

GEORGIA DOT RESEARCH PROJECT 14-28

FINAL REPORT

**CENTERLINE RUMBLE STRIPS SAFETY IMPACT
EVALUATION—PHASE 2**



**OFFICE OF PERFORMANCE-BASED
MANAGEMENT AND RESEARCH
15 KENNEDY DRIVE
FOREST PARK, GA 30297-2534**

GDOT Research Project 14-28

Final Report

CENTERLINE RUMBLE STRIPS SAFETY IMPACT EVALUATION—PHASE 2

By

Angshuman Guin, Ph.D., Principal Investigator
Michael O. Rodgers, Ph.D., co-Principal Investigator
Michael P. Hunter, Ph.D., co-Principal Investigator
Marisha Pena, Graduate Research Assistant
Jerome Sin, Graduate Research Assistant
Dong-yeon Lee, Graduate Research Assistant

School of Civil and Environmental Engineering
Georgia Institute of Technology

Contract with

Georgia Department of Transportation

In cooperation with

U.S. Department of Transportation
Federal Highway Administration

February 2018

The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Georgia Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation

Technical Report Documentation Page

1. Report No.: FHWA-GA-18-1428		2. Government Accession No.:		3. Recipient's Catalog No.:	
4. Title and Subtitle: Centerline Rumble Strips Safety Impact Evaluation—Phase 2				5. Report Date: February 2018	
				6. Performing Organization Code:	
7. Author(s): Angshuman Guin, Ph.D.; Michael O. Rodgers, Ph.D.; Michael P. Hunter, Ph.D.; Marisha Pena; Jerome Sin; Dong-yeon Lee				8. Performing Organization Report No. 14-28	
9. Performing Organization Name and Address: Georgia Tech Research Corporation School of Civil and Environ. Engineering 790 Atlantic Drive Atlanta, GA 30332-0355				10. Work Unit No.:	
				11. Contract or Grant No.: 0013528	
12. Sponsoring Organization Name and Address: Georgia Department of Transportation Office of Performance-Based Management & Research 15 Kennedy Drive Forest Park, GA 30297-2534				13. Type of Report & Period Covered: Final; June 2014 – January 2018	
				14. Sponsoring Agency Code:	
15. Supplementary Notes: Prepared in cooperation with the US Department of Transportation, Federal Highway Administration.					
16. Abstract: Centerline rumble strips are used by various states as a low-cost countermeasure for mitigating cross-over crashes on two-way highways. This study performs a safety impact evaluation using an empirical Bayesian analysis. The researchers obtained a crash modification factor of 0.58, indicating a 42% reduction in crashes involving centerline crossings associated with the installation of centerline rumble strips. The sample size of fatal and injury crashes was too small to obtain separate crash modification factors for fatal crashes and injury crashes. The favorable crash modification factor (0.58) found in this study supports wider use of centerline rumble strips as a safety measure to address crashes involving vehicles that cross the centerline of the roadway.					
17. Key Words Centerline rumble strips, safety, empirical Bayesian				18. Distribution Statement	
19. Security Class (This Report) Unclassified		20. Security Class (This Page) Unclassified		21. No of Pages 58	22. Price

Table of Contents

List of Tables.....	ix
List of Figures.....	xi
Executive Summary.....	xiii
Acknowledgements	xv
1 Introduction.....	1
1.1 Overview of Project	1
1.2 Project Objectives	3
2 Literature Review.....	5
2.1 Observational Before–After Studies	6
2.1.1 Naïve Before–After	7
2.1.2 Full Bayes.....	8
2.1.3 Empirical Bayes.....	9
2.1.4 Comparison of EB and FB	9
2.2 Evaluation of Safety Treatment with an Empirical Bayes Approach	11
3 Methodology.....	13
3.1 Site Selection	13
3.1.1 Georgia Project Database	13
3.1.2 Additional Sources for Authentication of Sites.....	15
3.1.3 Final Study Sites.....	19
3.1.4 Analysis Period.....	22
3.2 Crash Database	26
3.2.1 Treatment and Reference Crash Databases	26
3.2.2 Crash Database Verification.....	28

3.3 Empirical Bayes Method/Development of SPF	30
3.3.1 Before-Period SPF Parameters	30
3.3.2 After-Period SPF Parameters	33
3.3.3 Determination of Crash Modification Factor	34
4 Results.....	37
4.1 Crash Statistics	37
4.1.1 Total Target Crashes.....	37
4.1.2 Analysis of Crash Severities and Types	38
4.1.3 Naïve Analysis of Crashes by Collision Type.....	39
4.2 Empirical Bayes Method/Development of SPF	43
4.2.1 Before-Period SPF Parameters	44
4.2.2 After-Period SPF Parameters	47
4.2.3 Determination of CMF	50
4.3 Misclassified Crashes	51
5 Conclusions and Recommendations	53
6 References.....	55
Appendix	A-1

LIST OF TABLES

Table 1. Results Obtained from TransPI (Sin 2014).....	14
Table 2. TransPI Location Description by Installation Site (Sin 2014).....	15
Table 3. Project Plan Sheet Information for Project ID 0007077 (Sin 2014).....	17
Table 4. List of Study Sites.....	20
Table 5. CLRS Start and End Mileposts for Study Sites	23
Table 6. Begin and End Dates for CLRS Construction	25
Table 7. Filters Used to Create the Comparison Crashes Database.....	27
Table 8. Total Crashes	37
Table 9. Site-by-Site Comparison, All Crash Types.....	38
Table 10. Number of Individuals Injured	39
Table 11. Crash by Collision Type	40
Table 12. Site-by-Site Comparison by Collision Type.....	41
Table 13. Comparison of Head-on Crashes	42
Table 14. Comparison of Opposite-Direction Sideswipe Crashes.....	42
Table 15. Comparison of Not-a-Collision-with-a-Motor-Vehicle Crashes	43

Table 16. Crash Statistics.....	43
Table 17. Predicted Crash Frequency in Before Period.....	47
Table 18. Predicted Crash Frequency in After Period.....	50
Table 19. Misclassified Target Crashes.....	52

LIST OF FIGURES

Figure 1. Google Street View Verification of Centerline Rumble Strips (Sin 2014).....	16
Figure 2. Preconstruction Status Report for Project ID 0007077 (Sin 2014).....	17
Figure 3. Map of Project ID 0007077 Location from Project Plan Sheet.....	19
Figure 4. Locations of CLRS Sites (Sin 2014).....	21
Figure 5. Target Crash Exceptions.....	29
Figure 6. Percent Change of Crashes by Severity in Before versus After Period	39
Figure 7. Percent Change in Crashes by Type.....	40
Figure 8. BEFORE: $\ln(\text{crashes per VMT})$ versus $\ln(\text{AADT})$	45
Figure 9. AFTER: $\ln(\text{crashes per VMT})$ versus $\ln(\text{AADT})$	48

EXECUTIVE SUMMARY

Centerline rumble strips (CLRS) are often used as a low-cost countermeasure for reducing the frequency of cross-over crashes on two-way highways. Rumble strips are designed to provide motorists with an audible and tactile warning that they are either approaching a critical safety-related decision point or that their motor vehicles have partially or completely left the travel lane. Centerline rumble strips are used to reduce head-on crashes, opposite-direction sideswipe crashes, and, to some extent, single-vehicle-run-off-the-road-to-the-left crashes.

This study quantifies the effectiveness of the countermeasure through analysis of changes in observed cross-over crash rates where CLRS have been implemented in Georgia. While estimates of the safety benefits are available from other states implementing centerline rumble strips (Persaud et al. 2004, Outcalt 2001, Hallmark et al. 2009, Fontaine et al. 2009, Torbic et al. 2009), this study evaluates CLRS safety impacts in the context of driving and roadway characteristics within the state of Georgia.

In the first phase of the project, researchers performed a preliminary analysis of the crash data from several CLRS locations in the state of Georgia, noted inconsistencies in the availability and accuracy of location information, and developed a quality assurance procedure involving cross-checking the crash database with written police records. The current second-phase study used this same validation procedure to perform an empirical Bayesian (EB) analysis for evaluation of the safety impact of CLRS using a seasonally adjusted 24-month pre-deployment (before) period and a 24-month post-deployment (after) period for comparison. The EB analysis resulted in a crash modification factor of

0.58 for the CLRS treatment, indicating a 42% reduction in crashes associated with conditions that CLRS were designed to address (i.e., crashes involving crossing the centerline). However, the sample size was too small to obtain separate crash modification factors for fatal crashes and injury crashes.

The quality assurance procedure was the most resource-intensive part of the effort. Researchers manually checked the base crash data against the crash description recorded by the investigating police officer to verify crash type and to obtain a clearer indication of whether the crash could have been impacted by the presence of CLRS. This step was critical to improving the reliability of the CMF value, as it reduced crash misclassifications.

The involvement of multiple agencies in the recording of the crash data naturally introduces variability and non-uniformity in the crash data. These differences become critical when the results of safety evaluations are dependent on the correct categorization of the incidents and the correct association of the incidents to a safety measure. A broad methodological recommendation from the lessons learned in the study is to employ sufficient crash-verification procedures in any safety study that develops a crash modification factor, especially in cases where the sample size of the crashes is small, or if crash modification factors are desired for specific crash categories.

The favorable crash modification factor (0.58) obtained in this study clearly provides sufficient justification for the use of CLRS as a low-cost safety countermeasure to address crashes involving vehicles that cross the centerline of the roadway.

ACKNOWLEDGEMENTS

The authors thank the Georgia Department of Transportation (GDOT), in cooperation with the U.S. Department of Transportation (USDOT) Federal Highway Administration (FHWA), for support of this research under Research Project 14-28. The authors also thank David Adams, GDOT Office of Traffic Operations, for his support throughout the study.

1 Introduction

The objective of this research is to provide the Georgia Department of Transportation (GDOT) with an evaluation of the safety impacts of adding centerline rumble strips to an undivided highway facility.

Centerline rumble strips (CLRS) are often used as a low-cost countermeasure for reducing the frequency of cross-over crashes on two-way highways. The current study quantifies the effectiveness of the countermeasure through analysis of changes in observed cross-over crash rates where CLRS have been implemented in Georgia. While estimates of the safety benefits are available from other states implementing centerline rumble strips (Persaud et al. 2004, Outcalt 2001, Hallmark et al. 2009, Fontaine et al. 2009, Torbic et al. 2009), this study evaluates CLRS safety impacts in the context of driving and roadway characteristics within the state of Georgia.

1.1 Overview of Project

Rumble strips are designed to provide motorists with an audible and tactile warning that they are either approaching a critical safety-related decision point or that their motor vehicles have partially or completely left the travel lane. There are typically four types of rumble strip applications based on the placement of the rumble strips with reference to the travel path:

- Shoulder rumble strips
- Centerline rumble strips

- Mid-lane rumble strips
- Transverse rumble strips

Shoulder rumble strips are the most common application and have been used widely to address single-vehicle-run-off-the-road crashes. Similarly, centerline rumble strips are used to reduce head-on crashes, opposite-direction sideswipe crashes, and, to some extent, single-vehicle-run-off-the-road-to-the-left crashes.

In Phase 1 of the project (RP 12-12), the researchers conducted a nationwide survey on potential maintenance issues related to CLRS. This survey was performed in response to the perception that CLRS deployments are associated with increased maintenance requirements. The result of this survey indicated that most of the observed maintenance problems were associated with improper construction rather than with the CLRS themselves.

The Phase 1 project included a literature review on the safety impacts of centerline rumble strips. Most of the studies identified in this review were limited to specific roadway and/or crash types, and none were fully applicable to Georgia conditions. In RP 12-12 the researchers also conducted a preliminary analysis of crash data from nine Georgia CLRS locations. This preliminary analysis found both limited availability of certain crash data and identified inconsistencies in crash location information within the Georgia crash database. A methodology for investigating and mitigating biases related to these errors was developed that involves a manual review of sample sets of police crash reports and validation of the information in the crash database and provides the basis for the analysis used in this Phase 2 study.

1.2 Project Objectives

The overarching objective of this research is to provide the Georgia Department of Transportation with an evaluation of the safety impacts of adding a centerline rumble strip to an undivided highway facility. Specifically, the project aims to evaluate whether there is any decrease in the number of crashes or any change in the type or severity of crashes after installation of centerline rumble strips on highway facilities in Georgia.

To meet these objectives, the research team performed an updated literature review on CLRS safety. Georgia CLRS sites and corresponding control sites were identified, and associated roadway characteristics and crash data were obtained and checked for accuracy. During Phase 1, researchers observed issues regarding the completeness of the information in the crash records; in Phase 2 significant improvements in the overall quality of the crash data were noted relative to that of the preliminary study. Key sub-objectives of Phase 2 may be summarized as follows:

- Validate crash data for the chosen CLRS sites and sample control sites using police records and supplement with additional information as necessary, using the methods developed in Phase 1.
- Perform a before–after study using an EB analysis to evaluate the impact of a CLRS installation on those crash rates that CLRS are designed to mitigate.

2 Literature Review

With over 150 miles of centerline rumble strips on roadways throughout the state, Georgia has joined the ranks of states that use CLRS as a countermeasure to cross-centerline crashes, including head-on and opposite-direction sideswipe collisions. Many factors can lead to the aforementioned crash types, the most common being inattentive or drowsy drivers, which account for 86% of fatal head-on crashes on two-lane highways (Alexander and Gardner 1995). When coupled with rural roadway conditions, including higher traffic speeds, lower rates of seatbelt use, and longer emergency-response times, safety countermeasures such as CLRS become increasingly attractive. Though CLRS may be constructed in several forms, the majority of installations are of the milled-in type, which is cost effective and can be readily implemented on existing roadways. Alternatively, CLRS can be constructed directly on the centerline, extended into the travel lane, or on either side of the centerline pavement markings. CLRS may have the added benefits of improving safety in low-visibility driving conditions, especially in areas with wintry weather or when roadway markings are obscured.

A detailed literature review regarding the properties of rumble strips, benefits of CLRS, and concerns related to adverse impacts of CLRS is available in the Phase 1 (GDOT RP 12-12) project report (Guin et al. 2014). The following literature review focuses on the most common methodology for safety evaluation studies: observational before/after studies, which include naïve before and after, empirical Bayes, and full Bayes (FB).

2.1 Observational Before–After Studies

To evaluate the success of any roadway safety improvement program, it is essential to review the change in the number of motor vehicle crashes and the number of injuries and fatalities. At a minimum, a roadway safety project evaluation should include performance measures from both before and after the installation of a roadway treatment or other changes to the roadway (Herbel et al. 2010). Such a study of the effects of a roadway treatment should consider “what would have been the safety level” in the after period without treatment compared to the safety level with the treatment (Hauer 1997). The effect of the treatment is represented by the difference in the number of injuries or the crash rate, over time, relative to the after period with and without the treatment (Herbel et al. 2010).

The challenge in this type of comparison lies in estimating “what would have been” with no treatment. Natural variations in crash data and changes in site conditions produce limitations with any such estimates. Changes over time (e.g., weather, traffic patterns, physical changes to the site conditions, etc.) create fluctuations in crash data. These limitations generate bias and reduce the reliability of a comparative analysis.

These limitations potentially bias the after-period-without-treatment estimation, which uses the crash data from the before period. For instance, the before-period crash experience is likely the motivation behind the treatment site selection, and, thus, the after-period prediction is subject to a selection bias. That is, the treatment site is not a random selection but selected likely due to the observed crash rates.

Since crash rates can vary significantly from year to year, any estimates derived from these data are sensitive to a bias known as regression-to-the-mean (RTM). RTM is inherent in crash data. According to the Highway Safety Manual (HSM) (AASHTO 2010),

regression-to-the-mean is “the tendency for the occurrences of crashes at a particular site to fluctuate up or down, over the long term, and to converge to a long-term average. This tendency introduces regression-to-the-mean bias into crash estimation and analysis, making treatments at sites with extremely high crash frequency appear to be more effective than they truly are” (AASHTO 2010). RTM produces periods that may have a comparatively high or low crash frequency. Attributing a decline in crash frequency to a roadway treatment may be misleading because the overall trend of crash frequency may have already been in decline unrelated to the treatment. A proper comparative analysis effectively accounts for the RTM bias.

Observational before/after studies consist of three methods: naïve before–after, empirical Bayes, and full Bayes. A naïve before–after analysis is based on the assumption that nothing changed in the after period except the treatment in question. Therefore, the before-period crashes are used to predict what the after-period crashes would have been without treatment (Hauer 1997). The Bayesian methods (full and empirical) combine before-period data with the after-period data to develop the expected safety of a treatment (Persaud et al. 2010). Empirical Bayes and full Bayes are not different types of studies; they are simply two related approaches to combining prior and current information.

2.1.1 Naïve Before–After

In transportation safety, a naïve before–after study is one way (albeit not the most accurate way) to estimate the change between a parameter, such as crash frequency, during a before and after period. The naïve before–after study assumes that the passing of time has no effect on the after period and that the expected after crash rate without treatment would be the same as in the before period. However, the change in safety level can be attributed to

several factors in addition to the roadway treatment, such as weather, traffic patterns, driver behavior, driver inclination to report crashes, and RTM. All these other factors are assumed to be unchanged in a naïve before–after study, and any change in safety is assumed to be caused by the treatment only (Hauer 1997, Herbel et al. 2010). This assumption in a naïve before–after study is weak and usually unrealistic. In addition, this method is strongly influenced by the selection bias discussed previously. This method does not allow researchers to separate out crash rate change attributable to the treatment from the other factors that have also changed over the period of the study. Any conclusion from this study lessens a researcher’s ability to conclusively attribute the measured difference to the treatment of interest. This approach is generally not recommended for safety studies (Hauer 1997).

2.1.2 Full Bayes

Full Bayesian uses before data to predict future crashes at a treatment site had the treatment not been implemented. However, instead of a single-point estimate of the expected mean and its variance, it predicts a distribution of likely values. The estimate for expected crash frequency in the after period is determined by combining the distribution of likely values with the crash frequency of the specific study sites. The use of a distribution of likely values generally improves the overall estimate of likely crash rates (Persaud et al. 2010). As the researchers did not use the FB approach in this study, a detailed description is not provided; however, the next sections discuss the reasoning underlying the selection of empirical Bayes over full Bayes.

2.1.3 Empirical Bayes

The empirical Bayes method is a simplified version of the full Bayesian method and is well established and accepted by transportation professionals and researchers for roadway safety comparative studies (Carriquiry and Pawlovich 2004, Persaud et al. 2010). Through an EB study, the safety effectiveness of a treatment can then be based on the model rather than the raw crash data (Outcalt 2001). According to the HSM, the empirical Bayes approach combines “observed crash frequency data for a given site with predicted crash frequency data from many similar sites to estimate its expected crash frequency” (AASHTO 2010). The before-period data come from the evaluation sites and a reference group with similar roadway attributes to develop a calibrated safety performance function (Hauer 1997, AASHTO 2010, Persaud et al. 2010). The safety performance function (SPF) is an equation that represents the relationship between the expected number of target crashes and the roadway characteristics (Persaud et al. 2010). The expected crash frequency estimates are combined with “the site-specific crash count to obtain an improved estimate of a site’s long-term expected crash frequency” (Persaud et al. 2010). The EB approach uses the assumption that crashes follow a negative binomial (NB) distribution, and it employs the NB dispersion parameter in the estimation process (Persaud et al. 2010). Section 3.3 of this report provides a detailed walk-through of the EB method.

2.1.4 Comparison of EB and FB

Both EB and FB methods recognize that deriving estimates from just a few years of information from specific sites provides unreliable estimates. To remedy this, central to Bayesian analysis, comparison-group data for the same study period are used to

complement the treatment site's data. This addition of comparison-group data allows the analysis to formulate more robust estimates and account for RTM bias and traffic volume changes due to various factors, such as general trends, changes in crash reporting, weather, driver behavior, etc. (Carriquiry and Pawlovich 2004, Gross et al. 2010).

While both empirical Bayesian and full Bayesian approaches are suitable and effective methods for conducting a comparative analysis for traffic safety studies, their differences and comparative advantages render them most efficient in different scenarios of study (Persaud et al. 2010; Carriquiry and Pawlovich 2004). The FB approach is more complex than the EB approach, and some researchers believe that it more suitably accounts for uncertainty within crash data (Persaud et al. 2010). The EB approach simplifies the FB approach by assuming the study sites and the comparison sites have similar covariables, such as roadway geometry and signal control. These covariables are accounted for through the SPF derived in the EB method (Carriquiry and Pawlovich 2004). Furthermore, the FB approach requires substantially smaller dataset than the EB approach for the untreated control group sites. The FB approach “provides more detailed causal inferences” (Persaud et al. 2010) and “more flexibility in selecting crash count distributions” (Persaud et al. 2010), and it does not require the development of safety performance functions to obtain the predicted number of crashes.

However, the FB method is more cumbersome than the EB method. A high level of statistical knowledge is required to carry out the complexity of the full Bayes method. Additionally, it has been more difficult to develop statistical software for an FB application than for an EB application (Gross et al. 2010). Finally, research has shown that the EB approach produces similar results to the FB method and reliable analysis when an adequate

number of sites exists (Gross et al. 2010, Persaud et al. 2010). Thus, the research team in this effort chose EB for the safety analysis.

2.2 Evaluation of Safety Treatment with an Empirical Bayes Approach

In comparison to the naïve approach, the EB methodology enables a more precise estimation of the number of crashes that would have occurred at an individual treatment site in the after period had the treatment not been implemented (Harwood 1993). The EB method has been used in various roadway safety analyses (Persaud et al. 2010, Persaud et al. 2001, Kay et al. 2015, Porter et al. 2004, Sayed et al. 2010). These studies include diverse locations throughout the United States and Canada and varied roadway treatments, such as road diets, conversions of intersections to roundabouts, and shoulder rumble strips (SRS) and CLRS (Karkle et al. 2013, Kay et al. 2015, Porter et al. 2004, Persaud et al. 2001, Persaud et al. 2004, Persaud et al. 2010, Sayed et al. 2010).

For example, Persaud et al. (2001) conducted a before–after study using the EB procedure for the conversion of intersections to roundabouts. Their study estimated highly significant reductions of 40% for all crash severities combined and 80% for all injury crashes. Specifically, the crashes with fatalities and incapacitating injuries were reduced by 90%. In a later study, Persaud et al. (2010) used the EB and FB methods to examine the safety impacts of a road diet, which involved the conversion of four-lane roadways into three-lane roadways with a two-way left-turn lane in the middle. That study determined the estimated safety effects from both methods to be comparable.

Directly relevant to the current study, the effectiveness of CLRS has also been studied in various locations using EB. One study looked at the effects of 98 treatment sites of CLRS along approximately 210 miles of rural, two-lane roadways in seven states,

California, Colorado, Delaware, Maryland, Minnesota, Oregon, and Washington (Persaud et al. 2004). This analysis revealed that head-on and sideswipe accidents from the opposing direction experienced the most significant reduction, decreasing by 25%. In general, all crash types were reduced by 12%. From this analysis, the authors recommended a wider application of CLRS on rural, two-lane roads. Another study applied the EB method to evaluate the effectiveness of CLRS on undivided, rural two-lane arterials and divided, rural four-lane freeways in British Columbia (Sayed et al. 2010). The authors found that, overall, SRS and CLRS can significantly reduce severe collisions and specific collision types. The use of CLRS and SRS demonstrated a reduction of 18.0% of injuries. Individually, SRS reduced off-road right collisions by a statistically significant 22.5%, and CLRS showed a statistically significant reduction of 29.3% in off-road left and head-on collisions. Specifically, installing both CLRS and SRS on undivided, rural two-lane arterials reduced off-road right, off-road left, and head-on collisions combined by 21.4%. The authors concluded that rumble strips, whether just SRS or CLRS, or the combination of SRS and CLRS, are very effective safety measures to reduce the severity of crashes.

A recent study in Michigan assessed the safety impacts of a statewide CLRS implementation program carried out between 2008 and 2010. Using the EB method, the effectiveness of more than 4200 miles of CLRS installed along two-lane highways was assessed. Overall, CLRS were found to reduce target crossover collisions by 27.3% when used alone and by 32.8% when used in conjunction with SRS (Kay et al. 2015).

3 Methodology

3.1 Site Selection

In the Safety Action Plan, GDOT set out to develop 100 miles of CLRS in FY 2005. Using 2000–2003 crash data from the Accident Information System (AIS) database, locations with higher frequencies of centerline crossover crashes were identified as potential sites. The Office of Traffic Operations, in coordination with the Office of Maintenance, scheduled projects for CLRS installation in both stand-alone applications and in conjunction with resurfacing projects. Between the fall of 2005 and spring of 2006, GDOT carried out several CLRS installation projects. By 2006, there were nearly 200 miles of CLRS installed, primarily on rural two-lane, two-way roadways (Sin 2014).

3.1.1 Georgia Project Database

The CLRS sites for this study were selected from GDOT’s Transportation Project Information (TransPI) website in 2013. TransPI, now known as GeoPI (GDOT 2017), is a web-based database from which the public can access any related data or documentation for GDOT projects. Project managers and engineers submit project information, including documentation, financial information, and Geographical Information System views, into the TransPI/GeoPI system. That information is accessible to users both inside and outside of GDOT (Sin 2014).

For this study, an initial query for projects involving the installation of CLRS resulted in the eight projects listed in Table 1. Although there were only eight projects, several of those involved more than one installation site, such as the project on SR 36 that

had segments from SR 74 to SR 7 and also from SR 7 to I-75. After examining the 8 projects, the research team compiled at least 11 potential CLRS sites, which are listed in Table 2 with their corresponding beginning and ending descriptions (Sin 2014).

Table 1. Results Obtained from TransPI (Sin 2014)

Project ID	Project Accounting No.	Project Title	Counties
0006080	—	SR 25 SPUR EAST FM CR 583/SEA ISLAND DR TO E OF SR 25/US 17	Glynn
0006693	CSSTP-0006-00(693)	SR 14 SR 16 SR 154@SEV LOC IN CARROLL &COWETA [CENTERLINE]	Carroll, Coweta
0006945	CSSTP-0006-00(945)	SR 369 FM CHEROKEE CO TO HALL CO – CENTERLINE RUMBLE STRIPS	Forsyth
0006975	CSSTP-0006-00(975)	SR 42@SEV LOC IN HENRY BUTTS MONROE – CENTERLINE RUMBLE STRIPS	Butts, Henry, Monroe
0006976	CSSTP-0006-00(976)	SR 204 FM BRYAN COUNTY LINE TO I-95 – CENTERLINE RUMBLE STRIPS	Chatham
0007077	CSSTP-0007-00(077)	SR 36 FM SR 74 TO SR 7 & SR 36 FM SR 7 TO I-75	Butts, Lamar, Upson
0007079	CSSTP-0007-00(079)	SR 136 FROM SR 61/US 411 TO DAWSON COUNTY LINE	Gilmer, Gordon, Murray, Pickens
0007080	CSSTP-0007-00(080)	SR 26 FM E OF BULL RIVER BRIDGE TO TYBEE ISLAND CITY LIMITS	Chatham

Table 2. TransPI Location Description by Installation Site (Sin 2014)

Project ID	Centerline Rumble Strips Installation Site	Beginning Description	Ending Description
0006080	State Route 25 Spur	Sea Island Drive/ CR 583	State Route 25/US 17
0006693	State Route 14	Herring Road/CR 43	Johnston Circle/CR 7
0006693	State Route 16	Carrolton Bypass	Newnan Bypass
0006693	State Route 154	State Route 54	I-85
0006945	State Route 369	Forsyth County	Forsyth County
0006975	State Route 42	Several Locations in Henry, Butts, and Monroe Counties	Several Locations in Henry, Butts, and Monroe Counties
0006976	State Route 204	Bryan County Line	I-95
0007077	State Route 36	East Main Street	Peach Blossom Trail
0007077	State Route 36	Highway 41	I-75
0007079	State Route 136	State Route 61/US 411	Dawson County Line
0007080	State Route 26	East of Bull River Bridge	Tybee Island City Limits

3.1.2 Additional Sources for Authentication of Sites

The query for “centerline rumble strips” in TransPI yielded multiple entries for a single project; consequently, each entry’s project information required examination. Several project descriptions revealed that some projects consisted of multiple installation sites, and thus, provided conflicting information. The researchers used other sources to authenticate the discrepancies and confirm the details of each study site. To confirm that these roadways did have CLRS, they used Google Maps Street View® to verify its existence at the sites

returned by the TransPI query, as shown in Figure 1 for Project ID 0007077. Project Preconstruction Status Reports and Project Plan Sheets were requested from GDOT to gather additional project information, including the total mileage and various dates associated with the project such as the Management Let Date and the Project Completion Date (Sin 2014). An example of a Project Preconstruction Status Report and information taken from the Project Plan Sheets for Project ID 0007077 are shown in Figure 2 and Table 3.



Figure 1. Google Street View Verification of Centerline Rumble Strips (Sin 2014)

PRECONSTRUCTION STATUS REPORT						
PROJ ID	COUNTY	DESCRIPTION				
0006080	Glynn	SR 25 SPUR EAST FM CR 583/SEA ISLAND DR TO E OF SR 25/US 17				
This project consists of installing ground-in-place centerline indentation rumble strips centered about the centerline traffic stripe of State Route 25 Spur East from Sea Island Drive to east of State Route 25/US 17 in Glynn County (MP 6.15 to MP 10.53).						
PROJ NO.:	CSSTP-0006-00(080)	SPONSOR :	GDOT	Phase	Fiscal Year Approved	Approved FY Estimate
MPO TIP #:	BATS 06-01	PROJ MGR:	Cameron, Derrick	Engineering	2006	\$ 5,000.00
MPO:	Brunswick	DOT DIST:	5	*Inflation Included in Estimate		
PROJ LENGTH (MI):	3.93	CONG. DIST:	1			
PROGRAM TYPE:	Safety	House Dist :				
TYPE WORK:	Rumble Strips	Senate Dist :				
LET RESPONSIBILITY:	GDOT Let					
BIKE PROVISIONS INCLUDED?	N					

ACTIVITY	ACTUAL START	ACTUAL FINISH	PERCENT COMPLETE
Environmental Approval	1/5/2006	1/5/2006	100
FFPR Inspection	1/31/2006	1/31/2006	100
FFPR Inspection			0

<i>Right of Way Acquisition Information</i>			
Preliminary Parcel Count:	Total Parcel Count:	Acquired by:	N/R

Figure 2. Preconstruction Status Report for Project ID 0007077 (Sin 2014)

Table 3. Project Plan Sheet Information for Project ID 0007077 (Sin 2014)

Attribute	Description
Project Number	CSSTP-0007-00(077)
Project ID	0007077
Net Length	29.77
Starting Milepost	MP 8.12
Ending Milepost	MP 0.49
Starting County	Upson County
Ending County	Butts County

To determine the exact locations of CLRS along the roadways, the researchers examined maps from the Project Plan Sheets to verify the beginning and ending mileposts of some of the installation sites. However, they discovered that the maps in the Project Plan Sheets did not always match the descriptions found in the projects from the TransPI query. For example, the map in the Project Plan Sheets for Project ID 0007077, as seen in Figure 3, only showed one segment of CLRS installations, although the project actually had two sections. The two segments of CLRS were detailed in the project documents found in TransPI and verified in Google Maps and Street View. The Project Plan Sheets also listed a Detailed Quantities Estimate, which had values that were used to verify the existence of CLRS in the projects. While information from various sources was not always accurate, it served as reference for determining the correct locations of the CLRS (Sin 2014).



Figure 3. Map of Project ID 0007077 Location from Project Plan Sheet

3.1.3 Final Study Sites

After examination, 10 CLRS installation sites were chosen for the analysis. From the original query results listed in Table 1, Project ID 0007080 and Project ID 0006080 were listed as “cancelled” under the *Project Completion* category and eliminated from the list of potential sites. Project ID 0006975, SR 42, was considered as two separate sections for this analysis. The list of CLRS sites is provided in Table 4, and a map of their locations is presented as Figure 4.

Table 4. List of Study Sites

Project ID	Description
0006693	SR 14
0006693	SR 16
0006693	SR 154
0006945	SR 369
0006975	SR 42 Section A
0006975	SR 42 Section B
0006976	SR 204
0007077	SR 36 Section A
0007077	SR 36 Section B
0007079	SR 136

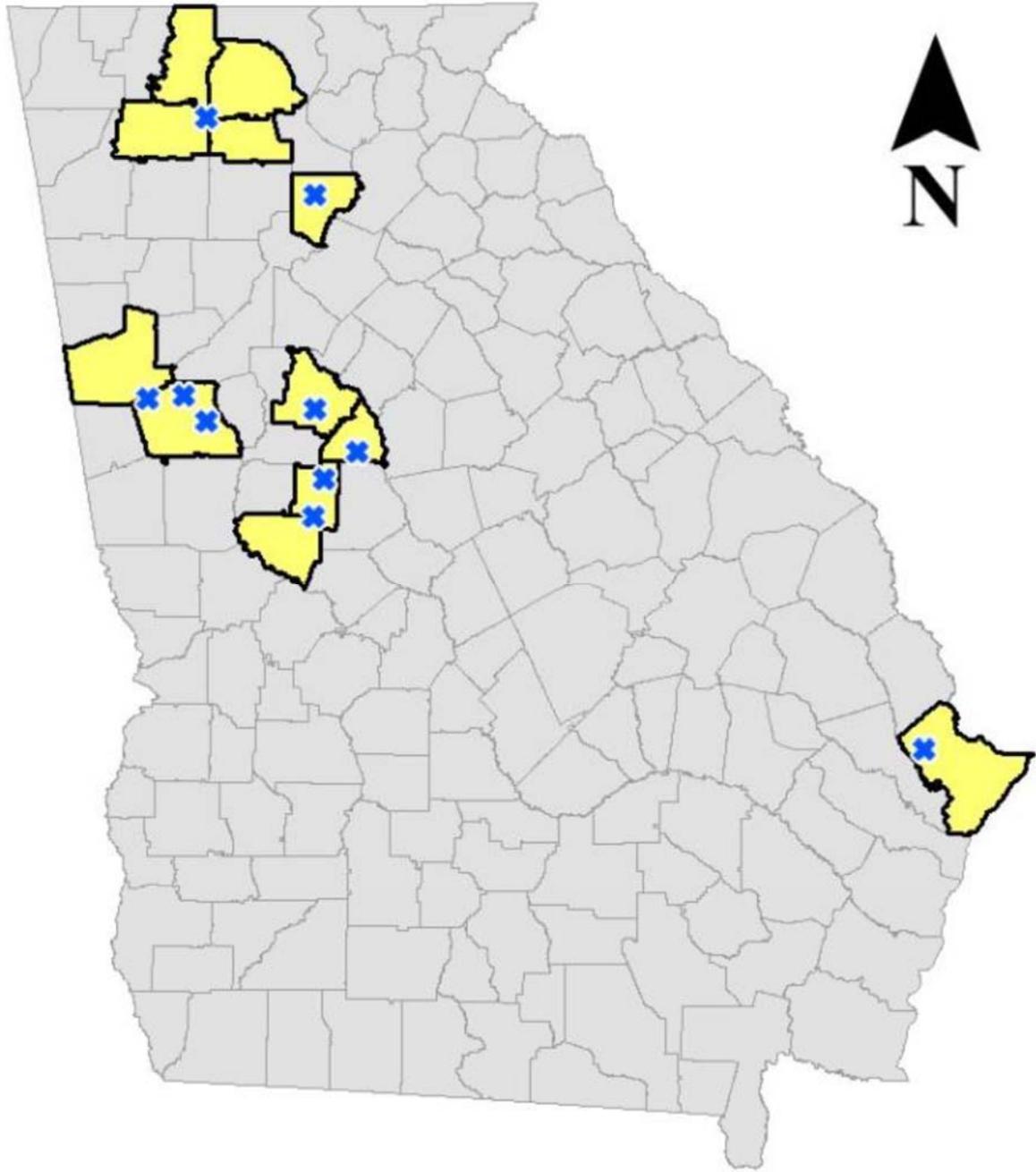


Figure 4. Locations of CLRS Sites (Sin 2014)

Once the CLRS segments were determined, the last step was to verify the exact locations of the start and end mileposts of the CLRS. Initially, the milepost information for each CLRS installation site was extracted from the Project Plan Sheets. However, after

Careful revision, most of the mileposts did not correspond with the mileposts noted in TransPI. To rectify these inaccuracies, the mileposts from the Project Plan Sheets were plotted in Google Earth® and verified using Google Street View®. Once the mileposts were confirmed, 126.46 miles along 10 routes were identified as the treatment sites.

Problems encountered during the automated association of crashes to treatment sites during the data reduction process led to a 13-mile decrease in the total number of miles of treatment sites studied. Table 5 shows the segments included in the preliminary sites. The segments that were not included in the final analysis are grayed out.

3.1.4 Analysis Period

To conduct an appropriate comparative analysis, the study periods must be determined to include time before the start of and after completion of CLRS installation on all study sites. The federal and TransPI reports had conflicting start and stop construction dates for each project. To clarify discrepancies, the construction completion dates were confirmed by GDOT engineers to be the *Time Charges Stop Date* from federal construction reports. The confirmed start and stop dates are listed in Table 6.

Table 5. CLRS Start and End Mileposts for Study Sites

Project ID/Roadway Description	County	Mileposts		Segment Length (mi.)	Total Study Site Length (mi.)
		Begin	End		
0006693/SR 14	Coweta	19.68	19.74	0.06	7.87
		19.74	23.17	3.43	
		23.17	26.72	3.55	
		26.72	27.55	0.83	
0006693/SR 16	Carroll	16.69	17.64	0.95	16.56 (18.24)
		17.64	22.65	5.01	
		22.65	26.19	3.54	
		26.19	27.87	1.68	
	Coweta	0.00	3.86	3.86	
		3.86	6.33	2.47	
		6.33	6.98	0.65	
		6.98	7.06	0.08	
0006693/SR 154	Coweta	0.11	0.56	0.45	0 (7.49)
		0.56	3.34	2.78	
		3.34	5.31	1.97	
		5.31	7.60	2.29	
0006945/SR 369	Forsyth	0.00	2.71	2.71	19.89
		2.71	5.80	3.09	
		5.80	6.43	0.63	
		6.43	10.07	3.64	
		10.07	11.08	1.01	
		11.08	11.86	0.78	
		11.86	12.82	0.96	
		12.82	19.89	7.07	
0006975/SR 42 A	Butts	0.00	3.18	3.18	7.68 (7.97)
		3.18	4.81	1.63	
		4.81	7.44	2.63	
		7.44	7.68	0.24	
		7.68	7.97	0.29	
0006975/SR 42 B	Henry	4.58	8.53	3.95	5.23
		8.53	9.81	1.28	
0006976/SR 204	Chatham	0.00	0.64	0.64	8.14
		0.64	8.14	7.50	

Table 5 (Cont.)

Project ID/Roadway Description	County	Mileposts		Segment Length (mi.)	Total Study Site Length (mi.)
0007077/SR 36 A	Upson	9.34	11.06	1.72	8.05 (13.87)
		11.06	15.72	4.66	
		15.72	19.11	3.39	
	Lamar	0.00	1.93	1.93	
		1.93	4.10	2.17	
0007077/SR 36 B	Lamar	7.21	13.51	6.30	11.84
		13.51	16.83	3.32	
		16.83	18.60	1.77	
		18.60	19.05	0.45	
0007079/SR 136	Gordon	23.56	24.07	0.51	28.25
	Murray	0.00	2.82	2.82	
	Gilmer	0.00	5.21	5.21	
	Pickens	0.00	3.67	3.67	
		3.67	6.32	2.65	
		6.32	7.25	0.93	
		7.25	12.01	4.76	
		12.01	14.14	2.13	
		14.14	17.96	3.82	
17.96	19.71	1.75			
				Total	113.51

Table 6. Begin and End Dates for CLRS Construction

CLRS Site	Start Date	Stop Date
SR 14	10/11/2005	10/31/2005
SR 16	10/11/2005	10/31/2005
SR 154	10/11/2005	10/31/2005
SR 369	03/06/2006	03/26/2006
SR 42 A	01/17/2006	05/31/2006
SR 42 B	01/17/2006	05/31/2006
SR 204	02/14/2006	02/28/2006
SR 36 A	01/17/2006	05/31/2006
SR 36 B	01/17/2006	05/31/2006
SR 136	01/17/2006	05/31/2006

Initially, the study periods were identified as two complete calendar years before (2003–2004) and two complete calendar years after (2007–2008) the construction of the CLRS sites. This study period would provide time to compensate for changes in driving patterns due to the unfamiliarity of the new roadway features (CLRS) or the presence of construction equipment and changes in roadways, such as closures or detours. However, for this study, police records corresponding to crash data along the CLRS sites and the control sites were available only for January 1, 2003, until May 31, 2008. This limited the use of data from the full 2008 calendar year. To maintain an analysis period that accounts for seasonal changes consistent with the before and after periods, the final dates for the comparative analysis were set to be:

- Before period: June 1, 2003, to May 31, 2005
- After period: June 1, 2006, to May 31, 2008

3.2 Crash Database

Investigating officers provide police reports documenting crash data involving motor vehicles, bicycles, and pedestrians. In Georgia, police agencies record motor vehicle crashes with the standardized Georgia Uniform Motor Vehicle Accident Report. The Georgia Department of Transportation (GDOT) and/or GDOT contractors (Police Report Archive) archive the images of these police reports. Additionally, GDOT and/or its contractors/collaborators extract data from the reports and retain the information in the GDOT Crash Database, a searchable database format to facilitate retrieval of important information for research and other purposes.

While both the Police Report Archive and GDOT Crash Database are public records, they contain sensitive, personally identifiable information that must be protected from inadvertent release. Researchers are granted access to both databases courtesy of GDOT based on approved data protection protocols. The protocols used in this study were originally developed by the Georgia Transportation Institute (GTI). To protect the anonymity of persons involved in crash reports, GTI requires that its research projects use sanitized versions of the databases with all sensitive, personally identifiable information redacted. When the research cannot be conducted using the sanitized databases, it requires approval from the Georgia Tech Institutional Review Board (IRB).

3.2.1 Treatment and Reference Crash Databases

The crash data were divided into two databases:

- Treatment crashes – crashes occurring along CLRS sites

- Reference crashes – crashes occurring along a control set of roadways with similar characteristics to those of the CLRS sites

Treatment (CLRS-associated) crash data were selected using a database query to filter crashes by the road characteristics and mileposts of the CLRS sites and study period dates. The *RCLink ID* in the crash database was used to associate road characteristics data with the crash data.

The reference crash data were chosen as all crashes along rural, undivided two-lane highways in Georgia from 2003 to 2008. Given the verification requirements for the crashes using the police records, a 25% random sampling was used to reduce the reference crash data. After sampling, 17,381 crashes were identified. Physical road characteristics were used to filter out irrelevant cases, as shown in Table 7. Additionally, crashes from treatment sites were excluded (Sin 2014). The final reference set consisted of 11,706 crashes.

Table 7. Filters Used to Create the Comparison Crashes Database

Variable	Filter
Accident Date	Same dates as the crashes in the Treatment Crashes database
Intersecting Road Type	Null
Dividing Highway Barrier Type	0 – No Barrier
Dividing Highway Median Type	0 – Undivided Road
Functional Classification	2 – Rural – Principal Arterial 6 – Rural – Minor Arterial 7 – Rural – Major Collector
Number of Left Lanes	1 – 1 lane on the left side of the roadway
Number of Right Lanes	1 – 1 lane on the right side of the roadway

The research team verified locations of each crash by comparing the results obtained from the crash database with the corresponding (sanitized) police report. The details of the sanitization process can be found in Appendix A. Because each police report was recorded by the investigating officer present at the time of the crash, the entries are subject to human error. The researchers assumed that the rate of errors was consistent throughout each study year. Each milepost reported in the police report was considered correct and used to determine if the crash was located in a treatment site in the verification process.

3.2.2 Crash Database Verification

Undergraduate research assistants were charged with verification of the crash database to identify target crashes needed for this research. In this case, target crashes are influenced by the presence of CLRS. CLRS are intended to prevent specific target crashes: those caused by vehicle centerline crossovers. This research identified target crashes as head-on collisions, opposite-direction sideswipe collisions, or collisions not with motor vehicles (Russell and Rys 2005, Sin 2014).

Additionally, the first few moments of these collision events likely involve crossing over the centerline due to inattentiveness, distraction, fatigue, or other conditions that are not intentional on the driver's part (Sin 2014). Therefore, target crashes exclude several centerline crossovers that occur due to other circumstances, such as vehicles that originally run off the right shoulder of the road and overcorrect and cross over the centerline, or crashes that occurred on locations that did not meet the two-lane, undivided requirement (e.g., at intersections or on three-lane or wider highways). A list of detailed exceptions is provided in Figure 5. Approximately, one-fifth (18.9%) of all target crashes (292 target

crashes) experienced some form of hydroplaning. Environmental conditions, such as water, rain, ice, and snow, and spilled fuel on the roadways that caused hydroplaning were considered as target crashes.

- Location outside of study scope
 - Intersections
 - On three-lane or wider highways
 - With separation or barriers between opposite directions of lanes
 - Two-way left-turn lanes
 - Raised medians
 - Turning lanes
- Overcorrection—vehicle first runs off to right
- Passing maneuvers
- Environmental/external factors

Figure 5. Target Crash Exceptions

Annual average daily traffic (AADT) and segment lengths for each crash analyzed in this study were identified from GDOT public data for 2003 to 2008. Due to the large number of crashes in the study set, a MatLab® code was written by the research team to extract this specific data for all target and non-target crashes by referencing the *RCLinkID* and location of crash as specified in the crash report.

3.3 Empirical Bayes Method/Development of SPF

The EB method was used to evaluate the effectiveness of the CLRS in preventing cross-over collisions. The safety performance functions for both before and after periods were derived to predict the number of expected crashes in the after period without the installation of CLRS. These estimates were based on all pre-installation (before period) crash data for the entire population of study segments, both treatment and reference sites. The basic input for this evaluation includes the number of collisions that occur on the study sites, and their respective AADT values and segment length. According to the Highway Safety Manual (AASHTO 2010), the SPF for predicted average crash frequency along rural two-lane, two-way roadway segments is given by Equation 1.

$$N_{spf\ rs} = e^{\alpha} AADT^{\beta} L \times 365 \times 10^{-6} \quad (1)$$

Where:

$N_{spf\ rs}$ = predicted total crash frequency for roadway segment base conditions;

$AADT$ = average annual daily traffic volume (vehicles per day); and

L = length of roadway segment (miles)

3.3.1 Before-Period SPF Parameters

The observed parameters for the SPF equation above were determined by a multistep process.

STEP 1: Select the before-period SPF

Based on the before period, the predicted number of crashes is found with the SPF, using Equation 2:

$$N_{predicted} = e^{(\alpha_0)} AADT_{before}^{\beta_0} L_{before} \times 365 \times 10^{-6} \quad (2)$$

Where:

α_0 = the relationship between crash frequency and roadway characteristic of rural, two-lane, undivided highways in the before period; and

β_0 = the relationship between crash frequency and AADT in the before period

The total number of collisions used in the SPF are derived from the treatment sites and control sites. This total number of collisions is affected by the AADT and other roadway characteristics (including the segment length). The effect of the AADT is evaluated by determining the β coefficient.

STEP 2: Determine the β coefficient for the before-period SPF

The specific values for the coefficients α and β are needed to complete the SPF. The SPF in Equation 2 is modified to include vehicle miles traveled (VMT) embedded within the AADT as shown in Equation 3. As per the U.S. Department of Transportation definition, vehicle miles traveled is the measurement of the total miles traveled by vehicles within a specific time-period (FHWA 2017).

VMT for two years was calculated as follows:

$$VMT = AADT \times L \times 730 \times 10^{-6} \quad (3)$$

VMT can be inserted into the before prediction SPF, Equation 2, and simplified as shown below:

$$N_{predicted} = e^{\alpha} AADT^{\beta-1} \times (VMT)$$

$$\frac{N_{predicted}}{VMT} = e^{\alpha} AADT^{\beta-1}$$

$$\ln\left(\frac{N_{predicted}}{VMT}\right) = \ln(e^{\alpha} AADT^{\beta-1})$$

$$\ln\left(\frac{N_{predicted}}{VMT}\right) = (\beta - 1)\ln(AADT) + \alpha \quad (4)$$

VMT is calculated by multiplying the amount of daily traffic on a roadway segment by the length of the segment. The relationship demonstrated in Equation 4, between the observed crashes during the before period and their respective AADT and segment lengths, is fitted to determine the appropriate β coefficient.

STEP 3: Determine the α coefficient of the before-period SPF

The α coefficient is determined by accounting for all roadways in the before–after set. Also, the AADT for each section must be corrected to account for all segments in the before period, even those with no crashes. The corrected AADT must be adjusted by a ratio of the standard AADT rate to the treatment AADT, referencing a weighted average.

Using the original SPF equation, the sums of the roadway segments and number of crashes should be included as follows:

$$\frac{\sum N_{observed\ before\ crashes,i}}{Total\ VMT\ for\ all\ roadways} = e^{(\alpha_0)} \overline{corrected\ AADT_t}^{\beta} \quad (5)$$

$$\overline{\text{corrected AADT}_i^\beta} = \text{ave} \left(\left(\frac{\text{AADT}_i}{\text{ave}(\text{AADT})} \right)^\beta \right) \quad (6)$$

$$\alpha_0 = \ln \left(\frac{\frac{\sum N_{\text{observed before crashes},i}}{\text{Total VMT}}}{\overline{\text{corrected AADT}_i^\beta}} \right) \quad (7)$$

STEP 4: Using the before-period SPF determined, calculate the predicted average crash frequency for the treatment group during the before period

3.3.2 After-Period SPF Parameters

STEP 5: Select the after-period SPF

The after-period SPF is calculated as:

$$N_{\text{clrs,observed}} = e^{(\alpha_1)} \text{AADT}_{\text{clrs,observed}}^{\beta_1} L_{\text{clrs,observed}} \times 730 \times 10^{-6} \quad (8)$$

Where:

α_1 = the relationship between crash frequency and roadway characteristics of rural, two-lane, undivided highways (including the CLRS) in the after period; and

β_1 = the relationship between crash frequency and AADT in the after period

STEP 6: Determine the β coefficients of the after-period SPF

The specific values for the coefficients, α_1 and β_1 , are needed to complete the SPF.

The prediction SPF for the after period is calculated as:

$$N_{predicted,after} = e^{(\alpha_1)} AADT_{after,clrs}^{\beta_1} L_{after,clrs} \times 730 \times 10^{-6} \quad (9)$$

$$\ln\left(\frac{N_{predicted}}{VMT}\right) = (\beta - 1)\ln(AADT) + \alpha \quad (10)$$

As with the before period, to determine the β coefficient in the after period, the relationship between the observed crashes during the after period and their respective AADT and segment lengths is fitted.

STEP 7: Determine the α coefficient of the after-period SPF

The α coefficient for the after period is determined by the same method as for the before period.

$$\overline{corrected\ AADT_i^\beta} = ave\left(\left(\frac{AADT_i}{ave(AADT)}\right)^\beta\right) \quad (11)$$

$$\alpha_0 = \ln\left(\frac{\sum N_{observed\ before\ crashes,i}}{\frac{Total\ VMT}{\overline{corrected\ AADT_i^\beta}}}\right) \quad (12)$$

STEP 8: Using the after-period SPF determined, calculate the predicted average crash frequency for the treatment group during the after period

3.3.3 Determination of Crash Modification Factor

The crash modification factor (CMF) for CLRS is determined by comparing the observed after-period data with the predictions from the associated SPFs.

STEP 9: Compare the observed number of crashes at the treatment sites with the predicted crashes in the before period

When comparing the observed collisions to the expected collisions, they are related by the site effects, which include all roadway characteristics found at the study sites, as seen in Equation 13.

$$\begin{aligned}
 (\text{Observed No. of Collisions})_{\text{before}} \\
 &= (\text{Predicted No. of Collisions})_{\text{before}} \\
 &\quad \times (\text{site effect})_{\text{before}}
 \end{aligned}
 \tag{13}$$

Site effects include all the factors that influence the crash rate at a certain site, such as roadway geometry, pavement condition, weather and environment, driver vehicle, etc.

$$\begin{aligned}
 \text{site effect} &= \text{effect of roadway geometry} \times \text{effect of weather} \\
 &\quad \times \text{effect driver behavior} \\
 &\quad \times \text{effect of pavement condition} \\
 &\quad \times \text{effect of vehicle fleet and age} \times \text{etc ...}
 \end{aligned}
 \tag{14}$$

The study period was chosen to be for the same duration and months so as to reflect the same effects in both the before and after periods. The only difference between the two periods is the addition of CLRS. The site effect in the after period is the same as that of the before period multiplied by the effect of the CLRS, which is quantified in the CMF (see Equation 15).

$$(\text{site effect})_{\text{after}} = (\text{site effect})_{\text{before}} \times CMF_{\text{clrs}}
 \tag{15}$$

The relationship between the after-period SPF and the before-period SPF is affected by the temporal trend in the data.

$$SPF_{\text{after}} = SPF_{\text{before}} \times \text{temporal trend}
 \tag{16}$$

The after-period SPF is used to predict the number of crashes at the treatment sites if the CLRS treatment was not installed.

$$(Predicted\ No.\ of\ Collisions)_{after} = SPF_{after} \times (site\ effect)_{after} \quad (17)$$

The observed number of collisions is found in a similar way as for the before period but now includes the presence of CLRS.

$$\begin{aligned} (Observed\ No.\ of\ Collisions)_{after} & \quad (18) \\ & = (Predicted\ No.\ of\ Collisions)_{after} \\ & \quad \times (site\ effect)_{before} \times CMF_{clrs} \end{aligned}$$

The before-period site effect is accounted for in the ratio of the average observed and the predicted crash frequencies in the before period.

$$(site\ effect)_{before} = \frac{(Average\ Observed\ No.\ of\ Collisions)_{before}}{(Predicted\ No.\ of\ Collisions)_{before}} \quad (19)$$

4 Results

4.1 Crash Statistics

This section examines comparative statistics of crashes in the before and after periods. This constitutes a naïve analysis.

4.1.1 Total Target Crashes

Overall, 1550 target crashes occurred on all segments during the study period. During the before period (June 1, 2003, to May 31, 2005), 98 and 739 target crashes occurred on CLRS and non-CLRS sites, respectively, for a total of 837 target crashes. In the after period (June 1, 2005, to May 31, 2008), 56 and 657 target crashes occurred on CLRS and non-CLRS sites, respectively, for a total of 713 target crashes. Table 8 summarizes these data by 12-month period. Table 9 shows a site-by-site comparison breakdown.

Table 8. Total Crashes

	Before		After		SUM	
	6/1/2003– 5/31/2004	6/1/2004– 5/31/2005	6/1/2006– 5/31/2007	6/1/2007– 5/31/2008	Before	After
CLRS	52	46	31	25	98	56
NON-CLRS	332	407	327	330	739	657

Table 9. Site-by-Site Comparison, All Crash Types

Study Sites	No. of Crashes in Before Period	No. of Crashes in After Period	Change	
SR 14	13	3	10	-76.9%
SR 16	28	9	19	-67.9%
SR 369	23	16	7	-30.4%
SR 42 A	6	1	5	-83.3%
SR 42 B	0	3	-3	—
SR 204	3	4	-1	+33.3%
SR 36 A	3	1	2	-66.7%
SR 36 B	11	3	8	-72.7%
SR 136	17	18	-1	+5.9%
Overall	104	58	46	-44.2%

4.1.2 Analysis of Crash Severities and Types

A naïve before–after analysis of severity types shows all sites saw a decline of the number of injuries and fatalities. Table 10 shows that CLRS sites experienced declines of -59.1% and -28.6% in injuries and fatalities, respectively. Though not as pronounced, non-CLRS sites had declines of -0.7% and -8.0% for injuries and fatalities.

Figure 6 shows that for CLRS sites, the proportion of crashes with injuries or fatalities declined by -8.17% and -0.34%, respectively. Non-CLRS sites experienced a slight increase (+0.11%) in the proportion of crashes with injuries and a slight decrease (-0.05%) in the proportion for fatalities.

Table 10. Number of Individuals Injured

		Before	After	% Change
Injuries	CLRS	88	36	-59.1%
	NON-CLRS	595	591	-0.7%
Fatalities	CLRS	7	5	-28.6%
	NON-CLRS	50	46	-8.0%

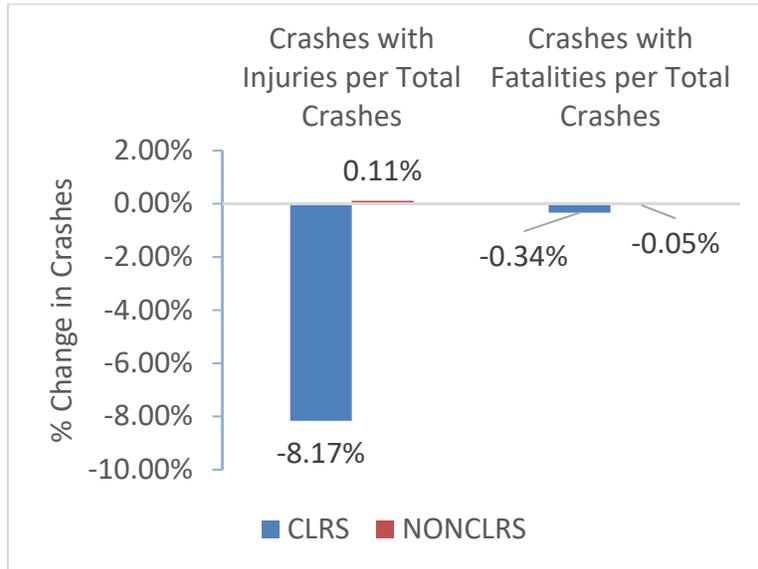


Figure 6. Percent Change of Crashes by Severity in Before versus After Period

4.1.3 Naïve Analysis of Crashes by Collision Type

Table 11 summarizes the number of target crashes in the before and after periods. With the exception of opposite-direction sideswipe collisions at the reference (non-CLRS) sites, all crash types showed decreases, albeit with greater decreases at the CLRS sites. Similarly, Figure 7 illustrates the change in these crashes as a portion of all crashes. Table 12 provides results for treatment sites by type of collision. Comparison of crash types for the before and after periods for all treatment sites for head-on crashes, opposite-direction sideswipe

crashes, and not-a-collision-with-a-motor-vehicle crashes is shown in Tables 13–15, respectively.

Table 11. Crash by Collision Type

Type of Collision		Before	After	% Change
Head On	CLRS	12	7	-41.7%
	NON-CLRS	70	55	-21.4%
Sideswipe—Opposite Direction	CLRS	30	6	-80.0%
	NON-CLRS	95	116	+22.1%
Not a Collision with a Motor Vehicle	CLRS	56	43	-23.2%
	NON-CLRS	570	484	-15.1%

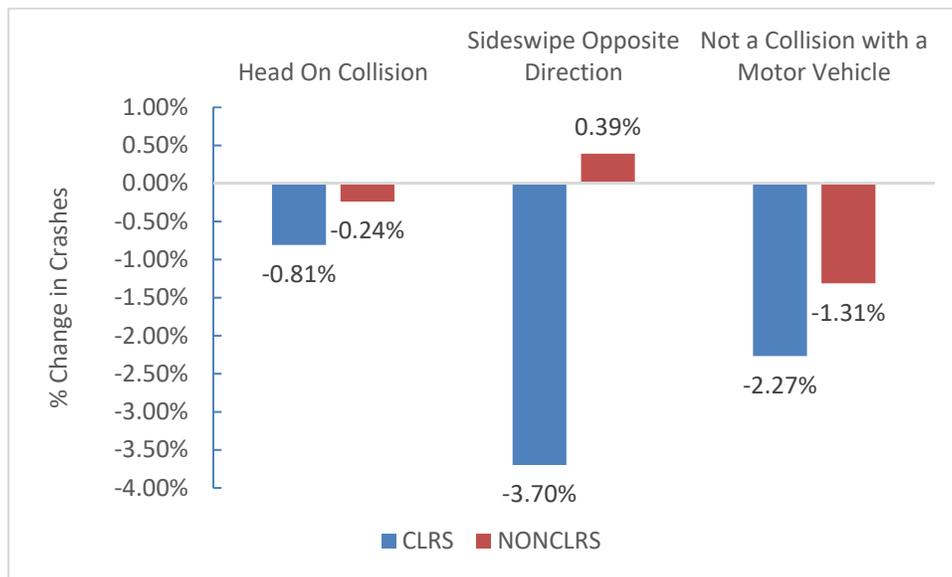


Figure 7. Percent Change in Crashes by Type

Table 12. Site-by-Site Comparison by Collision Type

Study Sites	Crash Type	Before	After
		Number of Crashes	Number of Crashes
SR 14	Head On	3	0
	Sideswipe—Opposite direction	2	0
	Not a collision with a motor vehicle	6	3
SR 16	Head On	3	1
	Sideswipe—Opposite direction	10	1
	Not a collision with a motor vehicle	14	7
SR 369	Head On	3	3
	Sideswipe—Opposite direction	12	3
	Not a collision with a motor vehicle	7	10
SR 42 A	Head On	0	0
	Sideswipe—Opposite direction	1	0
	Not a collision with a motor vehicle	5	1
SR 42 B	Head On	0	1
	Sideswipe—Opposite direction	0	0
	Not a collision with a motor vehicle	0	2
SR 204	Head On	0	0
	Sideswipe—Opposite direction	0	1
	Not a collision with a motor vehicle	3	3
SR 36 A	Head On	0	0
	Sideswipe—Opposite direction	1	0
	Not a collision with a motor vehicle	2	1
SR 36 B	Head On	2	0
	Sideswipe—Opposite direction	1	0
	Not a collision with a motor vehicle	6	2
SR 136	Head On	1	2
	Sideswipe—Opposite direction	3	1
	Not a collision with a motor vehicle	13	14

Table 13. Comparison of Head-on Crashes

Study Sites	Number of Target Crashes		Change in Crashes	
	Before	After	Number	Percent
SR 14	3	0	-3	-100%
SR 16	3	1	-2	-66.7%
SR 369	3	3	0	0.0%
SR 42 A	0	0	0	—
SR 42 B	0	1	+1	—
SR 204	0	0	0	—
SR 36 A	0	0	0	—
SR 36 B	2	0	-2	-100.0%
SR 136	1	2	+1	+100.0%
Total	12	7	-5	-41.7%

Table 14. Comparison of Opposite-Direction Sideswipe Crashes

Study Sites	Number of Target Crashes		Change in Crashes	
	Before	After	Number	Percent
SR 14	2	0	-2	-100.0%
SR 16	10	1	-9	-90.0%
SR 369	12	3	-9	-75.0%
SR 42 A	1	0	-1	-100.0%
SR 42 B	0	0	0	—
SR 204	0	1	+1	—
SR 36 A	1	0	-1	-100.0%
SR 36 B	1	0	-1	-100.0%
SR 136	3	1	-2	-66.7%
Total	30	6	-24	-80.0%

Table 15. Comparison of Not-a-Collision-with-a-Motor-Vehicle Crashes

Study Sites	Number of Target Crashes		Change in Crashes	
	Before	After	Number	Percent
SR 14	6	3	-3	-50.0%
SR 16	14	7	-7	-50.0%
SR 369	7	10	+3	+42.9%
SR 42 A	5	1	-4	-80.0%
SR 42 B	0	2	+2	—
SR 204	3	3	0	—
SR 36 A	2	1	-1	-50.0%
SR 36 B	6	2	-4	-66.7%
SR 136	13	14	+1	+7.7%
Total	56	43	-13	-23.2%

4.2 Empirical Bayes Method/Development of SPF

Table 16 lists the critical parameters used in the EB method to determine the respective SPF for each study period.

Table 16. Crash Statistics

	Before	After
Segments with crashes	623	532
All roadway segments w/o CLRS	2,414	2,318
Average AADT (vehicles)	4,302	4,217
Total VMT (millions)	20,384	19,869
Standard crash frequency rate (target crashes/yr/10⁶ VMT)	0.0431	0.0377

4.2.1 Before-Period SPF Parameters

STEP 1: Select the before-period SPF

Since the number of crashes in this analysis period is for two years, the general prediction SPF (Equation 2) is modified as shown in Equation 20:

$$N_{predicted,before} = e^{(\alpha_0)} AADT_{before}^{\beta_0} L_{before} \times 730 \times 10^{-6} \quad (20)$$

STEP 2: Determine the β coefficient of the before-period SPF

Equation 20 is modified to include vehicle miles traveled embedded within the AADT, as shown in Equation 3.

$$VMT = AADT \times L \times 730 \times 10^{-6} \quad (21)$$

VMT is incorporated into the before prediction SPF, Equation 20, and simplified as in Equation 4.

The study set included 623 roadway segments without CLRS that experienced crashes. The relationship demonstrated in Equation 4, between the observed crashes during the before period and their respective AADT and segment lengths, is used to determine the β coefficient as illustrated in Figure 8.

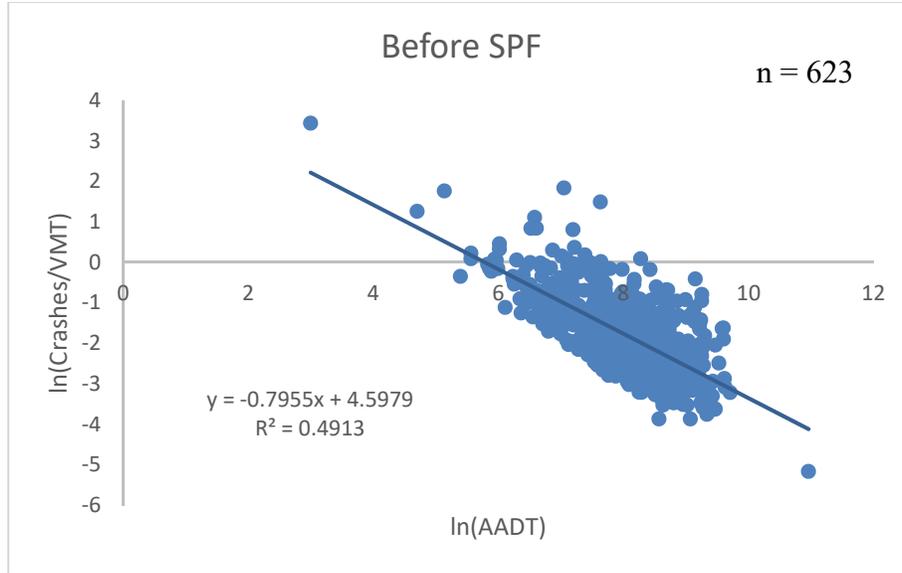


Figure 8. BEFORE: ln(crashes per VMT) versus ln(AADT)

$$-0.7955 = (\beta - 1), \text{ therefore } \beta = 0.2045 \quad (22)$$

STEP 3: Determine the α coefficient of the before-period SPF

The α coefficient is determined by accounting for all roadways in the before–after set, even those with no target crashes. The before-period study set comprised 2414 roadway segments without CLRS. The average AADT for all these segments was 4302 vehicles. AADT must be adjusted by the ratio of the base rate to the treatment AADT, referencing a weighted average. To do so, the AADT for each section was corrected by dividing each individual AADT by the average AADT. Then these values were replaced in the before-period SPF equation and raised to the β coefficient determined in step 2. The standard condition was determined by averaging these values, and was 0.95, as seen in Equation 23.

$$\overline{\text{corrected AADT}_i^\beta} = \text{ave} \left(\left(\frac{\text{AADT}_i}{\text{ave}(\text{AADT})} \right)^{0.2045} \right) = 0.95 \quad (23)$$

$$\text{ave}(\text{AADT}) = 4302 \text{ vehicles}$$

Alpha was then determined by incorporating the standard condition into the original before-period SPF, as shown in Equations 24 and 25.

$$\frac{\sum N_{observed\ before\ crashes,i}}{Total\ VMT\ for\ all\ roadways} = e^{(\alpha_0)} \overline{corrected\ AADT_i^\beta} \quad (24)$$

$$\alpha_0 = \ln \left(\frac{\frac{\sum N_{observed\ before\ crashes,i}}{Total\ VMT}}{corrected\ AADT_i^\beta} \right) = \ln \left(\frac{\frac{834}{20384}}{0.95} \right) = -3.1435 \quad (25)$$

The $e^{-3.1435}$ constant gives the standard crash frequency rate for the before period, which is 0.431 target crashes/yr/ 10^6 VMT.

STEP 4: Using the before-period SPF determined, calculate the predicted average crash frequency for the treatment group during the before period

Using the treatment before sites' AADT and segment lengths, the predicted number of crashes is calculated as follows:

$$N_{predicted,before} = e^{-3.1435} AADT_{before_i}^{0.2076} L_{before_i} \times 730 \times 10^{-6} \quad (26)$$

However, each segment has different values of AADT per year and only has a certain number of days in that year. Since the before-period study is from June 1, 2003, to May 31, 2005, the predicted frequency is more accurately calculated as shown in Equation 27.

$$\begin{aligned}
N_{\text{predicted,before}} &= (e^{-3.1435} AADT_{2003_i}^{0.2076} L_{\text{before}_i} \times 181 \times 10^{-6}) \\
&+ (e^{-3.1435} AADT_{2004_i}^{0.2076} L_{\text{before}_i} \times 365 \times 10^{-6}) \\
&+ (e^{-3.1435} AADT_{2005_i}^{0.2076} L_{\text{before}_i} \times 184 \times 10^{-6})
\end{aligned} \tag{27}$$

Table 17 displays the total predicted crashes on all nine sites.

Table 17. Predicted Crash Frequency in Before Period

Site	Predicted before total crash frequency (vehicles)
SR 14	3.44
SR 16	6.17
SR 369	11.02
SR 42 A	0.83
SR 42 B	1.93
SR 204	1.96
SR 36 A	1.24
SR 36 B	2.37
SR 136	1.61
Total	30.6

4.2.2 After-Period SPF Parameters

STEP 5: Select the after-period SPF

As in the before period, the general equation for the after-period SPF is:

$$N_{predicted,after} = e^{(\alpha_1)} AADT_{after}^{\beta_1} L_{after} \times 730 \times 10^{-6} \quad (28)$$

STEP 6: Determine the β coefficient of the after-period SPF

Using the same equation as Equation 4, but with the AADT and segment length values of the after-period comparison set, 532 roadway segments with crashes were plotted as shown in Figure 9.

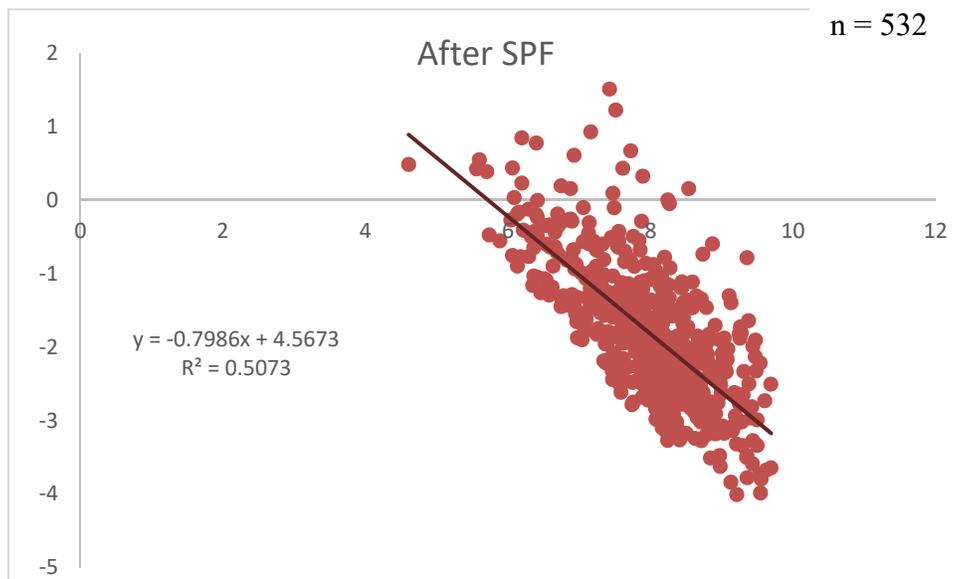


Figure 9. AFTER: ln(crashes per VMT) versus ln(AADT)

$$-0.7986 = (\beta - 1), \text{ therefore } \beta = 0.2014 \quad (29)$$

STEP 7: Determine the α coefficient of the after-period SPF

The α coefficient for the after period is determined by the same method as the before period. This set contained a total of 2318 roadway segments without CLRS. The average AADT for all these segments was 4218 vehicles. The standard condition is determined by averaging these values, which was 0.95, as seen in Equation 30.

$$\overline{\text{corrected AADT}_i^{\beta_1}} = \text{ave} \left(\left(\frac{\text{AADT}_i}{\text{ave}(\text{AADT})} \right)^{\beta_1} \right) = 0.95 \text{ t} \quad (30)$$

$$\text{ave}(\text{AADT}) = 4,218 \text{ vehicles}$$

$$\text{Total VMT} = 20,384 \times 10^6 \text{ miles}$$

$$\alpha_1 = \ln \left(\frac{\frac{\sum N_{\text{observed after crashes},i}}{\text{Total VMT}}}{\text{corrected AADT}_i^{\beta_1}} \right) = \ln \left(\frac{\frac{711}{19879}}{0.95} \right) = -3.2790 \quad (31)$$

The $e^{-3.2790}$ constant gives the standard crash frequency rate for the before period, which is 0.377 target crashes/yr/ 10^6 VMT.

STEP 8: Using the after-period SPF determined, calculate the predicted average crash frequency for the treatment group during the after period

Using the treatment after sites' AADT and segment lengths, the predicted number of crashes in the after period is generally calculated by Equation 32:

$$\begin{aligned} N_{\text{clrs,predicted,after}} \\ = e^{(-3.2790)} \text{AADT}_{\text{clrs,after}}^{0.2014} L_{\text{clrs,after}} \times 730 \times 10^{-6} \end{aligned} \quad (32)$$

The predicted frequency for the after period from June 1, 2006, to May 31, 2008, was more accurately calculated as shown in Equation 33.

$$\begin{aligned}
N_{\text{predicted,after}} = & (e^{-3.2790} AADT_{2006_i}^{0.2014} L_{\text{CLRS}_i} \times 181 \times 10^{-6}) \\
& + (e^{-3.2790} AADT_{2007_i}^{0.2014} L_{\text{CLRS}_i} \times 365 \times 10^{-6}) \\
& + (e^{-3.2790} AADT_{2008_i}^{0.2014} L_{\text{CLRS}_i} \times 184 \times 10^{-6})
\end{aligned} \tag{33}$$

Table 18 displays the total predicted crashes on all 9 sites, totaling 26.197 vehicles.

Table 18. Predicted Crash Frequency in After Period

Site	Predicted after total crash frequency (vehicles)
SR 14	2.8
SR 16	5.3
SR 369	9.2
SR 42 A	0.8
SR 42 B	1.9
SR 204	1.7
SR 36 A	1.0
SR 36 B	2.2
SR 136	1.4
Total	26.3

4.2.3 Determination of CMF

The CMF for CLRS is determined by analyzing the data from the before period and the after period.

STEP 9: Compare the observed number of crashes at the treatment sites with the predicted crashes in the before period

As discussed in Section 3.3.3 (Equations 13–19), the CMF is calculated as:

$$\begin{aligned} CMF_{clrs} &= \frac{(Observed\ No.\ of\ Collisions)_{after} \times (Predicted\ No.\ of\ Collisions)_{before}}{(Predicted\ No.\ of\ Collisions)_{after} \times (Observed\ No.\ of\ Collisions)_{before}} \\ &= \frac{56 \times 30.58}{29.96 \times 98} = 0.58329 \approx \mathbf{0.58} \end{aligned}$$

4.3 Misclassified Crashes

Table 19 shows the number of misclassified target crashes and the reason for their misclassification. Misclassifications were found in 6.73% of all target crashes. The category “not a collision with a motor vehicle” had the most misclassifications. This was mainly due to the use of wrong definitions for each classification. A head-on or angle collision involves more than one vehicle. When a motor vehicle collides with anything other than another motor vehicle, it is considered “not a collision with a motor vehicle.” Some police officers misinterpreted a motor vehicle crashing head on or at angle with an object as “head on” or “angle.” This error could easily be prevented in the future by more specific training.

Table 19. Misclassified Target Crashes

		Correct Classification			Subtotal
		Sideswipe— Opposite direction	Not a collision with a motor vehicle	Head on	
Misclassification	Sideswipe— Opposite direction		2	3	5
	Not a collision with a motor vehicle	2		—	2
	Head on	2	29		31
	Angle	23	21	14	58
	Rear	—	1	—	1
	No classification/ left blank	—	2	—	2
	Sideswipe— Same direction	3	1	1	5
Subtotal		30	56	18	104

5 Conclusions and Recommendations

This study used a seasonally adjusted 24-month pre-deployment (before) period and a 24-month post-deployment (after) period for comparison. The empirical Bayesian analysis resulted in a crash modification factor of 0.58 for the CLRS treatment, indicating a 42% reduction in crashes associated with those conditions that CLRS was designed to address (i.e., crashes involving crossing the centerline). However, the sample size was too small to obtain separate crash modification factors for fatal crashes and injury crashes.

The quality assurance procedure was the most resource-intensive part of the effort. The research team manually checked the base crash data against the crash description recorded by the investigating police officer to verify crash type, as well as obtain a clearer indication as to whether the crash could have been impacted by the presence of CLRS. This step was critical to improving the reliability of the CMF value, as it reduced crash misclassifications.

The involvement of multiple agencies in the recording of the crash data naturally introduces variability and non-uniformity in the crash data. These errors become critical, particularly when the results are dependent on the correct categorization of the incidents and the correct association of the incidents to a safety measure. A broader methodological recommendation from the lessons learned in the study is to employ sufficient crash verification procedures in any safety study that develops a crash modification factor, especially in cases where the sample size of the crashes is small, or if crash modification factors are desired for specific crash categories.

The Phase 1 report for this project showed that most of the maintenance and commonly cited deployment concerns can be addressed using some modification to the

design of the CLRS. Most of the state DOTs surveyed as part of that study were amenable to further use of CLRS as a safety countermeasure.

The favorable crash modification factor (0.58) obtained in this study clearly provides sufficient justification for the use of CLRS as a low-cost safety countermeasure to address crashes involving vehicles that cross the centerline of the roadway.

6 References

- AASHTO. (2010). Highway Safety Manual (1st ed., Vol. 2). Washington, DC: American Association of State Highway and Transportation Officials.
- Alexander, J., and P. Garder. (1995). "Continued Research on Continuous Rumble Strips." Final Technical Report 94-4, Maine Department of Transportation, Augusta.
- Carriquiry, A., and M. Pawlovich. (2004). "From Empirical Bayes to Full Bayes: Methods for Analyzing Traffic Safety Data." White Paper, Iowa State University.
- Federal Highway Administration. (2017). Planning Glossary. Retrieved November 2, 2017, [https://www.fhwa.dot.gov/planning/glossary/glossary_listing.cfm?TitleStart=V \(\)](https://www.fhwa.dot.gov/planning/glossary/glossary_listing.cfm?TitleStart=V)
- Fontaine, Michael D., and J. S. Gillespie. (2009). "Synthesis of Benefits and Costs of Alternative Lane Marking Strategies." Report Number: FHWA/VTRC 09-R24, Virginia Department of Transportation.
- GDOT. (2017). Office of IT Applications. GeoPI, <http://www.dot.ga.gov/DS/Maps/geopi>
- GDOT. (2013). "Transportation Project Information (TransPI)." Retrieved November 9, 2013, <http://www.dot.ga.gov/BS/Projects/ProjectSearch>
- GDOT. (2005). Safety Action Plan, Atlanta, GA.

- Gross, F., B. Persaud, and C. Lyon. (2010). "A Guide to Developing Quality Crash Modification Factors." Report FHWA-SA-10-032, Federal Highway Administration.
- Guin, A., M. Hunter, M. Rodgers, and J. Sin (2014). "Centerline Rumble Strips Safety and Maintenance Impacts." Final Report, FHWA-GA-14-1212, Georgia Department of Transportation.
- Hallmark, S., T. McDonald, and R. Sperry (2009). "Evaluation of Rumble Stripes on Low-volume Rural Roads in Iowa—Phase I." Final Report. Report Number: IHRB Project TR-577, Iowa Highway Research Board.
- Harwood, D. W. (1993). Use of rumble strips to enhance safety.
- Hauer, E. (1997). "Observational Before–After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety." Oxford: Pergamon.
- Hauer, E. (2001). "Overdispersion in Modelling Accidents on Road Sections and in Empirical Bayes Estimation." *Accident Analysis and Prevention* 33(6): 10.
- Herbel, S., L. Laing, and C. McGovern. (2010). *Highway Safety Improvement Program Manual*. U.S. Department of Transportation, Federal Highway Administration, Office of Safety.

- Kay, J., P. T. Savolainen, T. J. Gates, T. K. Datta, J. Finkelman, and B. Hamadeh. (2015). "Safety Impacts of a Statewide Centerline Rumble Strip Installation Program." *Transportation Research Record: Journal of the Transportation Research Board*.
- Karkle, D. E., et al. (2013). "Safety effectiveness of centerline rumble strips in Kansas." *Journal of Transportation Safety & Security* 5(1): 1-26.
- Rodgers, Michael O., A. M. Wilson., A. Shaw, and M. T. Barton. (2015). "Applying the Highway Safety Manual to Georgia." Forest Park, GA, Georgia Institute of Technology: 167.
- Outcalt, W. (2001). "Centerline Rumble Strips." Report Number CDOT-DTD-R-2001-8. Colorado Department of Transportation.
- Persaud, B., B. Lan, C. Lyon, and R. Bhim. (2010). "Comparison of Empirical Bayes and Full Bayes Approaches for Before–After Road Safety Evaluations." *Accident Analysis & Prevention* 42(1): 38–43.
- Persaud, B., R. Retting, P. Garder, and D. Lord. (2001). "Safety Effect of Roundabout Conversions in the United States: Empirical Bayes Observational Before–After Study." *Transportation Research Record: Journal of the Transportation Research Board* (1751): 1–8.
- Persaud, B. N., R. A. Retting, and C. A. Lyon. (2004). Crash Reduction Following Installation of Centerline Rumble Strips on Rural Two-lane Roads, *Accident Analysis & Prevention*, 36(6): 1073–1079.

- Porter, R., E. Donnell, and K. Mahoney. (2004). "Evaluation of Effects of Centerline Rumble Strips on Lateral Vehicle Placement and Speed." *Transportation Research Record: Journal of the Transportation Research Board* (1862): 10–16.
- Russell, E. R., and M. J. Rys. (2005). "NCHRP Synthesis 339: Centerline Rumble Strips." *Transportation Research Record: Journal of the Transportation Research Board*.
- Sayed, T., P. deLeur, and J. Pump. (2010). "Impact of Rumble Strips on Collision Reduction on Highways in British Columbia, Canada: Comprehensive Before-and-After Safety Study." *Transportation Research Record: Journal of the Transportation Research Board* (2148): 9–15.
- Sin, J. G. N. (2014). "Safety Impact Study of Centerline Rumble Strips in Georgia." Master of Science, Georgia Institute of Technology.
- Torbic, D. J. et al. (2009). "Guidance for the Design and Application of Shoulder and Centerline Rumble Strips." NCHRP Report 641, Transportation Research Board, Washington, D.C.

Appendix

Crash Database Sanitation

The police reports corresponding to the filtered control and treatment crashes contain personally identifiable information and, therefore, required a sanitization process. The first step of the sanitation process involved the use of a Perl script to redact sensitive information. Then, a GDOT-authorized database user manually reviewed the redacted version of the crash report. These reports were in standard two-page format with supplemental pages provided on some reports (e.g., when multiple vehicles were involved or injuries had occurred). The first page of the report always contained certain personally identifiable information, which was not pertinent to this research effort. Personally identifiable information may, or may not, be present in subsequent pages. Given the non-uniformity of the scanned reports, full automation of the sanitization process was challenging.

The Perl script accomplished several sanitization tasks. First, all police report image files were renamed to replace the crash ID with a unique ID used in the sanitized database. Once this step was completed, the table containing the link between the crash ID and the unique ID was securely destroyed. Hence, this unique ID had no link and could not be traced back to the original crash ID. Next, each police report image file was converted to a series of images, every image representing one page. Each image was identified with the unique ID that allows it to be linked back to the sanitized database where other non-personally identifiable information related to the crash is available. Since the information

on the first page of each report was not needed, it was then deleted, as it contained personally identifiable information. Subsequently, the second page image was verified to ensure proper orientation and was inverted, if necessary. Portions of the second page, where personally identifiable information was present, were then electronically blanked out. If the original record had only two pages, then the record quite likely was already fully sanitized; it needed to be verified for the existence of any unusual personally identifiable information since, often times, the officer includes information in the description such as the names of those involved, contact information, vehicle identification numbers (VINs), and other personally identifiable data pertaining to the individuals engaged in the crash. If the record had more than two pages, the remaining page images were manually checked by a GDOT-authorized database user to identify any personally identifiable information. Any images containing personally identifiable information that were not relevant for research were deleted. Any sensitive information, such as VINs or driver names and contact information, was manually removed with the software XnView®, a multi-format graphics viewer with image-processing capabilities. Removed data cannot be restored after they have been saved, thus ensuring the privacy of the individuals involved in the crash reports. Finally, the sanitized versions were made available to students and other researchers, as they were necessary for analysis in normal research applications.