected by the new data collection (summarized above), and included a list of the data that would be shared with the UC team. The bulletin also asked all PSP members to view the video recorded by Dr. Engel, described above, that was included in the CDR training.

Updates to the Data Collection Protocol

Data collection began department-wide on January 1, 2021. Throughout the year, the UC team periodically assessed the compiled stop data for errors and issues related to data integrity at the aggregate level. The UC team communicated these findings to the PSP in a series of memos and meetings. This allowed the PSP an opportunity to address and correct issues with the data collection process without directly identifying Troopers.

The UC team also provided a series of recommendations to the PSP regarding potentially important data fields to consider adding to the CDR form. The UC team provided these recommendations cognizant of the fact that every relevant factor that might explain traffic stops and stop outcomes cannot be realistically gathered in a single data collection system. However,

---

[6] Fridell, 2004.

the more relevant data elements that are included, the better we can understand statistical findings and provide meaningful policy and training recommendations. Ultimately, the PSP was responsive to many of these recommendations (described in detail below), while keeping in mind the delicate balance between including additional data fields and the need for an efficient data collection system in the field.

A comprehensive list of the additions and updates to the PSP data collection protocol is summarized in Table 2.1. Some of these changes were designed to address data integrity issues identified by the UC team, while others were made in the interest of adding potentially relevant explanatory factors that will assist the UC team in the analysis of post-stop outcomes. Like the pilot test that occurred prior to initiating department-wide data collection, changes to the data collection system made during June and December 2021 were implemented in two different PSP stations prior to department-wide rollout. Follow-up changes to the data collection system were made via Departmental Bulletin from the Director of the Bureau of Communication and Information Services and discussed during Roll Call.

After these revisions, the final CDR form (see Appendix) includes the following information:

- **Stop** – date/time, location (county and municipality, and latitude/longitude), type of roadway, use of canine, duration of the stop, reason(s) for the stop, whether the stop was related to a Special Traffic Enforcement program or Motor Carrier Safety Assistance program, and information related specifically to speeding violations (e.g., posted speed limit, amount over limit, etc.)

- **Driver –** gender, age, race/ethnicity, zip code of residency, compliant or resistant behavior, whether the driver was a foreign national,[7] whether the driver had limited English proficiency (LEP), and if yes, the type of language assistance utilized

- **Vehicle** – state of registration, number of passengers[8]

- **Outcome of the stop** – whether the driver and/or passenger was issued a citation (including the number of citations) or warning (including whether it was a verbal or written warning and the number of warnings), whether the driver and/or passenger was arrested and/or searched, and if a search was conducted roadside or following vehicle tow, reason(s) for the search[9] and whether property was seized

---

[7] If the driver or passenger is reported as a foreign national (DFN or PFN) a series of additional questions are required including the DFN race/ethnicity, whether the communications desk unit or supervisor was contacted, whether ICE was notified, and if yes, the reason and result, whether the DFN or PFN was detained and the reason and result, whether ICE has an administrative or criminal warrant for the DFN or PFN.

[8] If passengers are present, there are additional data fields for Troopers to complete, including the passenger's race, ethnicity, whether their identification was requested, and if yes, the type of identification provided.

[9] Modifications to the search initiated and search reason data fields were made as a result of the *Commonwealth v. Alexander*, 243 A.3d 177 (Pa. 2020) decision by the Pennsylvania Supreme Court, which deemed probable cause warrantless searches of motor vehicles unconstitutional (contrary to the Federal vehicle exception) unless exigent circumstances exist in addition to probable cause.

- **Identifying information** – Troopers' assigned station, employee identification number, and demographic characteristics

Note that the gender and racial/ethnic characteristics of drivers were determined by officers' perceptions rather than asking drivers to identify their gender, race, or ethnicity. This is consistent with the guidance of best practice guides regarding traffic stop data collection; identifying driver race/ethnicity based on officers' perceptions is the recommended method of data collection for examining racially biased policing. [10] The prevailing consensus on this matter is that asking drivers to self-identify their race/ethnicity has the potential to escalate tensions during police encounters with the public. Furthermore, although troopers may incorrectly perceive drivers' actual characteristics, this possible misperception is irrelevant for data collection analyses that seek to explain officer-decision making. Concerns regarding racial profiling are based on the presumption that officers treat people of color differently due to their personal bias. Therefore, proper data collection efforts must identify officers' perceptions of the race/ethnicity of the driver, not necessarily the driver's actual race/ethnicity. As stated by Ramirez and colleagues[11], "Whether the officer correctly ascertains the race or ethnicity of the driver is less important than being able to analyze whether, having perceived the driver is a person of color, the officer treats the person fairly."

---

[10] Lorie Fridell, Robert Lunney, Drew Diamond & Bruce Kubu, "Racially Biased Policing: A Principled Response, Police Executive Research Forum," (2001), https://www.policeforum.org/assets/docs/Free_Online_Documents/Racially-Biased_Policing/racially%20biased%20policing%20-%20a%20principled%20response%202001.pdf; Pryor et al., 2020; Ramirez et al., 2000.

[11] Ramirez et al., 2000, p. 47.

Table 2.1. Summary of Changes to Data Collection Protocol throughout 2021

Type of Change	Related Data Field(s)	Description of Change	Reason for Change	Effective Date
Update	Search initiated	Revised to indicate whether a search was conducted roadside or after vehicle tow	Assist in evaluating vehicle searches as a result of the *Commonwealth v. Alexander* decision.	6/29/2021
Update	Reason for search	PC now PC + exigency, K-9 alert and odor of drugs/alcohol eliminated, and consent separated to reflect verbal and written consent.	Modified to account for the *Commonwealth v. Alexander* decision.	6/29/2021
Addition	Whether K-9 was used	Documents canine use during the stop.	Assist in possibly explaining the duration of certain traffic stops.	6/29/2021
Addition	Driver behavior	New data field to document compliance and/or resistance with or toward the Trooper.	May improve the explanatory power of statistical models explaining stop outcomes.	6/29/2021
Addition	MCSAP	New data field to document stops related to the Motor Carrier Safety Assistance Program.	Assist in studying any enforcement differences related to the inspection of commercial motor vehicles.	6/29/2021
Update	Driver's zip code	Field revised to 1) record international zip codes as 99999 and 2) limit manual entry to five digits.	Correct data entry errors related to invalid zip codes and zip codes that were not 5 or 9 digits.	1) 6/29/2021 2) 2/16/2021
Addition	Driver is foreign national (DFN)	CDR linked to the RMS Foreign National Text Template to automatically extract information related to if DFN.	Assist in studying any variations in stops involving foreign nationals.	6/29/2021 1/11/2022
Update	Driver's actual speed	Field limited to three digits and maximum value of 165 mph.	Correct data entry errors related to three- or four-digit values that exceed 165 mph in a small number of cases.	12/16/2021
Addition	CDR form version	New data field to document which form version was used.	Assist in tracking effective dates of changes to CDR and missing data.	12/16/2021
Update	Driver's date of birth	Rule established to warn Trooper of possible error if age is less than 16 years old.	Minimize data entry errors related to invalid dates of birth.	12/16/2021
Addition	Driver warning type; driver # of warnings	New data field to distinguish verbal and written warnings; rule established to only count # of written warnings.	Initiated by the PSP to ensure CDRs were completed for all stops, including those without a formal written disposition.	12/16/2021
Addition	Dedicated Enforcement Team (DET)	New data field to document stops made by members of DET, and if yes, which one.	Assist in examining any enforcement differences related to assignment to a DET.	12/16/2021
Addition	Limited English proficiency (LEP)	New data field to document whether driver had LEP, and if yes, the type of language assistance employed.	Assist in possibly explaining the duration of certain traffic stops.	1/11/2022
Addition	Multiple passenger-related fields	New data fields to document passengers' race, ethnicity, LEP, whether ID requested, and if yes, type provided.	Assist in possibly explaining the duration of certain traffic stops.	1/11/2022
Addition	Target of search	New field to document who/what was searched (e.g., driver, passenger, and/or vehicle).	Assist in examining differences in search success rates by search target.	1/11/2022

III.   DATA AUDIT

The primary purpose of data collection is to provide the means for performing rigorous and robust analyses in which accurate conclusions can be drawn. In the case of traffic stop studies, the purpose is to better understand the factors that influence officers' stop initiation and enforcement decision-making and to assess whether the results of these decisions are equitable. To draw such conclusions, one must ensure that the data that are reliable, valid, and free of error.[12] Data integrity is, therefore, a crucial component to effective data analyses and ensuring that any conclusions drawn, and recommendations provided, are based on the highest quality data possible.

This section first describes the components of data integrity, threats to data integrity, and common data collection errors that can impact the reliability and validity of collected data. It then documents the UC team's audit of the PSP 2021 traffic stop data based on these standards.

Data Integrity

A recent best practice guide for traffic and pedestrian stop data collection identified the following elements of data integrity:[13]

1) For each reported stop, the data captured are complete (i.e., there is no missing data).
2) For each reported stop, the data captured are free from error (i.e., data is internally consistent and reliable).
3) Data are collected for all stops for which reporting is required.

There are several potential threats to data integrity, including "different officers collect data using different criteria, different officers use different methods to collect data, and different officers interpret data points differently."[14] Safeguarding against them requires a multi-faceted approach by the police agency to minimize the likelihood of these errors, including clear policy guidance, training, and supervisory oversight, as documented in Section 2.

Data collection errors can result from issues related to the front-end design of the data collection system as well as back-end user errors.[15] Design errors can include inconsistent coding of variables across time or across users, overlapping or not mutually exclusive categories for included variables, data fields that require manual data entry that could be categorical, and the inclusion of an "other" category without the ability for users to specify further. Most of these errors can be mitigated through proper planning and pilot testing, which can also minimize some user errors along with proper training. Typically, data audits for traffic stop studies check for several types of inaccuracies related to user error,[16] including:

---

[12] Eric Loken & Andrew Gelman, "Measurement error and the replication crisis," *Science, 355* (6325), 584 – 585, (2017).
[13] Pryor et al., 2020.
[14] Pryor et al., 2020, p.23.
[15] Pryor et al., 2020.
[16] Fridell, 2004; Pryor et al., 2020.

- Incorrect copying of information from one form to another (e.g., data transfer or entry errors)
- Missing specific information on individual forms (i.e., no information entered by the PSP member)
- Illogical information (e.g., date of birth in the future)
- Conflicting data entry between related data fields on individual forms (i.e., inconsistent information like search reason provided but search initiated reported as "no")
- Missing information about some member-initiated stops conducted (i.e., no CDR form generated)
- Data contains intentional misstatements of facts (e.g., Black motorist is recorded as White)

As described in Section 2, PSP's use of an electronic data collection system (TraCS) and its built-in data validation mitigates the likelihood of many of these inaccuracies.[17] For example, linking the stop data collection system with other existing department systems to maximize the use of auto-population decreases the chances of data transfer errors. Similarly, the PSP built rules into TraCS to warn users about conflicting data entry and illogical information, as well as to prevent submission of the CDR until all required fields are completed. Unfortunately, other types of inaccuracies are more difficult to detect. For example, electronic data entry cannot ensure that officers are completing the CDR for every required stop in the same way that it minimizes data entry errors.

There are also limited methods for directly assessing intentional distortion of collected data. For example, since officers are supposed to report their perceptions of driver race/ethnicity, it is impossible to determine if a discrepancy between recorded and actual race/ethnicity was intentional versus honest errors in perceptions. The research team and PSP included measures to increase the likelihood that PSP members collect data reliably.[18] First, the contractual agreement between the PSP and UC team guarantees confidentiality to each Trooper and prohibits any data analyses that would identify individual Troopers. Second, the lead researcher of the UC team, Dr. Engel, spoke to the leadership and members of the Pennsylvania State Troopers Association on two occasions. She initially met virtually with PSTA President and a small group of members in December 2020 and then virtually attended the annual PSTA meeting in January 2021; on both occasions Dr. Engel provided an overview of the CDR study and answered questions related to any concerns members had about the data collection effort.

It is worth noting that *PSP Special Order 2020-55,* which initiated the requirement for PSP members to complete CDR, included reference to the data collection effort as a demonstration of the *Pennsylvania State Police Core Values* of *Honor, Service, Integrity, Respect, Trust, Courage,*

---

[17] Pryor et al., 2020.

[18] These protections are included in the contract and approved by the University of Cincinnati Institutional Review Board, PSP legal team, and PSP union officials. The Principal Investigator advised PSP Troopers of this confidentiality agreement in a training video.

and Duty. The UC team recommends that departmental expectations regarding the completion of CDR data be clearly reinforced.

Data Audit

Data auditing is an important mechanism to assess data integrity prior to engaging in statistical analyses. It is the systematic process of evaluating the reliability and validity of the collected data. Data reliability refers to the stability or consistency of the items being measured (i.e., is the variable being measured consistently across cases). Having reliable data is important so that observed changes in the data can be confidently reported as real changes rather than simply changes in the data collection. Data validity refers to the overall accuracy of the measure (i.e., does it measure what it is supposed to be measuring).

No data collection is perfect, but the minimization of measurement errors (i.e., the difference between observed and true values) is critical because they can lead to biased or incorrect conclusions drawn from data analyses. It is particularly important to mitigate against systematic measurement error. Random measurement error is an error that tends to naturally find its way into a database due to chance factors; because it is inconsistent and unpredictable, its impact on conclusions is likely to be small, given that random errors are assumed to cancel each other out in an analysis.[19] Systematic measurement error, on the other hand, is an error in a database that produces a bias in the data because the error is consistent across all cases of the measure. Data that are inaccurately collected in a consistent manner may not affect the reliability of the measure, but validity will likely be severely impacted.[20]

From January 1, 2021 – December 31, 2021, PSP Troopers collected data on 439,104 member-initiated traffic stops. This section summarizes the results of a two-phase data audit of CDR data collected in 2021 that focuses primarily on missing data, logical inconsistencies, reliability of data fields, and the completeness of the data by comparing the number of stops in the electronic data with other independent sources of information.

Data Audit—Phase I

<u>**Description**</u>

Phase 1 of the audit for 2021 data assesses the degree to which the data captured by PSP Troopers are complete and free from error. This involves an examination of missing data[21] (i.e., no information entered by the officer), logical inconsistencies (i.e., fields with missing and/or incorrect entries that contradict other fields), and the reliability of the data collected. The fields

---

[19] Royce A. Singleton & Bruce C. Straits, *Approaches to social research* (6th edition), Oxford University Press, (2017); Madhubalan Viswanathan, *Measurement error and research design,* Sage, (2005).
[20] Singleton & Straits, 2005.
[21] Some missing data is to be expected and therefore not included in this table. For example, posted speed limit, drivers' actual speed, and amount over the posted limit should only be filled out for speeding stops. For data fields that should only be completed in certain circumstances, the missing data percentage is based on the valid number of stops for which it would be expected to be included.

analyzed in this data audit were assessed based on whether they conform with the guidelines provided in CDR Data Dictionary codebooks.[22]

Results

Table 3.1 reports the percent of missing data and erroneous or inconsistent data for most data fields in the extracted CDR data for 2021. Table 3.1 includes only data fields that were collected for the entire calendar year; as noted in Section 2, some data fields were added throughout the year.[23] The reason for this exclusion is that until December 2021, there was no method for identifying which version of the CDR form was in use, thus making it impossible to determine whether data was actually missing or whether the data field simply did not exist yet at the time of the stop. This issue is discussed further below.

In 2004, the Police Executive Research Forum, a police research and policy organization, published a comprehensive guide for analyzing data from traffic stops that remains a resource for law enforcement agencies nearly two decades later. In this guide, an error rate of less than 10% was recommended for traffic stop data.[24] Our research team recommends a more stringent standard of under 5%, with a goal of 2% missing/invalid data. Based on these standards, the results of this portion of the data audit are positive. As shown in Table 3.1, most of the variables examined have either no or very little missing or invalid data. Overall, the data validation built into the TraCS system, and the revisions made throughout 2021, have minimized the errors related to both missing and invalid data, with the exception of the data fields related to special traffic enforcement and number of warnings.

As data collection continues, the UC team recommends that PSP periodically review the default settings, validation rules, and error warnings in the TraCS system and incorporate minor adjustments when appropriate. For example, although many validation rules already exist, others may be warranted to warn users of the possibility of conflicting data entry (e.g., if arrest = no but search reason = incident to arrest, or search reason = plain view but contraband seized = no).

---

[22] Three versions of the codebook were provided to the UC team throughout 2021 as changes were made to the data collection protocol.

[23] These variables include driver behavior, K-9 utilized, MCSAP-related, dedicated enforcement team, and warning type. These variables will be assessed in data audits provided in the quarterly and annual reports moving forward.

[24] Fridell, 2004.

Table 3.1: Missing and Invalid Data from Member-Initiated Traffic Stops (n=439,104), Jan-Dec 2021

	% Missing	% Erroneous or Inconsistent
Stop Characteristics		
Date of Contact	0.00	0.00
Time of Contact	0.00	0.00
Location of Stop[25]	<0.00	0.00
Roadway Type	0.00	0.00
Duration of Stop	0.00	0.00
Reason for the Stop[26]	0.00	0.01
Special Traffic Enforcement[27]	9.04	0.00
Outcome of the Stop		
Number of driver warnings	2.06	0.00
Number of driver citations	0.00	0.00
Driver Arrest	0.00	0.00
Driver Characteristics		
Year of Birth	0.00	0.07[28]
Gender	0.00	0.00
Race	0.00	0.00
Ethnicity	0.00	0.00
Zip Code	0.00	0.22[29]
Vehicle Characteristics		
State of Registration	0.00	0.00
Number of Passengers	0.00	0.00
Trooper Characteristics[30]		
Employee ID number	0.00	0.00

Note: <0.00 reflects less than 0.005% error.

[25] A "valid location of stop" exists if there is a valid county and municipality code entered and/or valid latitude and longitude coordinates provided. Latitude and longitude are auto-populated from various TraCS forms (e.g., warning, citation, etc.), while county and municipality codes are auto-filled from the location selected in the TraCS Location Tool (TLT). If information is missing from original forms, it also appears as missing in CDR data. Although there were 496 CDRs missing county and municipality and 4,599 CDRs missing latitude and longitude, only 4 of the 439,104 CDRs were missing both types of location identifying data.

[26] These percentages reflect the inclusion of valid data for posted speed limit, actual speed, and amount over speed limit for stops made based on speeding violations. The 0.01% invalid represents 34 CDRs with speeds that exceed 165 mph.

[27] Special traffic enforcement (STE) was originally one of the response options under *reason for stop* but was removed from that data field's options in December 2020 during the pilot test phase because Troopers must have a legal reason for the stop (i.e., traffic violation) regardless of whether the stop is associated with a STE program.

[28] Although there is not missing data, there were 317 CDRs with erroneous entries for date of birth that were before 1/1/1921 or after 1/1/2011 (i.e., date of birth too new or old to be valid drivers).

[29] There were 956 CDRs that include zip codes with five digits not in the US zip code database and not equal to 99999, the PSP codebook designation for international addresses.

[30] The CDR form requires employee ID number, which links to an external personnel database and auto-populates the CDR data with information regarding Trooper gender, race, years of service, rank, current assignment/job code, and assigned station code. Therefore, the percent of missing and/or invalid data on employee ID number represents the percent of missing and/or invalid data for all Trooper characteristics.

Although the results of Table 3.1 are favorable, a separate significant concern for data integrity is not included those results. As described in Section 2, several updates to improve the data collection system occurred throughout 2021. Unfortunately, when a new CDR form is released, it is not automatically updated to all PSP technology on a single date. Although updates are mandatory, they must be manually downloaded by each individual PSP member; therefore, during 2021, the updates occurred over a period of time.[31] As a result, multiple CDR forms could be in use across the agency simultaneously. The TraCS system has a form identifier that is stored with the CDR data record. Unfortunately, this form identifier was not included as a data field in the data file provided to the UC team until December 2021 and could not be retroactively added. Due to this data collection design error, it was impossible for the UC team to determine the effective dates for variables that were collected for only part of the year (i.e., what the valid number of cases should be) or what the valid categories in use for data fields were on a specific date.

Table 3.2 documents the changes in the two data fields most significantly affected by this data collection design error: search initiated and search reason. As shown in the table, the measurement of these variables substantially changed. For example, "search initiated" was originally captured as yes = 1 and no = 2. Beginning with the 6/29/21 update, search initiated was changed to a three-category variable, with two options for yes denoting different types of searches (i.e., road search and search after tow) and a new option for no search (no = 3). Without knowing when each Trooper updated their system to the new form, it is impossible to determine when the "2" for search initiated reflects "no search" and when it indicates a "tow search." Similarly, the values for categories of search reasons also changed, with some reasons eliminated, others added, and the numeric codes for all categories differing from the previous CDR form to the updated form. Again, without knowing the form in use for each reported search, it is impossible to determine the actual reason for the search due to the inconsistency in coding.

When data are not collected in a consistent and systematic way, they become unreliable and will produce invalid results. As such, changes to a database need to be clearly documented so any potential errors can be appropriately accounted for in future analysis. Without the appropriate documentation of the timing of the changes to the PSP CDR form and the inconsistent measurement of the "search initiated" and "search reason" across time, these data fields are rendered unreliable. As a result, for the 2021 data, the UC team cannot reliably:

- assess whether there are missing data or logical inconsistencies for search-related data fields,
- count the number of searches conducted,
- examine the types of searches based on search authority and level of officer discretion, or

---

[31] Although all updates to the TraCS system have been mandatory, initially they could be skipped for up to two weeks. The decision not to mandate the update immediately was made due to the potential for the required update to negatively impact PSP operations. Since July 19, 2022, all system updates are mandatory and cannot be skipped. TraCS will not function until the update is completed.

- assess whether any racial/ethnic differences in searches or reason for search exist.

Given that the UC team's previous analyses of PSP data from 2002 – 2010 also documented data collection issues associated with similar fields and racial/ethnic disparities in searches, types of searches (i.e., search reason), and seizure rates, the inability to reliably examine this information in 2021 is a significant limitation for any substantive statistical analysis.

Table 3.2: Updates to CDR Data Fields Related to Searches

	Previous Numeric Value	New Numeric Value
Search Initiated Categories		
Yes	1	N/A
Yes, Roadside	N/A	1
Yes, Towed and Searched Elsewhere	N/A	2
No	2	3
Search Reason Categories		
Not Applicable	0	Eliminated[32]
Incident to Arrest	1	0
Inventory	2	1
K-9 Alert	3	Eliminated
Odor of Drugs/Alcohol	4	Eliminated
Officer Safety (Terry)	5	2
Plain View Contraband	6	3
Probable Cause	7	4 (changed to PC + exigency)
Search Warrant	8	5
Other	9	Eliminated
Consent	10	6 (Written Consent)
		7 (Verbal Consent)

Data Audit—Phase 2

Description

Phase 2 examines the data accuracy by comparing the number of stops in the electronic CDR data to the number of stops in an independent source of information to assess whether all stops recorded in the external source of information are represented in the CDR data. This type of audit determines the extent to which data collection forms are completed as required and addresses the question of data validity; that is, whether CDR data actually represents all member-initiated traffic stops regardless of outcome.

To determine whether the information is being recorded for all eligible traffic stops, an external data source that records the same stops is necessary. Typical comparison sources of data include computer-aided dispatch (CAD) data, citation data, written warning data, videotapes, or other departmental data.[33] Based on discussions with PSP personnel, it was determined that the most appropriate and comprehensive comparison data would be CAD calls coded as traffic stop

---

[32] With the revision, if no search is initiated, search reason does not open as a field for completion and is automatically set to missing data; that is, there is no need for a "not applicable" field.
[33] Fridell, 2004; Ramirez et al., 2000.

incidents provided by the PSP.[34] The reporting standards are almost identical between the two datasets; however, some exclusions were made from the CAD data to ensure an "apples to apples" comparison.[35]

Results

Table 3.3 compares the aggregate number of traffic stops included in CAD calls coded as traffic stops with the total number of traffic stops included in the CDR data for the PSP overall, as well as at the Area, Troop, and Station levels.[36] Note that the sum of the stops reported in each data source by each of the stations and specialized units does not equal the total number of stops reported department-wide in either data source because a small number of stops (0.09% of CDR and 0.8% of CAD) are made by PSP organizational units outside of the area commands or specialized SHIELD and Canine units.

A percent error is provided across organizational levels, representing the percentage of traffic stops that do not match across the two data sources. The percent error is calculated as follows, where the "observed value" equals the count of stops in the CDR data and the "true value" equals the count of stops in the CAD data:

$$\text{Percent Error} = \frac{V_{observed} - V_{true}}{V_{true}}$$

Positive error rates indicate the percent of stops that appear in the CDR data but not in the CAD records. Conversely, negative error rates indicate the percent of stops that appear in the CAD records but not in the CDR data.

Overall, the results displayed in Table 3.3 show that the percent error between the two datasets at the department level is -12.1%, which exceeds the PERF-recommended correspondence of 90% or more between two sources of information.[37] Using this same standard of 10% error to examine the smaller organizational units, the results of this audit are concerning. Only about one-third of the stations and the SHIELD specialized unit had error rates of 10% or lower. There is no consistent pattern across all stations; although the majority (75 of 88 stations and the two

---

[34] It is important to note that CAD codes for other types of traffic stops that are not Trooper-initiated are coded differently (e.g., when a dispatcher receives a report of a traffic violation like an erratic driver and assigns it to a Trooper for response). This is important because they can be distinguished from the CAD incidents when a Trooper initiates a traffic stop and self-generates a call number.

[35] Specifically, to ensure that only Trooper-initiated stops in the CAD data were compared to the CDR data, 6,815 motor carrier enforcement related stops and 820 disabled motorist related stops were excluded as these are not CDR-required stops. Further, to guarantee each CAD incident is only counted once, 1,259 duplicate incidents were excluded. Finally, to ensure that CAD incidents actually resulted in a stop that would generate a CDR, 130 CAD incidents that involved a pursuit without an apprehension and 818 canceled CAD incidents were excluded since they did not result in the stop of an individual.

[36] The PSP is organized into multiple managerial command levels, including 4 Areas, 16 Troops, and 88 Stations. Information in all reports produced by the UC team is presented for the PSP department, Area, Troop, and Station levels to illustrate differences across organizational units. Presenting information in this manner permits the identification of units that may appear as outliers, providing opportunities for closer examination and focused attention by PSP officials.

[37] Fridell, 2004.

specialized units) have fewer CDRs than stops in the CAD records, there are thirteen stations with the opposite pattern.

This confirms internal findings from routine CDR audits that the PSP conducted throughout 2021, which discovered discrepancies between computer-aided dispatch (CAD) records of traffic stops and CDR data. Initial estimates from the PSP indicated up to 10% of stops may not have been recorded on a CDR. The reasons for the discrepancies cannot be specifically determined, but the PSP attributes most of these issues to situations where Troopers issued verbal warnings.[38] Although the guidance regarding completion of CDRs in *Special Order 2020-55* and *FR 6-18, Contact Data Reporting* indicated they should be completed for all member-initiated stops regardless of enforcement outcomes or whether the member prepared a separate department report (i.e., citation, warning notice, etc.), verbal warnings were not specifically referenced. Therefore, it is plausible that this underreporting of stops resulting in verbal warnings was simply a misunderstanding of the reporting requirements.

PSP proactively responded to this internal discovery by reinforcing to its members the need to complete CDRs for all member-initiated traffic stops, even when no written disposition or formal enforcement outcome results. The PSP also added a data field to the CDR form to capture driver warning type (i.e., verbal or written) to reinforce the expectation that CDRs were to be completed for all stops, including those without a formal written disposition. It is unknown if there are other issues with underreporting that may be contributing to the findings reported in Table 3.3. It is recommended that the PSP continue to refine quality assurance measures and supervisory oversight to ensure members are completing the data collection form for every required contact. Continual supervisory oversight and regular data audits –like the system currently in development by PSP – are necessary to ensure the continued accuracy and validity of these data.

Unfortunately, the extent to which the CDRs do not accurately represent all member-initiated traffic stops by PSP members raises serious concerns regarding the validity of the 2021 data. Furthermore, there is the added concern that the stops for which CDRs were not completed differ in a systematic way from the stops with CDRs. That is, this is not a random measurement error. There is also the possibility that the systematic error could be related to drivers' race/ethnicity. For example, if drivers of color were significantly more (or less) likely to be issued a verbal warning compared to White drivers, then not having valid data for these stops could make it appear drivers of color were more (or less) likely to be stopped compared to White drivers than they actually were.

---

[38] It is possible that some verbal warnings may have been documented on the CDR as 1.4% of stops reported no outcomes (i.e., written warning, citation, or arrest) for either driver or passenger. Another possible explanation for a small percent of these stops without a CDR is related to the SHIELD program or troop special enforcement team being requested to initiate a stop by federal law enforcement partners, resulting in the initiation of federal charges but no enforcement action by the PSP. A recent guidebook on stop best practices suggests that agencies consider including a data field to indicate when another agency assumes enforcement responsibility on a stop initiated by one of their officers (Pryor et al., 2020).

Table 3.3: Comparison of Number of Stops in CDR and CAD Data Sets, 2021 (p.1 of 3)

	Traffic Stops in CDR	Traffic Stops in CAD	Percent Error
PSP Dept.	**439,104**	**499,443**	**-12.1%**
AREA I			
Troop B			
Belle Vernon	4,413	5,811	-24.1%
Pittsburgh	5,076	6,341	-20.0%
Uniontown	8,395	10,875	-22.8%
Washington	3,552	5,341	-33.5%
Waynesburg	3,711	4,742	-21.7%
Troop C			
Clarion	3,763	3,683	2.2%
Clearfield	3,991	4,573	-12.7%
Dubois	3,842	4,431	-13.3%
Lewis Run	5,515	5,256	4.9%
Marienville	2,527	3,144	-19.6%
Punxsutawney	4,095	4,478	-8.6%
Ridgway	3,441	3,609	-4.7%
Troop D			
Beaver	3,171	4,072	-22.1%
Butler	5,383	7,448	-27.7%
Kittanning	5,897	6,874	-14.2%
Mercer	2,850	3,329	-14.4%
New Castle	3,051	3,626	-15.9%
Troop E			
Corry	3,118	3,225	-3.3%
Erie	6,476	8,609	-24.8%
Franklin	2,657	2,977	-10.8%
Girard	6,421	7,294	-12.0%
Meadville	3,390	3,924	-13.6%
Warren	2,118	2,285	-7.3%
AREA II			
Troop A			
Ebensburg	2,014	2,131	-5.5%
Greensburg	5,426	5,816	-6.7%
Indiana	6,832	8,338	-18.1%
Kiski Valley	2,165	2,294	-5.6%
Somerset (A)	2,148	2,584	-16.9%
Troop G			
Bedford	4,744	5,793	-18.1%
Hollidaysburg	4,738	5,354	-11.5%
Huntingdon	4,484	5,305	-15.5%
Lewistown	4,433	5,045	-12.1%
McConnellsburg	3,058	3,213	-4.8%
Rockview	8,092	8,520	-5.0%

Table 3.3: Comparison of Number of Stops in CDR and CAD Data Sets, 2021 (p.2 of 3)

	Traffic Stops in CDR	Traffic Stops in CAD	Percent Error
Troop H			
Carlisle	12,065	12,534	-3.7%
Chambersburg	10,989	12,095	-9.1%
Gettysburg	9,444	12,131	-22.2%
Harrisburg	7,854	9,221	-14.8%
Lykens	2,589	3,190	-18.8%
Newport	3,631	4,284	-15.2%
Troop T			
Bowmansville	2,776	3,759	-26.2%
Everett	6,115	5,311	15.1%
Gibsonia	4,738	5,354	-11.5%
Highspire	0	205	-100.0%
King of Prussia	4,489	4,425	1.5%
New Stanton	5,505	5,927	-7.1%
Newville	3,992	4,609	-13.4%
Pocono	4,357	4,955	-12.1%
Somerset (T)	5,391	5,621	-4.1%
AREA III			
Troop F			
Coudersport	4,238	4,040	4.9%
Emporium	1,605	1,144	40.3%
Lamar	7,363	8,183	-10.0%
Mansfield	2,846	3,300	-13.8%
Milton	7,983	8,534	-6.5%
Montoursville	5,310	6,829	-22.2%
Selinsgrove	4,702	4,501	4.5%
Stonington	2,363	2,963	-20.3%
Troop N			
Bloomsburg	3,282	3,923	-16.3%
Fern Ridge	7,750	8,015	-3.3%
Hazleton	7,704	8,488	-9.2%
Lehighton	2,880	3,492	-17.5%
Stroudsburg	14,224	14,870	-4.3%
Troop P			
Laporte	1,823	1,444	26.3%
Shickshinny	1,586	1,857	-14.6%
Towanda	4,248	4,145	2.5%
Tunkhannock	1,777	1,737	2.3%
Wilkes-Barre	4,239	5,721	-25.9%

Table 3.3: Comparison of Number of Stops in CDR and CAD Data Sets, 2021 (p.3 of 3)

	Traffic Stops in CDR	Traffic Stops in CAD	Percent Error
Troop R			
Blooming Grove	5,716	6,966	-17.9%
Dunmore	3,910	4,253	-8.1%
Gibson	4,960	5,516	-10.1%
Honesdale	2,273	2,041	11.4%
AREA IV			
Troop J			
Avondale	7,380	9,083	-18.8%
Embreeville	6,041	6,404	-5.7%
Lancaster	5,950	6,791	-12.4%
York	9,963	10,585	-5.9%
Troop K			
Media	13,500	15,290	-11.7%
Philadelphia	7,959	8,601	-7.5%
Skippack	6,358	6,359	-0.0%
Troop L			
Frackville	2,724	3,557	-23.4%
Hamburg	2,677	2,883	-7.2%
Jonestown	4,314	4,837	-10.8%
Reading	5,234	5,848	-10.5%
Schuylkill Haven	5,453	5,359	1.8%
Troop M			
Belfast	4,626	4,435	4.3%
Bethlehem	4,604	5,486	-16.1%
Dublin	3,399	4,179	-18.7%
Fogelsville	5,802	6,574	-11.7%
Trevose	4,698	5,455	-13.9%
Specialized Units[39]			
SHIELD	4,587	5,010	-8.4%
Canine	1,735	2,565	-32.4%

Conclusion

In a data collection effort of this size and scope, it is common for issues to arise regarding data integrity. The PSP is to be commended for its responsiveness to improving the reliability and validity of the CDR data as issues were identified throughout 2021. Unfortunately, the results of the UC team's data audit demonstrate that the 2021 data collected by the PSP do not meet the required data integrity standards. Based on the identified limitations, **substantive statistical analyses of these data are not possible.**

---

[39] SHIELD is the Safe Highways Initiative thru Effective Law Enforcement and Detection program and involves PSP members who are specially trained to interdict criminal activity occurring on major highways.

IV. SUMMARY

The PSP proactively planned and initiated one of the most comprehensive traffic stop data collection efforts in the country in 2020 – 2021. The PSP also reinforced the importance of data collection through internal department communication, training, and changes in related department policies. Unfortunately, as documented by the results of the data audit in Section 3, the 2021 data collected by the PSP have significant issues with reliability and validity that prohibit any substantive analyses of these data. Instead, the UC team determined that 2021 would effectively be considered a year-long pilot test, during which several revisions and improvements to the data collection process were implemented based on issues identified both internally and by the UC team. Similarly, new data fields were added to increase the explanatory power of the UC team's future statistical analyses examining post-stop outcomes. Statistical analyses of PSP traffic stop data will begin at the conclusion of the collection of 2022 data. This ensures that any future decisions about organizational changes to training, policy, and supervision are based on the highest quality data possible, in which both the UC team and the PSP can be confident.

Future Reports

In late 2021, the PSP began the process of extending the collection of police-citizen contact data during all member-initiated traffic stops from 2022 – 2024. As a result of this contract extension with our independent, external research team, we will continue to analyze these data with the goal of providing the PSP with specific policy and training recommendations to reduce any racial/ethnic disparities identified. For each year from 2022 to 2024, the UC team will produce three quarterly reports and one annual report. The quarterly reports are designed strictly as ongoing data audits, while the final annual reports will include substantive and detailed statistical analyses. That is, the results presented in quarterly reports will be purely descriptive in nature and are designed to give feedback to PSP administrators regarding the status of the data collection process in each quarter, along with exploring initial trends and patterns in the data that may be utilized for data collection improvement, supervisory or training purposes.

The first annual report, which will be provided to the PSP on March 31, 2023,[40] will contain statistical analyses of 12 months of traffic stop data (Jan 1, 2022 – Dec 31, 2022), including: (1) comparisons of traffic stop data to appropriate benchmarks (if available) at the county and municipality levels; (2) bivariate statistical analyses of race/ethnicity and post-stop outcomes at the Area, Troop, and Station levels, and (3) multivariate statistical analyses of post-stop outcomes (warnings, citations, arrests, searches, and seizures) at the department level that account for multiple explanatory factors, and outcome test analyses of searches and seizures. This comprehensive examination of PSP's data will allow the research team to provide the PSP with answers to the research questions of interest, including:

---

[40] Future annual reports based on data collected in 2023 and 2024 will follow a similar schedule. The Year 3 final comprehensive written report will include the same types of statistical analyses described herein but will include all 36 months of traffic stop data collected by the PSP (Jan 1, 2022 – Dec 31, 2024) and will allow for comparisons of traffic stop data over time to examine long-term patterns and trends.

1. Does the rate of stops experienced by persons of different racial/ethnic groups align with those groups' representation among persons at-risk of being stopped by the PSP?
2. What factors or combination of factors predict the post-stop law enforcement actions received by individuals stopped by the PSP?
3. Is driver race or ethnicity related to law enforcement actions after accounting for other relevant individual, legal, situational, and environmental factors?

Based on these findings, the research team will provide actionable recommendations to address any patterns of racial/ethnic disparities identified during interactions with the public and present additional opportunities for the use of supplementary data and studies to aid in the interpretation of statistical findings.

Given the variety of influential factors on police stop and enforcement decisions, it is beneficial for agencies to identify and better understand trends and patterns to enhance their ability to interact with the public safely and fairly. Note, however, that the traffic stop data collected by PSP Troopers cannot be used to determine whether they have individually or collectively engaged in "racial profiling," nor can it be used to assess the legality of prior or future individual traffic stops. The continued data collection and analyses of all PSP member-initiated traffic stops, however, will provide police executives with the necessary information to identify potentially problematic areas and refocus training, supervision, and policies accordingly. In addition, as data collection continues over time, patterns and trends in traffic stops and post-stop outcomes will be examined from year to year to determine if changes in policies and training to reduce possible racial/ethnic disparities have the desired impact. Further, the proactive collection and analysis of racial/ethnic disparities in traffic stop data demonstrates dedication to transparency and accountability to the public, is widely considered a best practice and continues PSP's commitment toward evidence-based policing practices.[41]

---

[41] Pryor et al., 2020.

V.   APPENDIX

Figure A.1 CDR Form – Page 1 of 2

Contact Data Reporting
9.00.00, Form 12/22/2021, Rule 12/21/2021

CDR Number	Related Document Number
CBWVTTGTS22	
CAD/Case Number	RMS location

STOP LOCATION

Location

County Name		Municipality Code	Municipality Name
Stop Time HRS	Stop Date	Roadway Type	Registration State

GPS

Latitude: | Degrees | Minutes | Seconds | Decimal | Longitude: | Degrees - | Minutes | Seconds | Decimal

REASON

Reason for Stop "select all that apply"

Prior Posted Speed Limit	Driver Speed MPH	Amount Over Limit MPH	
Special Traffic Enforcement	Dedicated Enforcement Team	Assigned Enforcement Team	

DRIVER

Zip Code	Gender	Date of Birth	Race	Ethnicity
LEP?	LEP Language Access?			

STOP RESULT

Stop Duration minutes	MCSAP Related?	Number of Passengers	K-9 Utilized ?

DRIVER

Driver's Behavior	Driver Arrested?	Number of Citations	Warning Type	Number of Warnings

Figure A.1 CDR Form – Page 2 of 2

PASSENGER

Asked for Passenger Identification?	Passenger Identification Type?	Passenger Identification Justification?

Passenger Race	Passenger Ethnicity

LEP?	LEP Language Access?

Number of Warnings	Number of Citations	Passenger Arrested?	

ADD PASSENGER GROUP

SEARCH

Search Initiated?	Searched? (Mark all the apply)

Search Reason(s) "Mark all that apply"

Property Seized "Mark all that apply"

EMPLOYEE INFORMATION

Location Code	Employee Number
Z99	00111222