Chapter 9—Safety Effectiveness Evaluation

9.1. CHAPTER OVERVIEW

Evaluating the change in crashes from implemented safety treatments is an important step in the roadway safety evaluation process (see Figure 9-1). Safety evaluation leads to an assessment of how crash frequency or severity has changed due to a specific treatment or a set of treatments or projects. In situations where one treatment is applied at multiple similar sites, safety evaluation can also be used to estimate a crash modification factor (CMF) for the treatment. Finally, safety effectiveness evaluations have an important role in assessing how well funds have been invested in safety improvements. Each of these aspects of safety effectiveness evaluation may influence future decision-making activities related to allocation of funds and revisions to highway agency policies.

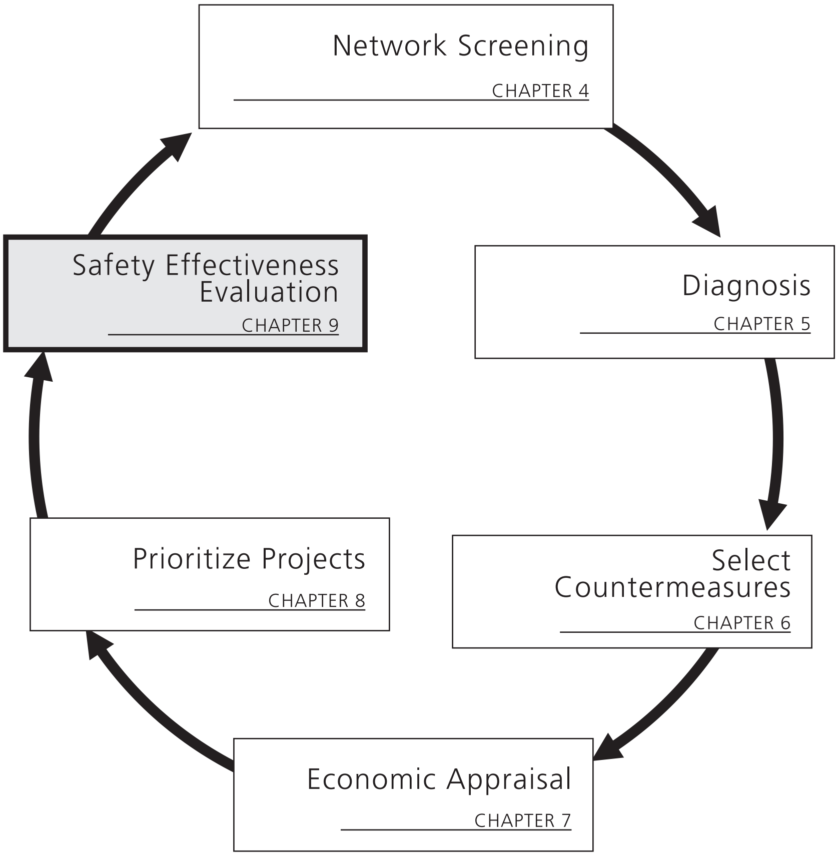


Figure 9-1. Roadway Safety Management Overview Process

The purpose of this chapter is to document and discuss the various methods for evaluating the effectiveness of a treatment, a set of treatments, an individual project, or a group of similar projects after improvements have been implemented to reduce crash frequency or severity. This chapter provides an introduction to the evaluation methods that can be used, highlights which methods are appropriate for assessing safety effectiveness in specific situations, and provides step-by-step procedures for conducting safety effectiveness evaluations.

9.2. SAFETY EFFECTIVENESS EVALUATION—DEFINITION AND PURPOSE

Safety effectiveness evaluation is the process of developing quantitative estimates of how a treatment, project, or a group of projects has affected crash frequencies or severities. The effectiveness estimate for a project or treatment is a valuable piece of information for future safety decision making and policy development.

Safety effectiveness evaluation may include:

- Evaluating a single project at a specific site to document the safety effectiveness of that specific project,
- Evaluating a group of similar projects to document the safety effectiveness of those projects,
- Evaluating a group of similar projects for the specific purpose of quantifying a CMF for a countermeasure, and
- Assessing the overall safety effectiveness of specific types of projects or countermeasures in comparison to their costs.

If a particular countermeasure has been installed on a systemwide basis, such as the installation of cable median barrier or shoulder rumble strips for the entire freeway system of a jurisdiction, a safety effectiveness evaluation of such a program would be conducted no differently than an evaluation of any other group of similar projects.

Safety effectiveness evaluations may use several different types of performance measures, such as a percentage reduction in crashes, a shift in the proportions of crashes by collision type or severity level, a CMF for a treatment, or a comparison of the safety benefits achieved to the cost of a project or treatment.

The next section presents an overview of available evaluation study designs and their corresponding evaluation methods. Detailed procedures for applying those methods are presented in Section 9.4 and Appendix 9A. Sections 9.5 through 9.8, respectively, describe how the evaluation study designs and methods for each of the evaluation types identified above are implemented.

9.3. STUDY DESIGN AND METHODS

To evaluate the effectiveness of a treatment in reducing crash frequency or severity, the treatment must have been implemented for at least one and, preferably, many sites. Selection of the appropriate study design for a safety effectiveness evaluation depends on the nature of the treatment, the type of sites at which the treatment has been implemented, and the time periods for which data are available for those sites (or will become available in the future). The evaluation is more complex than simply comparing before and after crash data at treatment sites because consideration is also given to what changes in crash frequency would have occurred at the evaluation sites between the time periods before and after the treatment even if the treatment had not been implemented. Many factors that can affect crash frequency may change over time, including changes in traffic volumes, weather, and driver behavior. General trends in crash frequency can also affect both improved and unimproved sites. For this reason, most evaluations use data for both treatment and nontreatment sites. Information can be directly obtained by collecting data on such sites or by making use of safety performance functions for sites with comparable geometrics and traffic patterns.

Table 9-1 presents a generic evaluation study design layout that will be used throughout the following discussion to explain the various study designs that can be used in safety effectiveness evaluation. As the exhibit indicates, study designs usually use data (crash and traffic volume) for both treatment and nontreatment sites and for time periods both before and after the implementation of the treatments. Even though no changes are made intentionally to the nontreatment sites, it is useful to have data for such sites during time periods both before and after improvement of the treatment sites so that general time trends in crash data can be accounted for.

Table 9-1. Generic Evaluation Study Design

Type of Site	Before Treatment	After Treatment
Treatment Sites		
Nontreatment Sites		

There are three basic study designs that are used for safety effectiveness evaluations:

- Observational before/after studies
- Observational cross-sectional studies
- Experimental before/after studies

Both observational and experimental studies are used in safety effectiveness evaluations. In observational studies, inferences are made from data observations for treatments that have been implemented by highway agencies in the normal course of efforts to improve the road system, not treatments that have been implemented specifically so they can be evaluated. By contrast, experimental studies consider treatments that have been implemented specifically so that their effectiveness can be evaluated. In experimental studies, sites that are potential candidates for improvement are randomly assigned to either a treatment group, at which the treatment of interest is implemented, or a comparison group, at which the treatment of interest is not implemented. Subsequent differences in crash frequency between the treatment and comparison groups are directly attributed to the treatment. Observational studies are much more common in road safety than experimental studies, because highway agencies are generally reluctant to use random selection in assigning treatments. For this reason, the focus of this chapter is on observational studies.

Each of the observational and experimental approaches to evaluation studies are explained below.

9.3.1. Observational Before/After Evaluation Studies

Observational before/after studies are the most common approach used for safety effectiveness evaluation. An example situation that warrants an observational before/after study is when an agency constructs left-turn lanes at specific locations on a two-lane highway where concerns about crash frequency had been identified. Table 9-2 shows the evaluation study design layout for an observational before/after study to identify the effectiveness of the left-turn lanes in reducing crash frequency or severity.

All observational before/after studies use crash and traffic volume data for time periods before and after improvement of the treated sites. The treatment sites do not need to have been selected in a particular way; they are typically sites of projects implemented by highway agencies in the course of their normal efforts to improve the operational and safety performance of the highway system. However, if the sites were selected for improvement because of unusually high crash frequencies, then using these sites as the treatment sites may introduce a selection bias which could result in a high regression-to-the-mean bias since treatment was not randomly assigned to sites. Chapter 3 provides more information about issues associated with regression-to-the-mean bias.

As shown in Table 9-2, the nontreatment sites (i.e., comparison sites)—sites that were not improved between the time periods before and after improvement of the treatment sites—may be represented either by SPFs or by crash and traffic volume data. Evaluation study design using these alternative approaches for consideration of non-treatment sites are not discussed below.

Table 9-2. Observational Before/After Evaluation Study Design

Type of Site	Before Treatment	After Treatment
Treatment Sites	✓	✓
Non-treatment Sites (SPF or comparison group)	✓	✓

If an observational before/after evaluation is conducted without any consideration of nontreatment sites (i.e., with no SPFs and no comparison group), this is referred to as a simple or naïve before/after evaluation. Such evaluations do not compensate for regression-to-the-mean bias (see Chapter 3) or compensate for general time trends in the crash data.

9.3.2. Observational Before/After Evaluation Studies Using SPFs—the Empirical Bayes Method

Observational before/after evaluation studies that include non-treatment sites are conducted in one of two ways. The Empirical Bayes method is most commonly used. This approach to evaluation studies uses SPFs to estimate what the average crash frequency at the treated sites would have been during the time period after implementation of the treatment, had the treatment not been implemented.

In cases where the treated sites were selected by the highway agency for improvement because of unusually high crash frequencies, this constitutes a selection bias which could result in a high regression-to-the-mean bias in the evaluation. The use of the EB approach, which can compensate for regression-to-the–mean bias, is particularly important in such cases.

Chapter 3 presents the basic principles of the EB method which is used to estimate a site's expected average crash frequency. The EB method combines a site's observed crash frequency and SPF-based predicted average crash frequency to estimate the expected average crash frequency for that site in the after period had the treatment not been implemented. The comparison of the observed after crash frequency to the expected average after crash frequency estimated with the EB method is the basis of the safety effectiveness evaluation.

A key advantage of the EB method for safety effectiveness evaluation is that existing SPFs can be used. There is no need to collect crash and traffic volume data for nontreatment sites and develop a new SPF each time a new evaluation is performed. However, if a suitable SPF is not available, one can be developed by assembling crash and traffic volume data for a set of comparable nontreatment sites.

The EB method has been explained for application to highway safety effectiveness evaluation by Hauer (5,6) and has been used extensively in safety effectiveness evaluations (2,8,10). The EB method implemented here is similar to that used in the FHWA SafetyAnalyst software tools (3). Detailed procedures for performing an observational before/after study with SPFs to implement the EB method are presented in Section 9.4.1 and Appendix 9A.

9.3.3. Observational Before/After Evaluation Study Using the Comparison-Group Method

Observational before/after studies may incorporate nontreatment sites into the evaluation as a comparison group. In a before/after comparison-group evaluation method, the purpose of the comparison group is to estimate the change in crash frequency that would have occurred at the treatment sites if the treatment had not been made. The comparison group allows consideration of general trends in crash frequency or severity whose causes may be unknown, but which are assumed to influence crash frequency and severity at the treatment and comparison sites equally. Therefore, the selection of an appropriate comparison group is a key step in the evaluation.

Comparison groups used in before/after evaluations have traditionally consisted of nontreated sites that are comparable in traffic volume, geometrics, and other site characteristics to the treated sites, but without the specific improvement being evaluated. Hauer (5) makes the case that the requirement for matching comparison sites with respect to site characteristics, such as traffic volumes and geometrics, is secondary to matching the treatment and comparison sites based on their crash frequencies over time (multiple years). Matching on the basis of crash frequency over time generally uses crash data for the period before treatment implementation. Once a set of comparison sites that are comparable to the treatment sites has been identified, crash and traffic volume data are needed for the same time periods as are being considered for the treated sites.

Obtaining a valid comparison group is essential when implementing an observational before/after evaluation study using the comparison-group method. It is therefore important that agreement between the treatment group and comparison-group data in the yearly time series of crash frequencies during the period before implementation of the treatment be confirmed. During the before period, the rate of change in crashes from year to year should be consistent between a particular comparison group and the associated treatment group. A statistical test using the yearly time series of crash frequencies at the treatment and comparison-group sites for the before period is generally used to assess this consistency. Hauer (5) provides a method to assess whether a candidate comparison group is suitable for a specific treatment group.

While the comparison-group method does not use SPF(s) in the same manner as the EB method, SPF(s) are desirable to compute adjustment factors for the nonlinear effects of changes in traffic volumes between the before and after periods.

The before/after comparison-group evaluation method has been explained for application to highway safety effectiveness evaluation by Griffin (1) and by Hauer (5). A variation of the before/after comparison-group method to handle adjustments to compensate for varying traffic volumes and study period durations between the before and after study periods and between the treatment and comparison sites was formulated by Harwood et al. (2). Detailed procedures for performing an observational before/after study with the comparison-group method are presented in Section 9.4.2 and Appendix 9A.

9.3.4. Observational Before/After Evaluation Studies to Evaluate Shifts in Collision Crash Type Proportions

An observational before/after evaluation study is used to assess whether a treatment has resulted in a shift in the frequency of a specific target collision type as a proportion of total crashes from before to after implementation of the treatment. The target collision types addressed in this type of evaluation may include specific crash severity levels or crash types. The procedures used to assess shifts in proportion are those used in the FHWA SafetyAnalyst software tools (3). The assessment of the statistical significance of shifts in proportions for target collision types is based on the Wilcoxon signed rank test (7). Detailed procedures for performing an observational before/after evaluation study to assess shifts in crash severity level or crash type proportions are presented in Section 9.4.3 and Appendix 9A.

9.3.5. Observational Cross-Sectional Studies

There are many situations in which a before/after evaluation, while desirable, is simply not feasible, including the following examples:

- When treatment installation dates are not available;
- When crash and traffic volume data for the period prior to treatment implementation are not available; or
- When the evaluation needs to explicitly account for effects of roadway geometrics or other related features by creating a CMF function rather than a single value for a CMF.

In such cases, an observational cross-sectional study may be applied. For example, if an agency wants to compare the safety performance of intersections with channelized right-turn lanes to intersections without channelized right-turn lanes and no sites are available that have been converted from one configuration to the other, then an observational cross-sectional study may be conducted comparing sites with these two configurations. Cross-sectional studies use statistical modeling techniques that consider the crash experience of sites with and without a particular treatment of interest (such as roadway lighting or a shoulder rumble strip) or with various levels of a continuous variable that represents a treatment of interest (such as lane width). This type of study is commonly referred to as a "with and without study." The difference in number of crashes is attributed to the presence of the discrete feature or the different levels of the continuous variable.

As shown in Table 9-3, the data for a cross-sectional study is typically obtained for the same period of time for both the treatment and comparison sites. Since the treatment is obviously in place during the entire study period, a cross-sectional study might be thought of as comparable to a before/after study in which data are only available for the time period after implementation of the treatment.

Table 9-3. Observational Cross-Sectional Evaluation Study Design

Type of Site	Before Treatment	After Treatment
Treatment Sites		✓
Nontreatment Sites	✓	

There are two substantial drawbacks to a cross-sectional study. First, there is no good method to compensate for the potential effect of regression-to-the-mean bias introduced by site selection procedures. Second, it is difficult to assess cause and effect and, therefore, it may be unclear whether the observed differences between the treatment and nontreatment sites are due to the treatment or due to other unexplained factors (4). In addition, the evaluation of the safety effectiveness requires a more involved statistical analysis approach. The recommended approach to performing observational before/after cross-sectional studies is presented in Section 9.4.4.

9.3.6. Selection Guide for Observational Before/After Evaluation Study Methods

Table 9-4 presents a selection guide to the observational before/after evaluation study methods. If, at the start of a safety evaluation, the user has information on both the safety measure to be evaluated and the types of data available, then the table indicates which type(s) of observational before/after evaluation studies are feasible. On the other hand, based on data availability, the information provided in Table 9-4 may also guide the user in assessing additional data needs depending on a desired safety measure (i.e., crash frequency or target collision type as a proportion of total crashes).

Table 9-4. Selection Guide for Observational Before/After Evaluation Methods

	Data availability					
	Treatment sites		Nontreatment sites			
Safety measure to be evaluated	Before period data	After period data	Before period data	After period data	SPF	Appropriate evaluation study method
Crash frequency	✓	✓			✓	Before/after evaluation study using the EB method
	✓	✓	✓	✓		Before/after evaluation study using either the EB method OR the comparison-group method
		✓		✓		Cross-sectional study
Target collision type as a proportion of total crashes	✓	✓				Before/after evaluation study for shift in proportions

9.3.7. Experimental Before/After Evaluation Studies

Experimental studies are those in which comparable sites with respect to traffic volumes and geometric features are randomly assigned to a treatment or nontreatment group. The treatment is then applied to the sites in the treatment group, and crash and traffic volume data is obtained for time periods before and after treatment. Optionally, data may also be collected at the nontreatment sites for the same time periods. For example, if an agency wants to evaluate the safety effectiveness of a new and innovative signing treatment, then an experimental study may be conducted. Table 9-5 illustrates the study design for an experimental before/after study.

Table 9-5. Experimental Before/After Evaluation Study Design

Type of Site	Before Treatment	After Treatment
Treatment Sites Required Data	✓	✓
Nontreatment Sites (Comparison Group) Optional Data		

The advantage of the experimental over the observational study is that randomly assigning individual sites to the treatment or nontreatment groups minimizes selection bias and, therefore, regression-to-the-mean bias. The disadvantage of experimental studies is that sites are randomly selected for improvement. Experimental before/after evaluations are performed regularly in other fields, such as medicine, but are rarely performed for highway safety improvements because of a reluctance to use random assignment procedures in choosing improvement locations. The layout of the study design for an experimental before/after study is identical to that for an observational before/after evaluation design and the same safety evaluation methods described above and presented in more detail in Section 9.4 can be used.

9.4. PROCEDURES TO IMPLEMENT SAFETY EVALUATION METHODS

This section presents step-by-step procedures for implementing the EB and comparison-group methods for observational before/after safety effectiveness evaluations. The cross-sectional approach to observational before/after evaluation and the applicability of the observational methods to experimental evaluations are also discussed. Table 9-6 provides a tabular overview of the data needs for each of the safety evaluation methods discussed in this chapter.

Table 9-6. Overview of Data Needs and Inputs for Safety Effectiveness Evaluations

	Safety Evaluation Method			
Data Needs and Inputs	EB Before/After	Before/After with Comparison Group	Before/After Shift in Proportion	Cross-Sectional
10 to 20 treatment sites	✓	✓	✓	✓
10 to 20 comparable non-treatment sites		✓		✓
A minimum of 650 aggregate crashes in non-treatment sites		✓		
3 to 5 years of crash and volume "before" data	✓	✓	✓	
3 to 5 years of crash and volume "after" data	✓	✓	✓	✓
SPF for treatment site types	✓	✓		
SPF for non-treatment site types		✓		
Target crash type			✓	

9.4.1. Implementing the EB Before/After Safety Evaluation Method

The Empirical Bayes (EB) before/after safety evaluation method is used to compare crash frequencies at a group of sites before and after a treatment is implemented. The EB method explicitly addresses the regression-to-the-mean issue by incorporating crash information from other but similar sites into the evaluation. This is done by using an SPF and weighting the observed crash frequency with the SPF-predicted average crash frequency to obtain an expected average crash frequency (see Chapter 3). Figure 9-2 provides a step-by-step overview of the EB before/after safety effectiveness evaluation method.

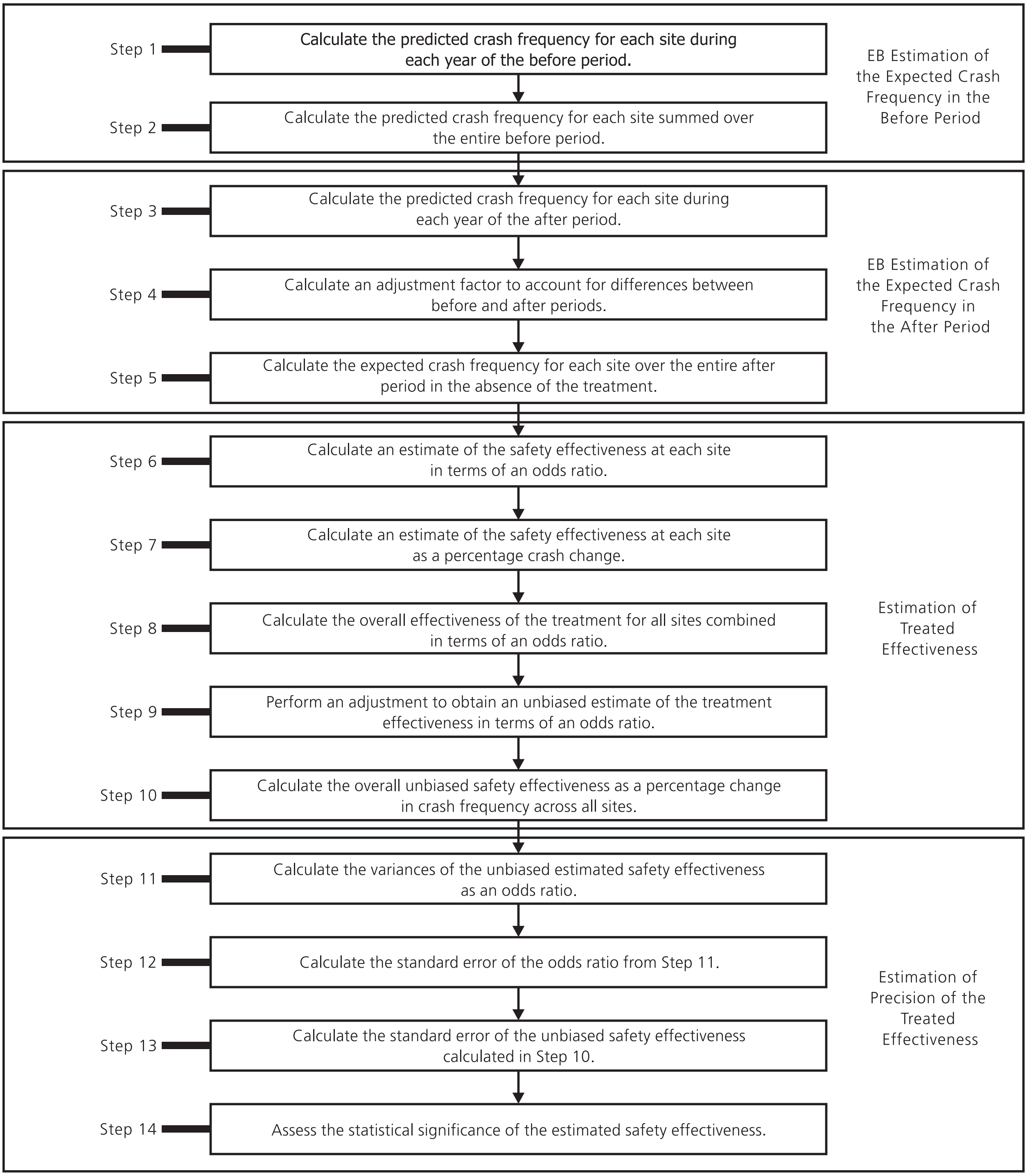


Figure 9-2. Overview of EB Before/After Safety Evaluation

Data Needs and Inputs

The data needed as input to an EB before/after evaluation include:

- At least 10 to 20 sites at which the treatment of interest has been implemented
- 3 to 5 years of crash and traffic volume data for the period before treatment implementation
- 3 to 5 years of crash and traffic volume for the period after treatment implementation
- SPF for treatment site types

An evaluation study can be performed with fewer sites or shorter time periods, or both, but statistically significant results are less likely.

Pre-Evaluation Activities

The key pre-evaluation activities are to:

- Identify the treatment sites to be evaluated.
- Select the time periods before and after treatment implementation for each site that will be included in the evaluation.
- Select the measure of effectiveness for the evaluation. Evaluations often use total crash frequency as the measure of effectiveness, but any specific crash severity level and/or crash type can be considered.
- Assemble the required crash and traffic volume data for each site and time period of interest.
- Identify (or develop) an SPF for each type of site being developed. SPFs may be obtained from SafetyAnalyst or they may be developed based on the available data as described in Part C. Typically, separate SPFs are used for specific types of roadway segments or intersections.

The before study period for a site must end before implementation of the treatment began at that site. The after study period for a site normally begins after treatment implementation is complete; a buffer period of several months is usually allowed for traffic to adjust to the presence of the treatment. Evaluation periods that are even multiples of 12 months in length are used so that there is no seasonal bias in the evaluation data. Analysts often choose evaluation periods consisting of complete calendar years because this often makes it easier to assemble the required data. When the evaluation periods consist of entire calendar years, the entire year during which the treatment was installed is normally excluded from the evaluation period.

Computational Procedure

A computational procedure using the EB method to determine the safety effectiveness of the treatment being evaluated, expressed as a percentage change in crashes, θ, and to assess its precision and statistical significance, is presented in Appendix 9A.

9.4.2. Implementing the Before/After Comparison-Group Safety Evaluation Method

The before/after comparison-group safety evaluation method is similar to the EB before/after method except that a comparison group is used, rather than an SPF, to estimate how safety would have changed at the treatment sites had no treatment been implemented. Figure 9-3 provides a step-by-step overview of the before/after comparison-group safety effectiveness evaluation method.

Data Needs and Inputs

The data needed as input to a before/after comparison-group evaluation include:

- At least 10 to 20 sites at which the treatment of interest has been implemented.
- At least 10 to 20 comparable sites at which the treatment has not been implemented and that have not had other major changes during the evaluation study period.

- A minimum of 650 aggregate crashes at the comparable sites at which the treatment has not been implemented.
- 3 to 5 years of crash data for the period before treatment implementation is recommended for both treatment and nontreatment sites.
- 3 to 5 years of crash data for the period after treatment implementation is recommended for both treatment and nontreatment sites.
- SPFs for treatment and nontreament sites.

An evaluation study can be performed with fewer sites or shorter time periods, or both, but statistically significant results are less likely.

Pre-Evaluation Activities
The key pre-evaluation activities are to:

- Identify the treatment sites to be evaluated.
- Select the time periods before and after treatment implementation for each site that will be included in the evaluation.
- Select the measure of effectiveness for the evaluation. Evaluations often use total crash frequency as the measure of effectiveness, but any specific crash severity level or crash type, or both can be considered.
- Select a set of comparison sites that are comparable to the treatment sites
- Assemble the required crash and traffic volume data for each site and time period of interest, including both treatment and comparison sites.
- Obtain SPF(s) applicable to the treatment and comparison sites. Such SPFs may be developed based on the available data as described in Part C or from SafetyAnalyst. In a comparison-group evaluation, the SPF(s) are used solely to derive adjustment factors to account for the nonlinear effects of changes in average daily traffic volume. This adjustment for changes in traffic volume is needed for both the treatment and comparison sites and, therefore, SPFs are needed for all site types included in the treatment and comparison sites. If no SPFs are available and the effects of traffic volume are assumed to be linear, this will make the evaluation results less accurate.

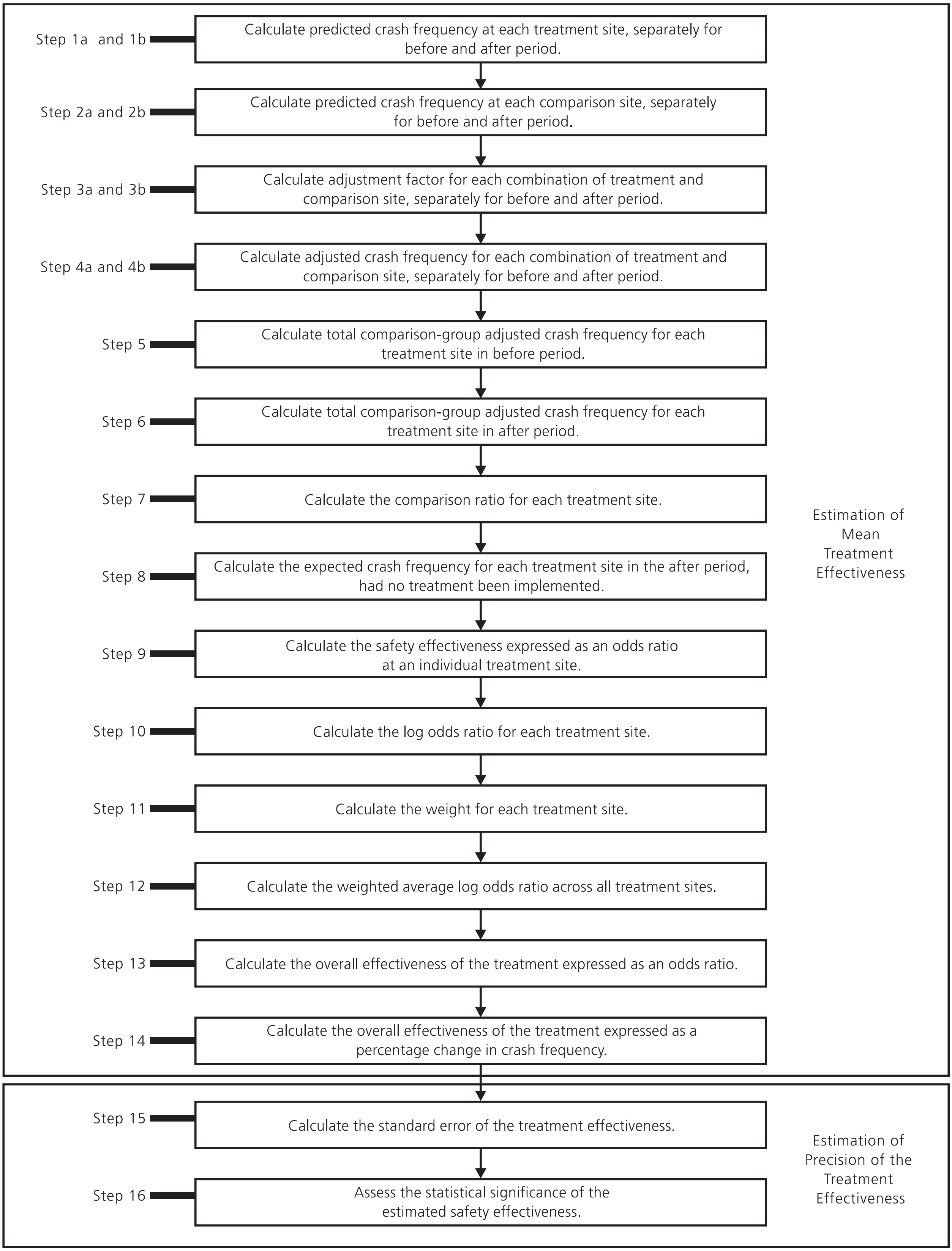


Figure 9-3. Overview of Before/After Comparison-Group Safety Evaluation

The before study period for a site must end before implementation of the treatment began at that site. The after study period for a site normally begins after treatment implementation is complete; a buffer period of several months is usually allowed for traffic to adjust to the presence of the treatment. Evaluation periods that are even multiples of 12 months in length are used so that there is no seasonal bias in the evaluation data. Analysts often choose evaluation periods that consist of complete calendar years because this often makes it easier to assemble the required data. When the evaluation periods consist of entire calendar years, the entire year during which the treatment was installed is normally excluded from the evaluation period.

The comparison-group procedures are based on the assumption that the same set of comparison-group sites are used for all treatment sites. A variation of the procedure that is applicable if different comparison-group sites are used for each treatment is presented by Harwood et al. (2). Generally, this variation would only be needed for special cases, such as multi-state studies where an in-state comparison group was used for each treatment site.

A weakness of the comparison-group method is that it cannot consider treatment sites at which the observed crash frequency in the period either before or after implementation of the treatment is zero. This may lead to an underestimate of the treatment effectiveness since sites with no crashes in the after treatment may represent locations at which the treatment was most effective.

Computational Procedure

A computational procedure using the comparison-group evaluation study method to determine the effectiveness of the treatment being evaluated, expressed as a percentage change in crashes, θ, and to assess its precision and statistical significance, is presented in the Appendix 9A.

9.4.3. Implementing the Safety Evaluation Method for Before/After Shifts in Proportions of Target Collision Types

The safety evaluation method for before/after shifts in proportions is used to quantify and assess the statistical significance of a change in the frequency of a specific target collision type expressed as a proportion of total crashes from before to after implementation of a specific countermeasure or treatment. This method uses data only for treatment sites and does not require data for nontreatment or comparison sites. Target collision types (e.g., run-off-the-road, head-on, rear-end) addressed by the method may include all crash severity levels or only specific crash severity levels (fatal-and-serious-injury crashes, fatal-and-injury-crashes, or property-damage-only crashes). Figure 9-4 provides a step-by-step overview of the method for conducting a before/after safety effectiveness evaluation for shifts in proportions of target collision types.

Data Needs and Inputs

The data needed as input to a before/after evaluation for shifts in proportions of target collision types include:

- At least 10 to 20 sites at which the treatment of interest has been implemented.
- 3 to 5 years of before-period crash data is recommended for the treatment sites.
- 3 to 5 years of after-period crash data is recommended for the treatment sites.

An evaluation study can be performed with fewer sites or shorter time periods, or both, but statistically significant results are less likely.

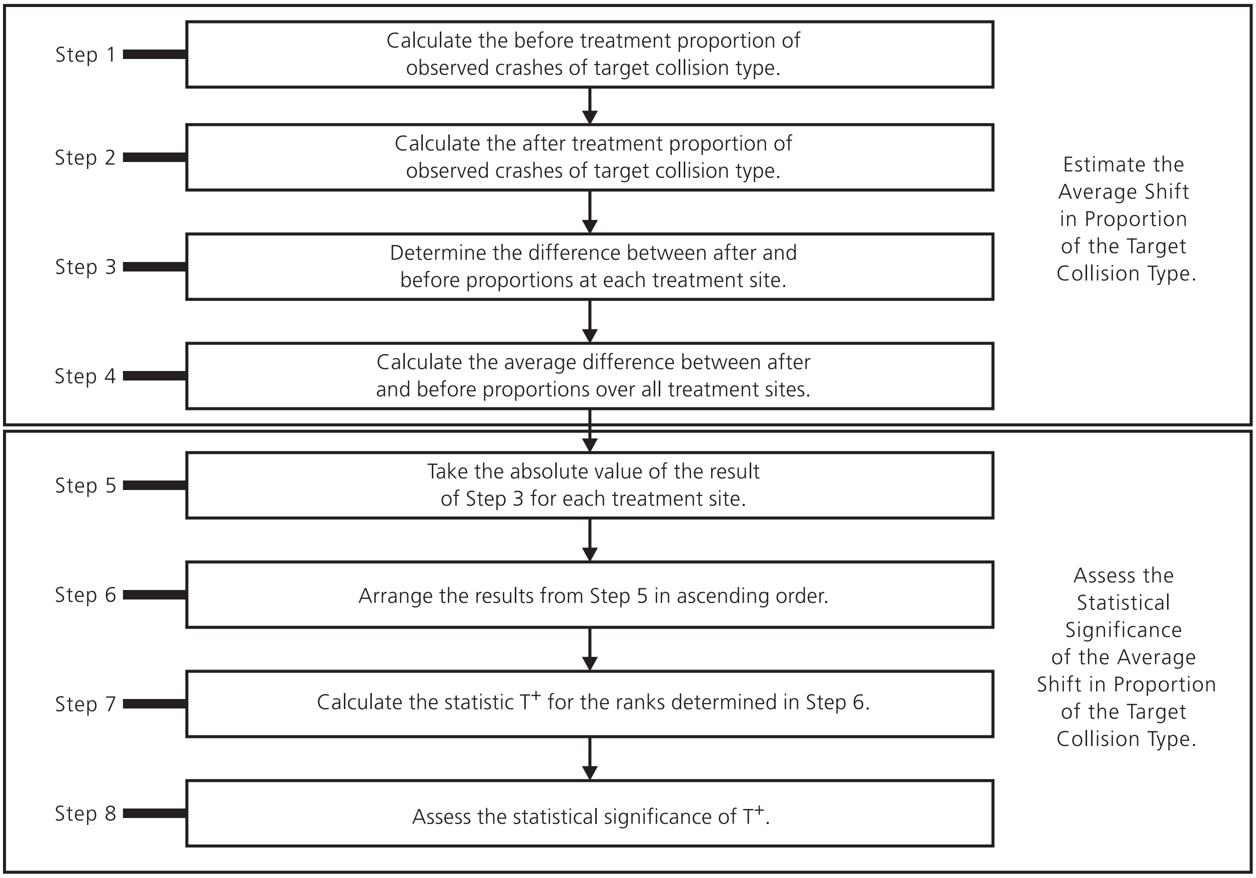


Figure 9-4. Overview Safety Evaluation for Before/After Shifts in Proportions

Pre-Evaluation Activities

The key pre-evaluation activities are to:

- Identify the treatment sites to be evaluated.
- Select the time periods before and after treatment implementation for each site that will be included in the evaluation.
- Select the target collision type for the evaluation.
- Assemble the required crash and traffic volume data for each site and time period of interest for the treatment sites.

The before study period for a site must end before implementation of the treatment began at that site. The after study period for a site normally begins after treatment implementation is complete; a buffer period of several months is usually allowed for traffic to adjust to the presence of the treatment. Evaluation periods that are even multiples of 12 months in length are used so that there is no seasonal bias in the evaluation data. Analysts often choose evaluation periods that consist of complete calendar years because this often makes it easier to assemble the required data. When the evaluation periods consist of entire calendar years, the entire year during which the treatment was installed is normally excluded from the evaluation period.

Computational Method

A computational procedure using the evaluation study method for assessing shifts in proportions of target collision types to determine the safety effectiveness of the treatment being evaluated, $AvgP_{(CT)diff}$, and to assess its statistical significance, is presented in Appendix 9A.

9.4.4. Implementing the Cross-Sectional Safety Evaluation Method

Definition

In the absence of before data at treatment sites, the cross-sectional safety evaluation method can be used to estimate the safety effectiveness of a treatment through comparison to crash data at comparable nontreatment sites. A cross-sectional safety evaluation generally requires complex statistical modeling and therefore is addressed here in general terms only.

Data Needs and Inputs

- 10 to 20 treatment sites are recommended to evaluate a safety treatment.
- 10 to 20 nontreatment sites are recommended for the nontreatment group.
- 3 to 5 years of crash data for both treatment and nontreatment sites is recommended.

Pre-Evaluation Activities

The key pre-evaluation activities are to:

- Identify the sites both with and without the treatment to be evaluated.
- Select the time periods that will be included in the evaluation when the conditions of interest existed at the treatment and nontreatment sites.
- Select the safety measure of effectiveness for the evaluation. Evaluations often use total crash frequency as the measure of effectiveness, but any specific crash severity level or crash type, or both, can be considered.
- Assemble the required crash and traffic volume data for each site and time period of interest.

Method

There is no step-by-step methodology for the cross-sectional safety evaluation method because this method requires model development rather than a sequence of computations that can be presented in equations. In implementing the cross-sectional safety evaluation method, all of the crash, traffic volume, and site characteristics data (including data for both the treatment and nontreatment sites) are analyzed in a single model including either an indicator variable for the presence or absence of the treatment at a site or a continuous variable representing the dimension of the treatment (e.g., lane width or shoulder width). A generalized linear model (GLM) with a negative binomial distribution and a logarithmic link function is a standard approach to model the yearly crash frequencies. Generally, a repeated-measures correlation structure is included to account for the relationship between crashes at a given site across years (temporal correlation). A compound symmetry, autoregressive, or other covariance structure can be used to account for within-site correlation. General estimating equations (GEE) may then be used to determine the final regression parameter estimates, including an estimate of the treatment effectiveness and its precision. An example of application of this statistical modeling approach is presented by Lord and Persaud (8). This approach may be implemented using any of several commercially available software packages.

The example below illustrates a generic application of a cross-sectional safety evaluation analysis.

Overview of a Cross-Sectional Analysis to Evaluate the Safety Effectiveness of a Treatment

A treatment was installed at 11 sites. Crash data, geometrics, and traffic volume data are available for a 4-year period at each site. Similar data are available for 9 sites without the treatment but with comparable geometrics and traffic volumes. The available data can be summarized as follows:

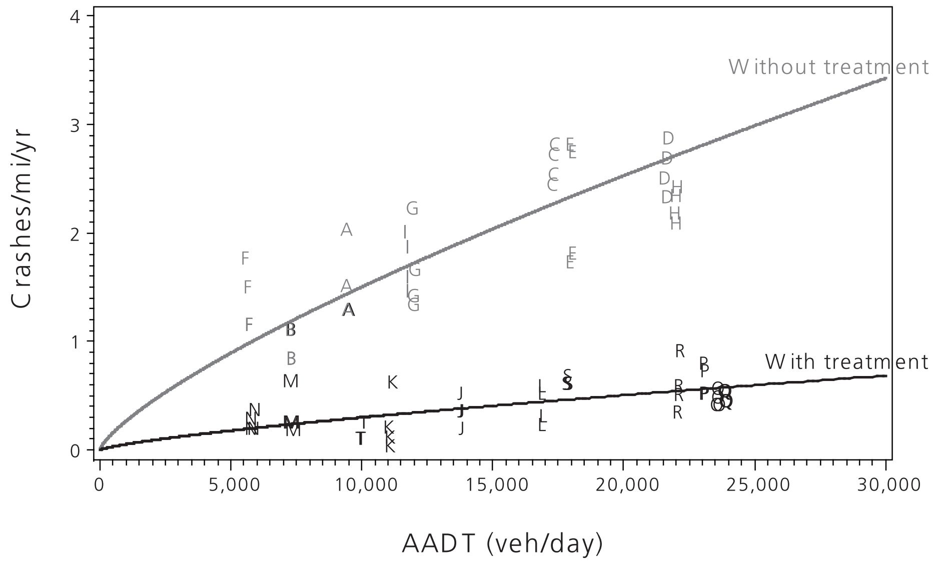
- 9 nontreatment sites (denoted A through I); 4 years of data at each site
- 11 treatment sites (denoted J through T); 4 years of data at each site

A negative binomial generalized linear model (GLM) was used to estimate the treatment effect based on the entire dataset, accounting for AADT and other geometric parameters (e.g., shoulder width, lane width, number of lanes, roadside hazard rating) as well as the relationship between crashes at a given site over the 4-year period (within-site correlation) using generalized estimating equations (GEE).

The graph illustrates the observed and predicted average crash frequency for the treatment and nontreatment sites. The safety effectiveness of the treatment is assessed by the statistical significance of the treatment effect on crash frequency. This effect is illustrated by the difference in the rate of change in the two curves. In this example, the installation of the treatment significantly reduced crash frequency.

Note that the data shown below are fictional crash and traffic data.

Observed and Predicted Crash Frequencies at Treatment and Nontreatment Sites



9.5. EVALUATING A SINGLE PROJECT AT A SPECIFIC SITE TO DETERMINE ITS SAFETY EFFECTIVENESS

An observational before/after evaluation can be conducted for a single project at a specific site to determine its effectiveness in reducing crash frequency or severity. The evaluation results provide an estimate of the effect of the project on safety at that particular site. Any of the study designs and evaluation methods presented in Sections 9.3 and 9.4, with the exception of cross-sectional studies which require more than one treatment site, can be applied to such an evaluation. The results of such evaluations, even for a single site, may be of interest to highway agencies in monitoring their improvement programs. However, results from the evaluation of a single site will not be very accurate and, with only one site available, the precision and statistical significance of the evaluation results cannot be assessed.

9.6. EVALUATING A GROUP OF SIMILAR PROJECTS TO DETERMINE THEIR SAFETY EFFECTIVENESS

Observational before/after evaluations can be conducted for groups of similar projects to determine their effectiveness reducing crash frequency or severity. The evaluation results provide an estimate of the overall safety effectiveness of the group of projects as a whole. Any of the study designs and evaluation methods presented in Sections 9.3 and 9.4, with the exception of cross-sectional studies, can be applied to such an evaluation. Cross-sectional studies are intended to make inferences about the effectiveness of a countermeasure or treatment when applied to other sites, not to evaluate the safety effectiveness of projects at particular sites. Therefore, cross-sectional studies are not appropriate when the objective of the evaluation is to assess the effectiveness of the projects themselves.

A safety effectiveness evaluation for a group of projects may be of interest to highway agencies in monitoring their improvement programs. Where more than one project is evaluated, the precision of the effectiveness estimate and the statistical significance of the evaluation results can be determined. The guidelines in Section 9.4 indicate that at least 10 to 20 sites generally need to be evaluated to obtain statistically significant results. While this minimum number of sites is presented as a general guideline, the actual number of sites needed to obtain statistically significant results can vary widely as a function of the magnitude of the safety effectiveness for the projects being evaluated and the site-to-site variability of the effect. The most reliable methods for evaluating a group of projects are those that compensate for regression-to-the-mean bias, such as the EB method.

9.7. QUANTIFYING CMFS AS A RESULT OF A SAFETY EFFECTIVENESS EVALUATION

A common application of safety effectiveness evaluation is to quantify the value of a CMF for a countermeasure by evaluating multiple sites where that countermeasure has been evaluated. The relationship between a CMF and safety effectiveness is given as CMF = (100 – Safety Effectiveness/100). Any of the study designs and evaluation methods presented in Sections 9.3 and 9.4 can be applied in quantifying a CMF value, although methods that compensate for regression-to-the-mean bias, such as the EB method, are the most reliable. The evaluation methods that can be used to quantify a CMF are the same as those described in Section 9.6 for evaluating a group of projects, except the cross-sectional studies may also be used, though they are less reliable than methods that compensate for regression-to-the- mean bias. As noted above, at least 10 to 20 sites generally need to be evaluated to obtain statistically significant results. While this minimum number of sites is presented as a general guideline, the actual number of sites needed to obtain statistically significant results can vary widely as a function of the magnitude of the safety effectiveness for the projects being evaluated and the site-to-site variability of the effect.

9.8. COMPARISON OF SAFETY BENEFITS AND COSTS OF IMPLEMENTED PROJECTS

Where the objective of an evaluation is to compare the crash reduction benefits and costs of implemented projects, the first step is to determine a CMF for the project, as described above in Section 9.7. The economic analysis procedures presented in Chapter 7 are then applied to quantify the safety benefits of the projects in monetary terms, using the CMF, and to compare the safety benefits and costs of the implemented projects. Figure 9-5 provides a graphical overview of this comparison.

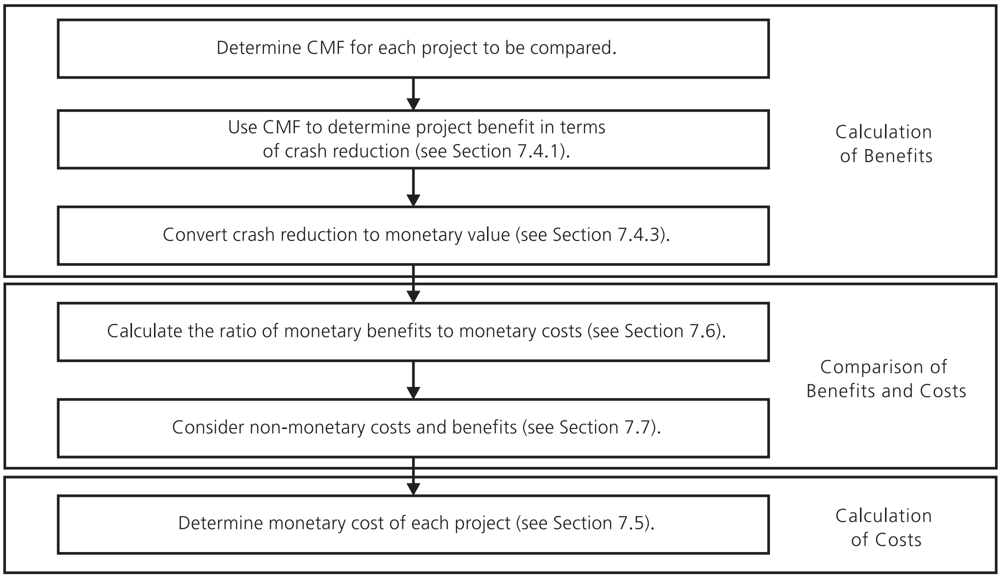


Figure 9-5. Overview of Safety Benefits and Costs Comparison of Implemented Projects

9.9. CONCLUSIONS

Safety effectiveness evaluation is the process of developing quantitative estimates of the reduction in the number of crashes or severity of crashes due to a treatment, project, or a group of projects. Evaluating implemented safety treatments is an important step in the roadway safety evaluation process, and provides important information for future decision making and policy development.

Safety effectiveness evaluation may include:

- Evaluating a single project at a specific site to document the safety effectiveness of that specific project,
- Evaluating a group of similar projects to document the safety effectiveness of those projects,
- Evaluating a group of similar projects for the specific purpose of quantifying a CMF for a countermeasure, and
- Assessing the overall safety effectiveness of specific types of projects or countermeasures in comparison to their costs.

There are three basic study designs that can be used for safety effectiveness evaluations:

- Observational before/after studies
- Observational cross-sectional studies
- Experimental before/after studies

Both observational and experimental studies may be used in safety effectiveness evaluations, although observational studies are more common among highway agencies.

This chapter documents and discusses the various methods for evaluating the effectiveness of a treatment, a set of treatments, an individual project, or a group of similar projects after safety improvements have been implemented. This chapter provides an introduction to the evaluation methods that can be used, highlights which methods are appropriate for assessing safety effectiveness in specific situations, and provides step-by-step procedures for conducting safety effectiveness evaluations.

9.10. SAMPLE PROBLEM TO ILLUSTRATE THE EB BEFORE/AFTER SAFETY EFFECTIVENESS EVALUATION METHOD

This section presents sample problems corresponding to the three observational before/after safety effectiveness evaluation methods presented in Chapter 9, including the EB method, the comparison-group method, and the shift in proportions method. The data used in these sample problems are hypothetical. Appendix 9A provides a detailed summary of the steps for each of these methods.

Passing lanes have been installed to increase passing opportunities at 13 rural two-lane highway sites. An evaluation is to be conducted to determine the overall effect of the installation of these passing lanes on total crashes at the 13 treatment sites.

Data for total crash frequencies are available for these sites, including five years of data before and two years of data after installation of the passing lanes. Other available data include the site length (L) and the before- and after-period traffic volumes. To simplify the calculations for this sample problem, AADT is assumed to be constant across all years for both the before and after periods. It is also assumed that the roadway characteristics match base conditions and, therefore, all applicable CMFs as well as the calibration factor (see Chapter 10) are equal to 1.0.

Column numbers are shown in the first row of all the tables in this sample problem; the description of the calculations refers to these column numbers for clarity of explanation. For example, the text may indicate that Column 10 is the sum of Columns 5 through 9 or that Column 13 is the sum of Columns 11 and 12. When columns are repeated from table to table, the original column number is kept. Where appropriate, column totals are indicated in the last row of each table.

9.10.1. Basic Input Data

The basic input data for the safety effectiveness evaluation, including the yearly observed before- and after-period crash data for the 13 rural two-lane road segments, are presented below:

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Site length	AADT (veh/day)		Observed before total crash frequency by year (crashes/site/year)					Observed crash frequency in before period	Observed after total crash frequency by year (crashes/site/year)		Observed crash frequency in after period
Site No.	(L) (mi)	Before	After	Y1	Y2	Y3	Y4	Y5		Y1	Y2	
1	1.114	8,858	8,832	4	4	1	5	2	16	1	1	2
2	0.880	11,190	11,156	2	0	0	2	2	6	0	2	2
3	0.479	11,190	11,156	1	0	2	1	0	4	1	1	2
4	1.000	6,408	6,388	2	5	4	3	2	16	0	1	1
5	0.459	6,402	6,382	0	0	1	0	0	1	0	1	1
6	0.500	6,268	6,250	1	1	0	2	1	5	1	0	1
7	0.987	6,268	6,250	4	3	3	4	3	17	6	3	9
8	0.710	5,503	5,061	4	3	1	1	3	12	0	0	0
9	0.880	5,523	5,024	2	0	6	0	0	8	0	0	0
10	0.720	5,523	5,024	1	0	1	1	0	3	0	0	0
11	0.780	5,523	5,024	1	4	2	1	1	9	3	2	5
12	1.110	5,523	5,024	1	0	2	4	2	9	4	2	6
13	0.920	5,523	5,024	3	2	3	3	5	16	0	1	1
Total				26	22	26	27	21	122	16	14	30

9.10.2. EB Estimation of the Expected Average Crash Frequency in the Before Period

Equation 10-6 provides the applicable SPF to predict total crashes on rural two-lane roads:

$$N_{spf\,rs} = \text{AADT} \times L \times 365 \times 10^{-6} \times e^{(-0.312)} \quad (10\text{–}6)$$

Where:

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

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

L = length of roadway segment (miles).

The overdispersion parameter is given by Equation 10-7 as:

$$k = \frac{0.236}{L} \quad (10\text{–}7)$$

Equation 10-1 presents the predicted average crash frequency for a specific site type x (roadway, rs, in this example). Note in this example all CMFs and the calibration factor are assumed to equal 1.0.

$$N_{\text{predicted}} = N_{spf\,x} \times (CMF_{1x} \times CMF_{2x} \times \ldots \times CMF_{yx}) \times C_x \quad (10\text{-}1)$$

Where:

$N_{predicted}$ = predicted average crash frequency for a specific year for site type x;

$N_{spf\,x}$ = predicted average crash frequency determined for base conditions of the SPF developed for site type x;

CMF_{yx} = Crash Modification Factors specific to site type x and specific geometric design and traffic control features y;

C_x = calibration factor to adjust SPF for local conditions for site type x.

Step 1—Using the above SPF and Columns 2 and 3, calculate the predicted average crash frequency for each site during each year of the before period.

Using the above SPF and Columns 2 and 3, calculate the predicted average crash frequency for each site during each year of the before period. The results appear in Columns 14 through 18. For use in later calculations, sum these predicted average crash frequencies over the five before years. The results appear in Column 19. Note that because in this example the AADT is assumed constant across years at a given site in the before period, the predicted average crash frequencies do not change from year to year since they are simply a function of segment length and AADT at a given site. This will not be the case in general, when yearly AADT data are available.

(1)	(14)	(15)	(16)	(17)	(18)	(19)
	Predicted before total crash frequency by year (crashes/year)					Predicted average crash frequency in before period
Site No.	Y1	Y2	Y3	Y4	Y5	
1	2.64	2.64	2.64	2.64	2.64	13.18
2	2.63	2.63	2.63	2.63	2.63	13.15
3	1.43	1.43	1.43	1.43	1.43	7.16
4	1.71	1.71	1.71	1.71	1.71	8.56
5	0.79	0.79	0.79	0.79	0.79	3.93
6	0.84	0.84	0.84	0.84	0.84	4.19
7	1.65	1.65	1.65	1.65	1.65	8.26
8	1.04	1.04	1.04	1.04	1.04	5.22
9	1.30	1.30	1.30	1.30	1.30	6.49
10	1.06	1.06	1.06	1.06	1.06	5.31
11	1.15	1.15	1.15	1.15	1.15	5.75
12	1.64	1.64	1.64	1.64	1.64	8.19
13	1.36	1.36	1.36	1.36	1.36	6.79
Total	19.24	19.24	19.24	19.24	19.24	96.19

Step 2—Calculate the Weighted Adjustment, *w*, for each site for the before period.

Using Equation 9A.1-2, the calculated overdispersion parameter (shown in Column 20), and Column 19 (Step 1), calculate the weighted adjustment, w, for each site for the before period. The results appear in Column 21. Using Equation 9A.1-1, Columns 21, 19 (Step 1), and 10 (Basic Input Data), calculate the expected average crash frequency for each site, summed over the entire before period. The results appear in Column 22.

(1)	(20)	(21)	(22)
Site No.	Overdispersion parameter, *k*	Weighted adjustment, *w*	Expected average crash frequency in before period
1	0.212	0.264	15.26
2	0.268	0.221	7.58
3	0.493	0.221	4.70
4	0.236	0.331	13.54
5	0.514	0.331	1.97
6	0.472	0.336	4.73
7	0.239	0.336	14.06
8	0.332	0.366	9.52
9	0.268	0.365	7.45
10	0.328	0.365	3.84
11	0.303	0.365	7.82
12	0.213	0.365	8.70
13	0.257	0.365	12.64
Total			111.81

9.10.3. EB Estimation of the Expected Average Crash Frequency in the After Period in the Absence of the Treatment

Step 3—Calculate the Predicted Average Crash Frequency for each site during each year of the after period. Using the above SPF and Columns 2 and 4, calculate the predicted average crash frequency for each site during each year of the after period. The results appear in Columns 23 and 24. For use in later calculations, sum these predicted average crash frequencies over the two after years. The results appear in Column 25.

(1)	(23)	(24)	(25)	(26)	(27)
	Predicted after total crash frequency (crashes/year)				
Site No.	Y1	Y2	Predicted average crash frequency in after period	Adjustment factor, *r*	Expected average crash frequency in after period without treatment
1	2.63	2.63	5.26	0.399	6.08
2	2.62	2.62	5.25	0.399	3.02
3	1.43	1.43	2.86	0.399	1.87
4	1.71	1.71	3.41	0.399	5.40
5	0.78	0.78	1.57	0.399	0.79
6	0.83	0.83	1.67	0.399	1.89
7	1.65	1.65	3.30	0.399	5.61
8	0.96	0.96	1.92	0.368	3.50
9	1.18	1.18	2.36	0.364	2.71
10	0.97	0.97	1.93	0.364	1.40
11	1.05	1.05	2.09	0.364	2.84
12	1.49	1.49	2.98	0.364	3.17
13	1.23	1.23	2.47	0.364	4.60
Total	18.53	18.53	37.06		42.88

Step 4—Calculate the Adjustment Factor, *r*, to account for the differences between the before and after periods in duration and traffic volume at each site.
Using Equation 9A.1-3 and Columns 25 and 19, calculate the adjustment factor, *r*, to account for the differences between the before and after periods in duration and traffic volume at each site. The results appear in Column 26 in the table presented in Step 3.

Step 5—Calculate the Expected Average Crash Frequency for each Site over the Entire after Period in the Absence of the Treatment.
Using Equation 9A.1-4 and Columns 22 and 26, calculate the expected average crash frequency for each site over the entire after period in the absence of the treatment. The results appear in Column 27 in the table presented in Step 3.

9.10.4. Estimation of the Treatment Effectiveness

Step 6—Calculate an Estimate of the Safety Effectiveness of the Treatment at each site in the form of an odds ratio.
Using Equation 9A.1-5 and Columns 13 and 27, calculate an estimate of the safety effectiveness of the treatment at each site in the form of an odds ratio. The results appear in Column 28.

(1)	(13)	(27)	(28)	(29)	(30)
Site No.	Observed crash frequency in after period	Expected average crash frequency in after period without treatment	Odds ratio	Safety effectiveness (%)	Variance term (Eq. 9A.1-11)
1	2	6.08	0.329	67.13	1.787
2	2	3.02	0.662	33.84	0.939
3	2	1.87	1.068	–6.75	0.582
4	1	5.40	0.185	81.47	1.440
5	1	0.79	1.274	–27.35	0.209
6	1	1.89	0.530	46.96	0.499
7	9	5.61	1.604	–60.44	1.486
8	0	3.50	0.000	100.00	0.817
9	0	2.71	0.000	100.00	0.627
10	0	1.40	0.000	100.00	0.323
11	5	2.84	1.758	–75.81	0.657
12	6	3.17	1.894	–89.44	0.732
13	1	4.60	0.217	78.26	1.063
Total	30	42.88			11.162

Step 7—Calculate the Safety Effectiveness as a percentage crash change at each site.
Using Equation 9A.1-6 and Column 28, calculate the safety effectiveness as a percentage crash change at each site. The results appear in Column 29 in the table presented in Step 6. A positive result indicates a reduction in crashes; conversely, a negative result indicates an increase in crashes.

Step 8—Calculate the Overall Effectiveness of the Treatment for all sites combined, in the form of an odds ratio.
Using Equation 9A.1-7 and the totals from Columns 13 and 27 (Step 6), calculate the overall effectiveness of the treatment for all sites combined, in the form of an odds ratio:

$$OR' = \frac{30}{42.88} = 0.700$$

Step 9—Calculate each Term of Equation 9A.1-9.

Using Columns 26 (Step 3), 22 (Step 2), and 21 (Step 2), calculate each term of Equation 9A.1-9. The results appear in Column 30 in the table presented in Step 6. Sum the terms in Column 30. Next, using Equations 9A.1-8 and 9A.1-9, the value for *OR'* from Step 8, and the sums in Column 30 and 27 in Step 6, calculate the final adjusted odds ratio:

$$OR = \frac{0.700}{1+\frac{11.162}{(42.88)^2}} = 0.695$$

Since the odds ratio is less than 1, it indicates a reduction in crash frequency due to the treatment.

Step 10—Calculate the Overall Unbiased Safety Effectiveness as a percentage change in crash frequency across all sites.

Using Equation 9A.1-10 and the above result, calculate the overall unbiased safety effectiveness as a percentage change in crash frequency across all sites:

Safety Effectiveness = 100 × (1 – 0.695) = 30.5%

9.10.5. Estimation of the Precision of the Treatment Effectiveness

Step 11—Calculate the Variance of *OR*.

Using Equation 9A.1-11, the value for *OR'* from Step 8, and the sums from Columns 13, 30, and 27 in Step 6, calculate the variance of *OR*:

$$Var(OR) = \frac{(0.700)^2\left[\frac{1}{30}+\frac{11.162}{(42.88)^2}\right]}{\left[1+\frac{11.162}{(42.88)^2}\right]} = 0.019$$

Step 12—Calculate the Standard Error of *OR*.

Using Equation 9A.1-12 and the result from Step 11, calculate the standard error of *OR*:

$$SE(OR) = \sqrt{0.019} = 0.138$$

Step 13—Calculate the Standard Error of the Safety Effectiveness.

Using Equation 9A.1-13 and the result from Step 12, calculate the standard error of the Safety Effectiveness:

SE(Safety Effectiveness) =100 × 0.138 = 13.8%

Step 14—Assess the Statistical Significance of the Estimated Safety Effectiveness.

Assess the statistical significance of the estimated safety effectiveness by calculating the quantity:

$$Abs\frac{\text{Safety Effectiveness}}{SE\left(\text{Safety Effectiveness}\right)} = \frac{30.5}{13.85} = 2.20$$

Since *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] ≥ 2.0, conclude that the treatment effect is significant at the (approximate) 95 percent confidence level. The positive estimate of Safety Effectiveness, 30.5 percent, indicates a positive effectiveness, i.e., a reduction, in total crash frequency.

In summary, the evaluation results indicate that the installation of passing lanes at the 13 rural two-lane highway sites reduced total crash frequency by 30.5 percent on average, and that this result is statistically significant at the 95 percent confidence level.

9.11. SAMPLE PROBLEM TO ILLUSTRATE THE COMPARISON-GROUP SAFETY EFFECTIVENESS EVALUATION METHOD

Passing lanes have been installed to increase passing opportunities at 13 rural two-lane highway sites. An evaluation is to be conducted to determine the overall effect of the installation of these passing lanes on total crashes at the 13 treatment sites.

9.11.1. Basic Input Data for Treatment Sites

Data for total crash frequencies are available for the 13 sites, including five years of data before and two years of data after installation of the passing lanes. Other available data include the site length (*L*) and the before- and after-period traffic volumes. To simplify the calculations for this sample problem, AADT is assumed to be constant across all years for both the before and after periods. The detailed step-by-step procedures in Appendix 9A show how to handle computations for sites with AADTs that vary from year to year.

Column numbers are shown in the first row of all the tables in this sample problem; the description of the calculations refers to these column numbers for clarity of explanation. When columns are repeated from table to table, the original column number is kept. Where appropriate, column totals are indicated in the last row of each table.

Organize the observed before- and after-period data for the 13 rural two-lane road segments as shown below based on the input data for the treatment sites shown in the sample problem in Section 9.10:

(1)	(2)	(3)	(4)	(5)	(6)
Treatment Sites					
		AADT (veh/day)		Observed crash frequency in before Period (5 years) ($N_{observed}$)	Observed crash frequency in after period (2 years) (*L*)
Site No.	Site length (*L*) (mi)	Before	After		
1	1.114	8,858	8,832	16	2
2	0.880	11,190	11,156	6	2
3	0.479	11,190	11,156	4	2
4	1.000	6,408	6,388	16	1
5	0.459	6,402	6,382	1	1
6	0.500	6,268	6,250	5	1
7	0.987	6,268	6,250	17	9
8	0.710	5,503	5,061	12	0
9	0.880	5,523	5,024	8	0
10	0.720	5,523	5,024	3	0
11	0.780	5,523	5,024	9	5
12	1.110	5,523	5,024	9	6
13	0.920	5,523	5,024	16	1
Total	10.539			122	30

9.11.2. Basic Input Data for Comparison-Group Sites

A comparison group of 15 similar, but untreated, rural two-lane highway sites has been selected. The length of each site is known. Seven years of before-period data and three years of after-period data (crash frequencies and before- and after-period AADTs) are available for each of the 15 sites in the comparison group. As above, AADT is assumed to be constant across all years in both the before and after periods for each comparison site. The same comparison group is assigned to each treatment site in this sample problem.

Organize the observed before- and after-period data for the 15 rural two-lane road segments as shown below:

(7)	(8)	(9)	(10)	(11)	(12)
Comparison Group					
		AADT (veh/day)			
Site No.	Site length (L) (mi)	Before	After	Observed crash frequency in before period (7 years)	Observed crash frequency in after period (3 years)
1	1.146	8,927	8,868	27	4
2	1.014	11,288	11,201	5	5
3	0.502	11,253	11,163	7	3
4	1.193	6,504	6,415	21	2
5	0.525	6,481	6,455	3	0
6	0.623	6,300	6,273	6	1
7	1.135	6,341	6,334	26	11
8	0.859	5,468	5,385	12	4
9	1.155	5,375	5,324	20	12
10	0.908	5,582	5,149	33	5
11	1.080	5,597	5,096	5	0
12	0.808	5,602	5,054	3	0
13	0.858	5,590	5,033	4	10
14	1.161	5,530	5,043	12	2
15	1.038	5,620	5,078	21	2
Total	14.004			205	61

9.11.3. Estimation of Mean Treatment Effectiveness

Equation 10-6 provides the applicable SPF for total crashes on rural two-lane roads:

$$N_{spf\,rs} = \text{AADT} \times L \times 365 \times 10^{-6} \times e^{(-0.312)} \qquad (10\text{–}6)$$

The overdispersion parameter for this SPF is not relevant to the comparison-group method.

Equation 10-1 presents the predicted average crash frequency for a specific site type x (roadway, rs, in this example). Note in this example all CMFs and the calibration factor are assumed to equal 1.0.

$$N_{\text{predicted}} = N_{spf\,x} \times (CMF_{1x} \times CMF_{2x} \times \ldots \times CMF_{yx}) \times C_x \qquad (10\text{-}1)$$

Where:

$N_{\text{predicted}}$ = predicted average crash frequency for a specific year for site type x;

$N_{spf\,x}$ = predicted average crash frequency determined for base conditions of the *SPF* developed for site type x;

CMF_{yx} = Crash Modification Factors specific to site type x and specific geometric design and traffic control features y;

C_x = calibration factor to adjust *SPF* for local conditions for site type x.

Step 1a—Calculate the Predicted Average Crash Frequency at each treatment site in the 5-year before period. Using the above SPF and Columns 2 and 3, calculate the predicted average crash frequency at each treatment site in the 5-year before period. The results appear in Column 13 in the table below. For use in later calculations, sum these predicted average crash frequencies over the 13 treatment sites.

Step 1b—Calculate the predicted average crash frequency at each treatment site in the 2-year after period. Similarly, using the above SPF and Columns 2 and 4, calculate the predicted average crash frequency at each treatment site in the 2-year after period. The results appear in Column 14. Sum these predicted average crash frequencies over the 13 treatment sites.

(1)	(13)	(14)
	Treatment Sites	
Site No.	Predicted average crash frequency at treatment site in before period (5 years)	Predicted average crash frequency at treatment site in after period (2 years)
1	13.18	5.26
2	13.15	5.25
3	7.16	2.86
4	8.56	3.41
5	3.93	1.57
6	4.19	1.67
7	8.26	3.30
8	5.22	1.92
9	6.49	2.36
10	5.31	1.93
11	5.75	2.09
12	8.19	2.98
13	6.79	2.47
Total	96.19	37.06

Step 2a—Calculate the Predicted Average Crash Frequency for each comparison site in the 7-year before period. Using the above SPF and Columns 8 and 9, calculate the predicted average crash frequency for each comparison site in the 7-year before period. The results appear in Column 15 in the table below. Sum these predicted average crash frequencies over the 15 comparison sites.

Step 2b—Calculate the Predicted Average Crash Frequency for each comparison site in the 3-year after period. Similarly, using the above SPF and Columns 8 and 10, calculate the predicted average crash frequency for each comparison site in the 3-year after period. The results appear in Column 16. Sum these predicted average crash frequencies over the 15 comparison sites.

(7)	(15)	(16)
	Comparison Group	
Site No.	Predicted average crash frequency at comparison site in before period (7 years)	Predicted average crash frequency at comparison site in after period (3 years)
1	19.13	8.14
2	21.40	9.10
3	10.56	4.49
4	14.51	6.13
5	6.37	2.72
6	7.34	3.13
7	13.46	5.76
8	8.79	3.71
9	11.62	4.93
10	9.48	3.75
11	11.30	4.41
12	8.46	3.27
13	8.97	3.46
14	12.01	4.69
15	10.91	4.22
Total	174.29	71.93

Step 3a—Calculate the 13 Before Adjustment Factors for each of the 15 comparison sites.
Using Equation 9A.2-1, Columns 13 and 15, the number of before years for the treatment sites (5 years), and the number of before years for the comparison sites (7 years), calculate the 13 before adjustment factors for each of the 15 comparison sites. The results appear in Columns 17 through 29.

(7)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)
	Comparison Group—Before Adjustment Factors (Equation 9A.2-1)												
Site No.	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.49	0.49	0.27	0.32	0.15	0.16	0.31	0.19	0.24	0.20	0.21	0.31	0.25
2	0.44	0.44	0.24	0.29	0.13	0.14	0.28	0.17	0.22	0.18	0.19	0.27	0.23
3	0.89	0.89	0.48	0.58	0.27	0.28	0.56	0.35	0.44	0.36	0.39	0.55	0.46
4	0.65	0.65	0.35	0.42	0.19	0.21	0.41	0.26	0.32	0.26	0.28	0.40	0.33
5	1.48	1.48	0.80	0.96	0.44	0.47	0.93	0.59	0.73	0.60	0.65	0.92	0.76
6	1.28	1.28	0.70	0.83	0.38	0.41	0.80	0.51	0.63	0.52	0.56	0.80	0.66
7	0.70	0.70	0.38	0.45	0.21	0.22	0.44	0.28	0.34	0.28	0.31	0.43	0.36
8	1.07	1.07	0.58	0.70	0.32	0.34	0.67	0.42	0.53	0.43	0.47	0.67	0.55
9	0.81	0.81	0.44	0.53	0.24	0.26	0.51	0.32	0.40	0.33	0.35	0.50	0.42
10	0.99	0.99	0.54	0.65	0.30	0.32	0.62	0.39	0.49	0.40	0.43	0.62	0.51
11	0.83	0.83	0.45	0.54	0.25	0.26	0.52	0.33	0.41	0.34	0.36	0.52	0.43
12	1.11	1.11	0.60	0.72	0.33	0.35	0.70	0.44	0.55	0.45	0.49	0.69	0.57
13	1.05	1.05	0.57	0.68	0.31	0.33	0.66	0.42	0.52	0.42	0.46	0.65	0.54
14	0.78	0.78	0.43	0.51	0.23	0.25	0.49	0.31	0.39	0.32	0.34	0.49	0.40
15	0.86	0.86	0.47	0.56	0.26	0.27	0.54	0.34	0.43	0.35	0.38	0.54	0.44
Total	0.49	0.49	0.27	0.32	0.15	0.16	0.31	0.19	0.24	0.20	0.21	0.31	0.25

Step 3b—Calculate the 13 After Adjustment Factors for each of the 15 comparison sites.
Using Equation 9A.2-2, Columns 14 and 16, the number of after years for the treatment sites (2 years), and the number of after years for the comparison sites (3 years), calculate the 13 after adjustment factors for each of the 15 comparison sites. The results appear in Columns 30 through 42.

(7)	(30)	(31)	(32)	(33)	(34)	(35)	(36)	(37)	(38)	(39)	(40)	(41)	(42)
	Comparison Group—After Adjustment Factors (Equation 9A.2-2)												
Site No.	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.43	0.43	0.23	0.28	0.13	0.14	0.27	0.16	0.19	0.16	0.17	0.24	0.20
2	0.39	0.38	0.21	0.25	0.11	0.12	0.24	0.14	0.17	0.14	0.15	0.22	0.18
3	0.78	0.78	0.42	0.51	0.23	0.25	0.49	0.29	0.35	0.29	0.31	0.44	0.37
4	0.57	0.57	0.31	0.37	0.17	0.18	0.36	0.21	0.26	0.21	0.23	0.32	0.27
5	1.29	1.29	0.70	0.84	0.38	0.41	0.81	0.47	0.58	0.47	0.51	0.73	0.61
6	1.12	1.12	0.61	0.73	0.33	0.36	0.70	0.41	0.50	0.41	0.45	0.63	0.53
7	0.61	0.61	0.33	0.39	0.18	0.19	0.38	0.22	0.27	0.22	0.24	0.34	0.29
8	0.94	0.94	0.51	0.61	0.28	0.30	0.59	0.35	0.42	0.35	0.38	0.54	0.44
9	0.71	0.71	0.39	0.46	0.21	0.23	0.45	0.26	0.32	0.26	0.28	0.40	0.33
10	0.94	0.93	0.51	0.61	0.28	0.30	0.59	0.34	0.42	0.34	0.37	0.53	0.44
11	0.79	0.79	0.43	0.52	0.24	0.25	0.50	0.29	0.36	0.29	0.32	0.45	0.37
12	1.07	1.07	0.58	0.70	0.32	0.34	0.67	0.39	0.48	0.39	0.43	0.61	0.50
13	1.01	1.01	0.55	0.66	0.30	0.32	0.64	0.37	0.46	0.37	0.40	0.57	0.48
14	0.75	0.75	0.41	0.49	0.22	0.24	0.47	0.27	0.34	0.27	0.30	0.42	0.35
15	0.83	0.83	0.45	0.54	0.25	0.26	0.52	0.30	0.37	0.31	0.33	0.47	0.39
Total	0.43	0.43	0.23	0.28	0.13	0.14	0.27	0.16	0.19	0.16	0.17	0.24	0.20

Step 4a—Calculate the Expected Average Crash Frequencies in the before period for an individual comparison site.

(7)	(43)	(44)	(45)	(46)	(47)	(48)	(49)	(50)	(51)	(52)	(53)	(54)	(55)
	Comparison Group—Before Adjusted Crash Frequencies (Equation 9A.2-3)												
Site No.	1	2	3	4	5	6	7	8	9	10	11	12	13
1	13.29	13.26	7.22	8.63	3.96	4.22	8.33	5.26	6.55	5.36	5.80	8.26	6.84
2	2.20	2.20	1.19	1.43	0.66	0.70	1.38	0.87	1.08	0.89	0.96	1.37	1.13
3	6.24	6.23	3.39	4.05	1.86	1.98	3.91	2.47	3.08	2.52	2.73	3.88	3.21
4	13.63	13.60	7.40	8.85	4.06	4.33	8.54	5.40	6.71	5.49	5.95	8.47	7.02
5	4.44	4.43	2.41	2.88	1.32	1.41	2.78	1.76	2.19	1.79	1.94	2.76	2.28
6	7.69	7.68	4.18	5.00	2.29	2.44	4.82	3.05	3.79	3.10	3.36	4.78	3.96
7	18.18	18.14	9.88	11.81	5.41	5.77	11.40	7.20	8.96	7.33	7.94	11.30	9.36
8	12.86	12.83	6.98	8.35	3.83	4.08	8.06	5.09	6.33	5.18	5.61	7.99	6.62
9	16.21	16.18	8.81	10.53	4.83	5.15	10.16	6.42	7.99	6.53	7.08	10.07	8.35
10	32.78	32.71	17.81	21.29	9.76	10.41	20.55	12.98	16.15	13.21	14.31	20.37	16.88
11	4.16	4.16	2.26	2.70	1.24	1.32	2.61	1.65	2.05	1.68	1.82	2.59	2.14
12	3.34	3.33	1.81	2.17	0.99	1.06	2.09	1.32	1.64	1.35	1.46	2.07	1.72
13	4.20	4.19	2.28	2.73	1.25	1.33	2.63	1.66	2.07	1.69	1.83	2.61	2.16
14	9.41	9.39	5.11	6.11	2.80	2.99	5.90	3.73	4.64	3.79	4.11	5.85	4.85
15	18.13	18.09	9.85	11.77	5.40	5.76	11.37	7.18	8.93	7.31	7.91	11.26	9.34
Total	166.77	166.42	90.59	108.30	49.66	52.97	104.55	66.03	82.14	67.21	72.81	103.61	85.87

Using Equation 9A.2-3, Columns 17 through 29, and Column 11, calculate the adjusted crash frequencies in the before period for an individual comparison site. The results appear in Columns 43 through 55.

Step 4b—Calculate the Expected Average Crash Frequencies in the after period for an individual comparison site.
Similarly, using Equation 9A.2-4, Columns 30 through 42, and Column 12, calculate the adjusted crash frequencies in the after period for an individual comparison site. The results appear in Columns 56 through 68.

(7)	(56)	(57)	(58)	(58)	(60)	(61)	(62)	(63)	(64)	(65)	(66)	(67)	(68)
	Comparison Group—After Adjusted Crash Frequencies (Equation 9A.2-4)												
Site No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1	1.72	1.72	0.94	1.12	0.51	0.55	1.08	0.63	0.77	0.63	0.69	0.98	0.81
2	1.93	1.92	1.05	1.25	0.57	0.61	1.21	0.70	0.87	0.71	0.77	1.09	0.90
3	2.34	2.34	1.27	1.52	0.70	0.74	1.47	0.86	1.05	0.86	0.93	1.33	1.10
4	1.14	1.14	0.62	0.74	0.34	0.36	0.72	0.42	0.51	0.42	0.46	0.65	0.54
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	1.12	1.12	0.61	0.73	0.33	0.36	0.70	0.41	0.50	0.41	0.45	0.63	0.53
7	6.69	6.67	3.63	4.34	1.99	2.12	4.19	2.44	3.01	2.46	2.66	3.79	3.14
8	3.78	3.77	2.05	2.45	1.13	1.20	2.37	1.38	1.70	1.39	1.51	2.14	1.78
9	8.53	8.51	4.63	5.54	2.54	2.71	5.35	3.12	3.83	3.14	3.40	4.83	4.01
10	4.68	4.67	2.54	3.04	1.39	1.49	2.93	1.71	2.10	1.72	1.86	2.65	2.20
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	10.13	10.11	5.50	6.58	3.02	3.22	6.35	3.70	4.55	3.72	4.03	5.74	4.76
14	1.49	1.49	0.81	0.97	0.44	0.47	0.94	0.55	0.67	0.55	0.60	0.85	0.70
15	1.66	1.66	0.90	1.08	0.49	0.53	1.04	0.61	0.75	0.61	0.66	0.94	0.78
Total	45.21	45.11	24.56	29.35	13.46	14.36	28.35	16.51	20.32	16.62	18.01	25.63	21.24

Step 5—Calculate the Total Expected Comparison-Group Crash Frequencies in the before period for each treatment site.
Applying Equation 9A.2-5, sum the crash frequencies in each of the Columns 43 through 55 obtained in Step 4a. These are the 13 total comparison-group adjusted crash frequencies in the before period for each treatment site. The results appear in the final row of the table presented with Step 4a.

Step 6—Calculate the Total Expected Comparison-group Crash Frequencies in the after period for each treatment site.
Similarly, applying Equation 9A.2-6, sum the crash frequencies in each of the Columns 56 through 68 obtained in Step 4b. These are the 13 total comparison-group adjusted crash frequencies in the after period for each treatment site. The results appear in the final row of the table presented with Step 4b.

Step 7—Reorganize the Treatment Site Data by transposing the column totals (last row) of the tables shown in Steps 4a and 4b.
For ease of computation, reorganize the treatment site data (*M* and *N*) as shown below by transposing the column totals (last row) of the tables shown in Steps 4a and 4b.

Using Equation 9A.2-7, Columns 69 and 70, calculate the comparison ratios. The results appear in Column 71.

(1)	(69)	(70)	(71)	(72)	(6)	(73)
Treatment Sites						
Site No.	**Comparison-group adjusted crash frequency in before period**	**Comparison-group adjusted crash frequency in after period**	**Comparison ratio**	**Expected average crash frequency in after period without treatment**	**Observed crash frequency in after period**	**Odds ratio**
1	166.77	45.21	0.271	4.34	2	0.461
2	166.42	45.11	0.271	1.63	2	1.230
3	90.59	24.56	0.271	1.08	2	1.845
4	108.30	29.35	0.271	4.34	1	0.231
5	49.66	13.46	0.271	0.27	1	3.689
6	52.97	14.36	0.271	1.36	1	0.738
7	104.55	28.35	0.271	4.61	9	1.953
8	66.03	16.51	0.250	3.00	0	0.000
9	82.14	20.32	0.247	1.98	0	0.000
10	67.21	16.62	0.247	0.74	0	0.000
11	72.81	18.01	0.247	2.23	5	2.246
12	103.61	25.63	0.247	2.23	6	2.695
13	85.87	21.24	0.247	3.96	1	0.253
Total	1,216.93	318.72		31.75	30	

Step 8—Calculate the Expected Average Crash Frequency for each treatment site in the after period had no treatment been implemented.

Using Equation 9A.2-8, Columns 5 and 71, calculate the expected average crash frequency for each treatment site in the after period had no treatment been implemented. The results appear in Column 72 in the table presented in Step 7. Sum the frequencies in Column 72.

Step 9—Calculate the Safety Effectiveness, Expressed as an odds ratio, *OR*, at an individual treatment site.

Using Equation 9A.2-9, Columns 6 and 72, calculate the safety effectiveness, expressed as an odds ratio, *OR*, at an individual treatment site. The results appear in Column 73 in the table presented in Step 7.

9.11.4. Estimation of the Overall Treatment Effectiveness and its Precision

Step 10—Calculate the Log Odds Ratio (*R*) for each treatment site.
Using Equation 9A.2-11 and Column 73, calculate the log odds ratio (*R*) for each treatment site. The results appear in Column 74.

(1)	(74)	(75)	(76)	(77)
		Treatment Sites		
Site No.	**Log odds ratio, *R***	**Squared standard error of log odds ratio**	**Weighted Adjustment, *w***	**Weighted product**
1	–0.774	0.591	1.69	–1.31
2	0.207	0.695	1.44	0.30
3	0.612	0.802	1.25	0.76
4	–1.467	1.106	0.90	–1.33
5	1.305	2.094	0.48	0.62
6	–0.304	1.289	0.78	–0.24
7	0.669	0.215	4.66	3.12
8	[a]	[a]	[a]	[a]
9	[a]	[a]	[a]	[a]
10	[a]	[a]	[a]	[a]
11	0.809	0.380	2.63	2.13
12	0.992	0.326	3.06	3.04
13	–1.376	1.121	0.89	–1.23
Total			17.78	5.86

[a] Quantities cannot be calculated because zero crashes were observed in after period at these treatment sites.

Step 11—Calculate the Squared Standard Error of the log odds ratio at each treatment site.
Using Equation 9A.2-13, Columns 5, 6, 69, and 70, calculate the squared standard error of the log odds ratio at each treatment site. The results appear in Column 75 of the table presented with Step 10.

Using Equation 9A.2-12 and Column 75, calculate the weight *w* for each treatment site. The results appear in Column 76 of the table presented with Step 10. Calculate the product of Columns 75 and 76. The results appear in Column 77 of the table presented with Step 10. Sum each of Columns 76 and 77.

Step 12—Calculate the Weighted Average Log Odds ratio, *R*, across all treatment sites.
Using Equation 9A.2-14 and the sums from Columns 76 and 77, calculate the weighted average log odds ratio (*R*) across all treatment sites:

$$R = \frac{5.86}{17.78} = 0.33$$

Step 13—Calculate the Overall Effectiveness of the Treatment expressed as an odds ratio.
Using Equation 9A.2-15 and the result from Step 12, calculate the overall effectiveness of the treatment, expressed as an odds ratio, *OR*, averaged across all sites:

$$OR = e^{(0.33)} = 1.391$$

Step 14—Calculate the Overall Safety Effectiveness, expressed as a percentage change in crash frequency, *CMF*, averaged across all sites.
Using Equation 9A.2-16 and the results from Step 13, calculate the overall safety effectiveness, expressed as a percentage change in crash frequency, Safety Effectiveness, averaged across all sites:

Safety Effectiveness = 100 × (1 – 1.391) = –39.1%

Note—The negative estimate of the Safety Effectiveness indicates a negative effectiveness, i.e., an increase in total crashes.

Step 15—Calculate the Precision of the Treatment Effectiveness.
Using Equation 9A.2-17 and the results from Step 13 and the sum from Column 76, calculate the precision of the treatment effectiveness:

$$SE(\text{Safety Effectiveness}) = 100\left(\frac{1.391}{\sqrt{17.78}}\right) = 33.0\%$$

Step 16—Assess the Statistical Significance of the Estimated Safety Effectiveness.
Assess the statistical significance of the estimated safety effectiveness by calculating the quantity:

$$Abs\left(\frac{\text{Safety Effectivness}}{SE\left(\text{Safety Effectiveness}\right)}\right) = \frac{39.1}{33.0} = 1.18$$

Since *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] < 1.7, conclude that the treatment effect is not significant at the (approximate) 90 percent confidence level.

In summary, the evaluation results indicate that an average increase in total crash frequency of 39.1 percent was observed after the installation of passing lanes at the rural two-lane highway sites, but this increase was not statistically significant at the 90 percent confidence level. This sample problem provided different results than the EB evaluation in Section B.1 for two primary reasons. First, a comparison group rather than an SPF was used to estimate future changes in crash frequency at the treatment sites. Second, the three treatment sites at which zero crashes were observed in the period after installation of the passing lanes could not be considered in the comparison-group method because of division by zero. These three sites were considered in the EB method. This illustrates a weakness of the comparison-group method which has no mechanism for considering these three sites where the treatment appears to have been most effective.

9.12. SAMPLE PROBLEM TO ILLUSTRATE THE SHIFT OF PROPORTIONS SAFETY EFFECTIVENESS EVALUATION METHOD

Passing lanes have been installed to increase passing opportunities at 13 rural two-lane highway sites. An evaluation is to be conducted to determine the overall effect of the installation of these passing lanes on the proportion of fatal-and-injury crashes at the 13 treatment sites.

Data are available for both fatal-and-injury and total crash frequencies for each of the 13 rural two-lane highway sites for five years before and two years after installation of passing lanes. These data can be used to estimate fatal-and-injury crash frequency as a proportion of total crash frequency for the periods before and after implementation of the treatment.

As before, column numbers are shown in the first row of all the tables in this sample problem; the description of the calculations refers to these column numbers for clarity of explanation. When columns are repeated from table to table, the original column number is kept. Where appropriate, column totals are indicated in the last row of each table.

9.12.1. Basic Input Data

Organize the observed before- and after-period total and fatal-and-injury (FI) crash frequencies for the 13 rural two-lane road segments as follows in Columns 1 through 5:

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Site No.	Crash frequency in before period (5 years)		Crash frequency in after period (2 years)		Proportion of FI/total crashes		Difference in proportions
	Total	FI	Total	FI	Before	After	
1	17	9	3	3	0.53	1.000	0.471
2	6	3	3	2	0.50	0.667	0.167
3	6	2	3	2	0.33	0.667	0.333
4	17	6	3	2	0.35	0.667	0.314
5	1	1	2	1	1.00	0.500	–0.500
6	5	2	3	0	0.40	0.000	–0.400
7	18	12	10	3	0.67	0.300	–0.367
8	12	3	2	1	0.25	0.500	0.250
9	8	1	1	1	0.13	1.000	0.875
10	4	3	1	0	0.75	0.000	–0.750
11	10	1	6	2	0.10	0.333	0.233
12	10	3	7	1	0.30	0.143	–0.157
13	18	4	1	1	0.22	1.000	0.778
Total	132	50	45	19			1.247

9.12.2. Estimate the Average Shift in Proportion of the Target Collision Type

Step 1—Calculate the Before Treatment Proportion.
Using Equation 9A.3-1 and Columns 2 and 3, calculate the before treatment proportion. The results appear in Column 6 above.

Step 2—Calculate the After Treatment Proportion.
Similarly, using Equation 9A.3-2 and Columns 4 and 5, calculate the after treatment proportion. The results appear in Column 7 above.

Step 3—Calculate the Difference between the After and Before Proportions at each treatment site.
Using Equation 9A.3-3 and Columns 6 and 7, calculate the difference between the after and before proportions at each treatment site. The results appear in Column 8 above. Sum the entries in Column 8.

Step 4—Calculate the Average Difference between After and Before Proportions over all n treatment sites.
Using Equation 9A.3-4, the total from Column 8, and the number of sites (13), calculate the average difference between after and before proportions over all *n* treatment sites:

$$AvgP_{(FI)diff} = \frac{1.247}{13} = 0.10$$

This result indicates that the treatment resulted in an observed change in the proportion of fatal-and-injury crashes of 0.10, i.e., a 10 percent increase in proportion.

9.12.3. Assess the Statistical Significance of the Average Shift in Proportion of the Target Collision Type

Step 5—Obtain the Absolute Value of the Differences in Proportion in Column 8.
Using Equation 9A.3-5, obtain the absolute value of the differences in proportion in Column 8. The results appear in Column 9 in the table presented in Step 6.

Step 6—Sort the Data in ascending order of the absolute values in Column 9.
Sort the data in ascending order of the absolute values in Column 9. Assign the corresponding rank to each site. The results appear in Column 10. [**Note**—sum the numbers in Column 10; this is the maximum total rank possible based on 13 sites.] Organize the data as shown below:

(1)	(8)	(9)	(10)	(11)
Site No.	**Difference in proportions**	**Absolute difference in proportions**	**Rank**	**Rank corresponding to positive difference**
12	–0.157	0.157	1	0
2	0.167	0.167	2	2
11	0.233	0.233	3	3
8	0.250	0.250	4	4
4	0.314	0.314	5	5
3	0.333	0.333	6	6
7	–0.367	0.367	7	0
6	–0.400	0.400	8	0
1	0.471	0.471	9	9
5	–0.500	0.500	10	0
10	–0.750	0.750	11	0
13	0.778	0.778	12	12
9	0.875	0.875	13	13
Total			91	54

Step 7—Calculate the Value of the T^+ Statistic.
Replace all ranks (shown in Column 10) associated with negative difference (shown in Column 8) with zero. The results appear in Column 11 in the table presented in Step 6. Sum the ranks in Column 11. This is the value of the T^+ statistic in Equation 9A.3-6:

$T^+ = 54$

Step 8—Assess the Statistical Significance of T^+ Using a two-sided significance test at the 0.10 level (90 percent confidence level).
Assess the statistical significance of T^+ using a two-sided significance test at the 0.10 level (90 percent confidence level). Using Equation 9A.3-7 and Table 9A.3-1, obtain the upper and lower critical limits as:

- Upper limit—$t(\alpha_2,13) = 70$; this corresponds to an α_2 of 0.047, the closest value to 0.10/2
- Lower limit—$91 - t(\alpha_1,13) = 91 - 69 = 22$; here 69 corresponds to an α_1 of 0.055, for a total α of 0.047 + 0.055 = 0.102, the closest value to the significance level of 0.10

Since the calculated T^+ of 54 is between 22 and 70, conclude that the treatment has not significantly affected the proportion of fatal-and-injury crashes relative to total crashes.

In summary, the evaluation results indicate that an increase in proportion of fatal-and-injury crashes of 0.10 (i.e., 10 percent) was observed after the installation of passing lanes at the 13 rural two-lane highway sites, but this increase was not statistically significant at the 90 percent confidence level.

9.13 REFERENCES

(1) Griffin, L. I., and R. J. Flowers. *A Discussion of Six Procedures for Evaluating Highway Safety Projects*. Federal Highway Administration, U.S. Department of Transportation, Washington, DC, December 1997.

(2) Harwood, D. W., K. M. Bauer. I. B. Potts., D. J. Torbic. K. R. Richard, E. R. Kohlman Rabbani, E. Hauer, and L. Elefteriadou. *Safety Effectiveness of Intersection Left- and Right-Turn Lanes*. Report No. FHWA-RD-02-089. Federal Highway Administration, U.S. Department of Transportation, Washington, DC, April 2002.

(3) Harwood, D. W., et al. *SafetyAnalyst: Software Tools for Safety Management of Specific Highway Sites*. Federal Highway Administration, U.S. Department of Transportation, Washington, DC. More information available from http://www.safetyanalyst.org.

(4) Hauer, E. Cause and Effect in Observational Cross-Section Studies on Road Safety. *Transportation Research Board Annual Meeting CD-ROM*. TRB, National Research Council, Washington, DC, 2005.

(5) Hauer, E. *Observational Before-after Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*. Pergamon Press, Elsevier Science Ltd, Oxford, UK, 1997.

(6) Hauer, E., D. W. Harwood, F. M. Council., and M. S. Griffith. Estimating Safety by the Empirical Bayes Method: A Tutorial. In *Transportation Research Record 1784*. TRB, National Research Council, Washington, DC, 2002.

(7) Hollander, M., and D. A. Wolfe. *Nonparametric Statistical Methods*. John Wiley & Sons, Inc., Hoboken, NJ, 1973.

(8) Lord, D. and B. N. Persaud, 2000. Accident Prediction Models with and without Trend: Application of the Generalized Estimating Equation Procedure. In *Transportation Research Record 1717*. TRB, National Research Council, Washington, DC, pp. 102–108.

(9) Lyon, C., B. N. Persaud, N. X. Lefler, D. L. Carter, and K. A. Eccles. Safety Evaluation of Installing Center Two-Way Left-Turn Lanes on Two-Lane Roads. In *Transportation Research Record 2075*, TRB, National Research Council, Washington, DC, 2008, pp. 34–41.

(10) Persaud, B. N., R. A. Retting, P. E. Garder, and D. Lord. Safety Effect of Roundabout Conversions in the United States: Empirical Bayes Observational Before-After Studies. In *Transportation Research Record 1751*. TRB, National Research Council, Washington, DC, 2001.

APPENDIX 9A—COMPUTATIONAL PROCEDURES FOR SAFETY EFFECTIVENESS EVALUATION METHODS

This appendix presents computational procedures for three observational before/after safety evaluation methods presented in this chapter, including the EB method, the comparison-group method, and the shift in proportions method.

9A.1. COMPUTATIONAL PROCEDURE FOR IMPLEMENTING THE EB BEFORE/AFTER SAFETY EFFECTIVENESS EVALUATION METHOD

A computational procedure using the EB method to determine the safety effectiveness of the treatment being evaluated, expressed as a percentage change in crashes, θ, and to assess its precision and statistical significance, is presented as follows.

All calculations are shown in Steps 1 through 13 in this section for the total crash frequencies for the before period and after periods, respectively, at a given site. The computational procedure can also be adapted to consider crash frequencies on a year-by-year basis for each site [e.g., see the computational procedure used in the FHWA SafetyAnalyst software (3)].

EB Estimation of the Expected Average Crash Frequency in the Before Period

Step 1—Using the applicable SPF, calculate the predicted average crash frequency, $N_{predicted}$, for site type x during each year of the before period. For roadway segments, the predicted average crash frequency will be expressed as crashes per site per year; for intersections, the predicted average crash frequency is expressed as crashes per intersection per year. Note that:

$$N_{\text{predicted}} = N_{spf} \times (CMF_{1x} \times CMF_{2x} \times \ldots \times CMF_{yx}) \times C_x$$

However, for this level of evaluation, it may be assumed that all* CMF*s and* C_x *are equal to 1.0.

Step 2—Calculate the expected average crash frequency, $N_{expected}$, for each site i, summed over the entire before period. For roadway segments, the expected average crash frequency will be expressed as crashes per site; for intersections, the expected average crash frequency is expressed as crashes per intersection.

$$N_{\text{expected},B} = w_{i,B}\, N_{\text{predicted}} + (1 - w_{i,B}) N_{\text{observed},B} \tag{9A.1-1}$$

Where the weight, $w_{i,B}$, for each site i, is determined as:

$$w_{i,B} = \frac{1}{1 + k \sum\limits_{\substack{\text{Before}\\\text{years}}} N_{\text{predicted}}} \tag{9A.1-2}$$

and:

$N_{expected}$ = Expected average crash frequency at site i for the entire before period

$N_{spf\,x}$ = Predicted average crash frequency determined with the applicable SPF (from Step 1)

$N_{observed,B}$ = Observed crash frequency at site i for the entire before period

k = Overdispersion parameter for the applicable SPF

Note—If no SPF is available for a particular crash severity level or crash type being evaluated, but that crash type is a subset of another crash severity level or crash type for which an SPF is available, the value of $PR_{i,y,B}$ can be determined by multiplying the SPF-predicted average crash frequency by the average proportion represented by the crash severity level or crash type of interest. This approach is an approximation that is used when an SPF for the crash severity level or crash type of interest cannot be readily developed. If an SPF from another jurisdiction is available, consider calibrating that SPF to local conditions using the calibration procedure presented in the Appendix to Part C.

EB Estimation of the Expected Average Crash Frequency in the After Period in the Absence of the Treatment

Step 3—Using the applicable SPF, calculate the predicted average crash frequency, $PR_{i,y,A}$, for each site i during each year y of the after period.

Step 4—Calculate an adjustment factor, r_i, to account for the differences between the before and after periods in duration and traffic volume at each site i as:

$$r_i = \frac{\sum\limits_{\substack{\text{After}\\\text{years}}} N_{\text{predicted},A}}{\sum\limits_{\substack{\text{Before}\\\text{years}}} N_{\text{predicted},B}} \tag{9A.1-3}$$

Step 5—Calculate the expected average crash frequency, N_{expected}, for each site *i*, over the entire after period in the absence of the treatment as:

$$N_{\text{expected},A} = N_{\text{expected},B} \times r_i \tag{9A.1-4}$$

Estimation of Treatment Effectiveness

Step 6—Calculate an estimate of the safety effectiveness of the treatment at each site *i* in the form of an odds ratio, OR_i, as:

$$OR_i = \frac{N_{\text{observed},A}}{N_{\text{expected},A}} \tag{9A.1-5}$$

Where:

OR_i = Odd ration at site *i*

$N_{\text{observed},A}$ = Observed crash frequency at site *i* for the entire after period

Step 7—Calculate the safety effectiveness as a percentage crash change at site *i* as:

$$\text{Safety Effectiveness}_i = 100 \times (1 - OR_i) \tag{9A.1-6}$$

Step 8—Calculate the overall effectiveness of the treatment for all sites combined, in the form of an odds ratio, *OR′*, as follows:

$$OR' = \frac{\sum\limits_{\text{All sites}} N_{\text{observed},A}}{\sum\limits_{\text{All sites}} N_{\text{expected},A}} \tag{9A.1-7}$$

Step 9—The odds ratio, *OR′*, calculated in Equation 9A.1-7 is potentially biased; therefore, an adjustment is needed to obtain an unbiased estimate of the treatment effectiveness in terms of an adjusted odds ratio, *OR*. This is calculated as follows:

$$OR = \frac{OR'}{1 + \dfrac{Var\left(\sum\limits_{\text{All sites}} N_{\text{expected},A}\right)}{\left(\sum\limits_{\text{All sites}} N_{\text{expected},A}\right)^2}} \tag{9A.1-8}$$

Where:

$$Var\left(\sum_{\text{All sites}} N_{\text{expected},A}\right) = \sum_{\text{All sites}} \left[(r_i)^2 \times N_{\text{expected},B} \times \left(1 - w_{i,B}\right)\right] \tag{9A.1-9}$$

and $w_{i,B}$ is defined in Equation 9A.1-2 and r_i is defined in Equation 9A.1-3.

Step 10—Calculate the overall unbiased safety effectiveness as a percentage change in crash frequency across all sites as:

$$\text{Safety Effectiveness} = 100 \times (1 - OR) \qquad (9A.1\text{-}10)$$

Estimation of the Precision of the Treatment Effectiveness

To assess whether the estimated safety effectiveness of the treatment is statistically significant, one needs to determine its precision. This is done by first calculating the precision of the odds ratio, *OR*, in Equation 9A.1-8. The following steps show how to calculate the variance of this ratio to derive a precision estimate and present criteria assessing the statistical significance of the treatment effectiveness estimate.

Step 11—Calculate the variance of the unbiased estimated safety effectiveness, expressed as an odds ratio, *OR*, as follows:

$$Var(OR) = \frac{(OR')^2 \left[\dfrac{1}{N_{\text{observed},A}} + \dfrac{Var\left(\sum\limits_{\text{All sites}} N_{\text{expected},A} \right)}{\left(\sum\limits_{\text{All sites}} N_{\text{expected},A} \right)^2} \right]}{\left[1 + \dfrac{Var\left(\sum\limits_{\text{All sites}} N_{\text{expected},A} \right)}{\left(\sum\limits_{\text{All sites}} N_{\text{expected},A} \right)^2} \right]} \qquad (9A.1\text{-}11)$$

Step 12—To obtain a measure of the precision of the odds ratio, *OR*, calculate its standard error as the square root of its variance:

$$SE(OR) = \sqrt{Var(OR)} \qquad (9A.1\text{-}12)$$

Step 13—Using the relationship between *OR* and Safety Effectiveness shown in Equation 9A.1-10, the standard error of Safety Effectiveness, *SE*(Safety Effectiveness), is calculated as:

$$SE(\text{Safety Effectiveness}) = 100 \times SE(OR) \qquad (9A.1\text{-}13)$$

Step 14—Assess the statistical significance of the estimated safety effectiveness by making comparisons with the measure *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] and drawing conclusions based on the following criteria:

- If *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] < 1.7, conclude that the treatment effect is not significant at the (approximate) 90 percent confidence level.
- If *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] ≥ 1.7, conclude that the treatment effect is significant at the (approximate) 90 percent confidence level.
- If *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] ≥ 2.0, conclude that the treatment effect is significant at the (approximate) 95 percent confidence level.

9A.2 COMPUTATIONAL PROCEDURE FOR IMPLEMENTING THE COMPARISON-GROUP SAFETY EFFECTIVENESS EVALUATION METHOD

A computational procedure using the comparison-group evaluation study method to determine the safety effectiveness of the treatment being evaluated, expressed as a percentage change in crashes, θ, and to assess its precision and statistical significance, is presented below.

Note—The following notation will be used in presenting the computational procedure for the comparison-group method. Each individual treatment site has a corresponding comparison group of sites, each with their own AADT and number of before and after years. The notation is as follows:

- Subscript i denotes a treatment site, i=1,…,n, where n denotes the total number of treatment sites
- Subscript j denotes a comparison site, j=1,…,m, where m denotes the total number of comparison sites
- Each treatment site i has a number of before years, Y_{BT}, and a number of after years, Y_{AT}
- Each comparison site j has a number of before years, Y_{BC}, and a number of after years, Y_{AC}
- It is assumed for this section that Y_{BT} is the same across all treatment sites; that Y_{AT} is the same across all treatment sites; that Y_{BC} is the same across all comparison sites; and that Y_{AC} is the same across all comparison sites. Where this is not the case, computations involving the durations of the before and after periods may need to vary on a site-by-site basis.

The following symbols are used for observed crash frequencies, in accordance with Hauer's notation (5):

	Before Treatment	After Treatment
Treatment Site	$N_{\text{observed},T,B}$	$N_{\text{observed},T,A}$
Comparison Group	$N_{\text{observed},C,B}$	$N_{\text{observed},C,A}$

Estimation of Mean Treatment Effectiveness

Step 1a—Using the applicable SPF and site-specific AADT, calculate $\sum N_{\text{predicted},T,B}$, the sum of the predicted average crash frequencies at treatment site *i* in before period.

Step 1b—Using the applicable SPF and site-specific AADT, calculate $\sum N_{\text{predicted},T,A}$, the sum of the predicted average crash frequencies at treatment site *i* in after period.

Step 2a—Using the applicable SPF and site-specific AADT, calculate $\sum N_{\text{predicted},C,B}$, the sum of the predicted average crash frequencies at comparison site *j* in before period.

Step 2b—Using the applicable SPF and site-specific AADT, calculate $\sum N_{\text{predicted},C,A}$, the sum of the predicted average crash frequencies at comparison site *j* in after period.

Step 3a—For each treatment site *i* and comparison site *j* combination, calculate an adjustment factor to account for differences in traffic volumes and number of years between the treatment and comparison sites during the before period as follows:

$$Adj_{i,j,B} = \frac{N_{\text{predicted},T,B}}{N_{\text{predicted},C,B}} \times \frac{Y_{BT}}{Y_{BC}} \tag{9A.2-1}$$

Where:

$N_{\text{predicted},T,B}$ = Sum of predicted average crash frequencies at treatment site i in before period using the appropriate SPF and site-specific AADT;

$N_{\text{predicted},C,B}$ = Sum of predicted average crash frequencies at comparison site j in before period using the same SPF and site-specific AADT;

Y_{BT} = Duration (years) of before period for treatment site *i*; and

Y_{BC} = Duration (years) of before period for comparison site *j*.

Step 3b—For each treatment site *i* and comparison site *j* combination, calculate an adjustment factor to account for differences in AADTs and number of years between the treatment and comparison sites during the after period as follows:

$$Adj_{i,j,A} = \frac{N_{\text{predicted},T,A}}{N_{\text{predicted},C,A}} \times \frac{Y_{AT}}{Y_{AC}} \qquad (9A.2\text{-}2)$$

Where:

$N_{\text{predicted},T,A}$ = Sum of predicted average crash frequencies at treatment site *i* in after period using the appropriate SPF and site-specific AADT;

$N_{\text{predicted},C,A}$ = Sum of predicted average crash frequencies at comparison site *j* in the after period using the same SPF and site-specific AADT;

Y_{AT} = Duration (years) of after period for treatment site *i*; and

Y_{AC} = Duration (years) of after period for comparison site *j*

Step 4a—Using the adjustment factors calculated in Equation 9A.2-1, calculate the expected average crash frequencies in the before period for each comparison site *j* and treatment site *i* combination, as follows:

$$N_{\text{expected},C,B} = \sum_{\text{All sites}} N_{\text{observed},C,B} \times Adj_{i,j,B} \qquad (9A.2\text{-}3)$$

Where:

$\sum N_{\text{observed},C,B}$ = Sum of observed crash frequencies at comparison site *j* in the before period

Step 4b—Using the adjustment factor calculated in Equation 9A.2-2, calculate the expected average crash frequencies in the after period for each comparison site *j* and treatment site *i* combination, as follows:

$$N_{\text{expected},C,A} = \sum_{\text{All sites}} N_{\text{observed},C,A} \times Adj_{i,j,A} \qquad (9A.2\text{-}4)$$

Where:

N_j = Sum of observed crash frequencies at comparison site *j* in the after period

Step 5—For each treatment site *i*, calculate the total comparison-group expected average crash frequency in the before period as follows:

$$N_{\text{expected},C,B,\text{total}} = \sum_{\text{All comparison sites}} N_{\text{expected},C,B} \qquad (9A.2\text{-}5)$$

Step 6—For each treatment site *i*, calculate the total comparison-group expected average crash frequency in the after period as follows:

$$N_{\text{expected},C,A,\text{total}} = \sum_{\text{All comparison sites}} N_{\text{expected},C,A} \tag{9A.2-6}$$

Step 7—For each treatment site *i*, calculate the comparison ratio, r_{iC}, as the ratio of the comparison-group expected average crash frequency after period to the comparison-group expected average crash frequency in the before period at the comparison sites as follows:

$$r_{i,C} = \frac{N_{\text{expected},C,A,\text{total}}}{N_{\text{expected},C,B,\text{total}}} \tag{9A.2-7}$$

Step 8—Using the comparison ratio calculated in Equation 9A.2-7, calculate the expected average crash frequency for a treatment site *i* in the after period, had no treatment been implement as follows:

$$N_{\text{expected},T,A} = \sum_{\text{All sites}} N_{\text{observed},T,B} \times r_{iC} \tag{9A.2-8}$$

Step 9—Using Equation 9A.2-9, calculate the safety effectiveness, expressed as an odds ratio, OR_i, at an individual treatment site *i* as the ratio of the expected average crash frequency with the treatment over the expected average crash frequency had the treatment not been implemented, as follows:

$$OR_i = \sum_{\text{All sites}} \frac{N_{\text{observed},T,A}}{N_{\text{expected},T,A}} \tag{9A.2-9}$$

or alternatively,

$$OR_i = \frac{N_{\text{observed},T,A,\text{total}}}{N_{\text{observed},T,B,\text{total}}} \times \frac{N_{\text{expected},C,B,\text{total}}}{N_{\text{expected},C,A,\text{total}}} \tag{9A.2-10}$$

Where:

$N_{\text{observed},T,A}$,total and $N_{\text{observed},T,B,\text{total}}$ represent the total treatment group observed crash frequencies at treatment site *i* calculated as the sum of $N_{\text{observed},T,A}$ and $N_{\text{observed},T,B}$ for all sites;

The next steps show how to estimate weighted average safety effectiveness and its precision based on individual site data.

Step 10—For each treatment site *i*, calculate the log odds ratio, R_i, as follows:

$$R_i = \ln(OR_i) \tag{9A.2-11}$$

Where the ln function represents the natural logarithm.

Step 11—For each treatment site *i*, calculate the weight w_i as follows:

$$w_i = \frac{1}{R_{i(se)}^2} \tag{9A.2-12}$$

Where:

$$R^2_{i(SE)} = \frac{1}{N_{\text{observed},T,B,\text{total}}} + \frac{1}{N_{\text{observed},T,A,\text{total}}} + \frac{1}{N_{\text{expected},C,B,\text{total}}} + \frac{1}{N_{\text{expected},C,A,\text{total}}} \qquad (9A.2\text{-}13)$$

Step 12—Using Equation 9A.2-14, calculate the weighted average log odds ratio, *R*, across all *n* treatment sites as:

$$R = \frac{\sum_n w_i R_i}{\sum_n w_i} \qquad (9A.2\text{-}14)$$

Step 13—Exponentiating the result from Equation 9A.2-14, calculate the overall effectiveness of the treatment, expressed as an odds ratio, *OR*, averaged across all sites, as follows:

$$OR = e^R \qquad (9A.2\text{-}15)$$

Step 14—Calculate the overall safety effectiveness, expressed as a percentage change in crash frequency averaged across all sites as:

$$\text{Safety Effectiveness} = 100 \times (1 - R) \qquad (9A.2\text{-}16)$$

Step 15—To obtain a measure of the precision of the treatment effectiveness, calculate its standard error, *SE*(Safety Effectiveness), as follows:

$$SE(\text{Safety Effectiveness}) = 100 \frac{OR}{\sqrt{\sum_n w_i}} \qquad (9A.2\text{-}17)$$

Step 16—Assess the statistical significance of the estimated safety effectiveness by making comparisons with the measure *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] and drawing conclusions based on the following criteria:

- If *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] < 1.7, conclude that the treatment effect is not significant at the (approximate) 90 percent confidence level.
- If *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] ≥ 1.7, conclude that the treatment effect is significant at the (approximate) 90 percent confidence level.
- If *Abs*[Safety Effectiveness/*SE*(Safety Effectiveness)] ≥ 2.0, conclude that the treatment effect is significant at the (approximate) 95 percent confidence level.

9A.3 COMPUTATIONAL PROCEDURE FOR IMPLEMENTING THE SHIFT OF PROPORTIONS SAFETY EFFECTIVENESS EVALUATION METHOD

A computational procedure using the evaluation study method for assessing shifts in proportions of target collision types to determine the safety effectiveness of the treatment being evaluated, $AvgP_{(CT)diff}$, and to assess its statistical significance, follows.

This step-by-step procedure uses the same notation as that used in the traditional comparison-group safety evaluation method. All proportions of specific crash types (subscript "*CT*") are relative to total crashes (subscript "total").

- $N_{\text{observed},B,\text{total}}$ denotes the observed number of total crashes at treatment site *i* over the entire before treatment period.
- $N_{\text{observed},B,CT}$ denotes the observed number of *CT* crashes of a specific crash type at treatment site *i* over the entire before treatment period.
- $N_{\text{observed},A,\text{total}}$ denotes the observed number of total crashes at treatment site *i* over the entire after treatment period.
- $N_{\text{observed},A,CT}$ denotes the observed number of *CT* crashes of a specific crash type at treatment site *i* over the entire after treatment period.

Estimate the Average Shift in Proportion of the Target Collision Type

Step 1—Calculate the before treatment proportion of observed crashes of a specific target collision type (*CT*) relative to total crashes (total) at treatment site *i*, $P_{i(CT)B}$, across the entire before period as follows:

$$P_{i(CT)B} = \frac{N_{\text{obseved},B,CT}}{N_{\text{obseved},B,\text{total}}} \tag{9A.3-1}$$

Step 2—Similarly, calculate the after treatment proportion of observed crashes of a specific target collision type of total crashes at treatment site *i*, $P_{i(CT)A}$, across the entire after period as follows:

$$P_{i(CT)A} = \frac{N_{\text{obseved},A,CT}}{N_{\text{obseved},A,\text{total}}} \tag{9A.3-2}$$

Step 3—Determine the difference between the after and before proportions at each treatment site *i* as follows:

$$P_{i(CT)diff} = P_{i(CT)A} - P_{i(CT)B} \tag{9A.3-3}$$

Step 4—Calculate the average difference between after and before proportions over all *n* treatment sites as follows:

$$AvgP_{(CT)diff} = \frac{1}{n} \sum_{\substack{\text{Treat}\\ \text{sites}}} P_{i(CT)diff} \tag{9A.3-4}$$

Assess the Statistical Significance of the Average Shift in Proportion of the Target Collision Type

The following steps demonstrate how to assess whether the treatment significantly affected the proportion of crashes of the collision type under consideration. Because the site-specific differences in Equation 9A.3-4 do not necessarily come from a normal distribution and because some of these differences may be equal to zero, a nonparametric statistical method, the Wilcoxon signed rank test, is used to test whether the average difference in proportions calculated in Equation 9A.3-4 is significantly different from zero at a predefined confidence level.

Step 5—Take the absolute value of the non-zero $P_{i(CT)diff}$ calculated in Equation 9A.3-3. For simplicity of notation, let Z_i denote the absolute value of $P_{i(CT)diff}$, thus:

$$Z_i = abs(P_{i(CT)\,diff}) \tag{9A.3-5}$$

Where:

i = 1,…,n*, with n* representing the (reduced) number of treatment sites with non-zero differences in proportions.

Step 6—Arrange the n^* Z_i values in ascending rank order. When multiple Z_i have the same value (i.e., ties are present), use the average rank as the rank of each tied value of Z_i. For example, if three Z_i values are identical and would rank, say, 12, 13, and 14, use 13 as the rank for each. If the ranks would be, for example, 15 and 16, use 15.5 as the rank for each. Let R_i designate the rank of the Z_i value.

Step 7—Using only the ranks associated with positive differences (i.e., positive values of $P_{i(CT)diff}$), calculate the statistic T^+ as follows:

$$T^+ = \sum_{n^*} R_i^+ \qquad (9A.3\text{-}6)$$

Step 8—Assess the statistical significance of T^+ using a two-sided significance test at the α level of significance (i.e., [1 – α] confidence level) as follows:

- Conclude that the treatment is statistically significant if:

$$T^+ \geq t(a_2, n^*) \quad \text{or} \quad T^+ \leq \frac{n^*(n^*+1)}{2} - t(a_1, n^*) \qquad (9A.3\text{-}7)$$

Where:

$\alpha = \alpha_1 + \alpha_2$

- Otherwise, conclude that the treatment is not statistically significant.

The quantities $t(\alpha_1,n^*)$ and $t(\alpha_2,n^*)$ are obtained from the table of critical values for the Wilcoxon signed rank test, partially reproduced in Table 9A.3-1. Generally, α_1 and α_2 are approximately equal to $\alpha/2$. Choose the values for α_1 and α_2 so that $\alpha_1 + \alpha_2$ is closest to α in Table 9A.3-1 and α_1 and α_2 are each closest to $\alpha/2$. Often, $\alpha_1 = \alpha_2$ are the closest values to $\alpha/2$.

Table 9A.3-1 presents only an excerpt of the full table of critical values shown in Hollander and Wolfe (8). A range of significance levels (α) has been selected to test a change in proportion of a target collision type—approximately 10 to 20 percent. Although 5 to 10 percent are more typical significance levels used in statistical tests, then a 20 percent significance level has been included here because the Wilcoxon signed rank test is a conservative test (i.e., it is difficult to detect a significant effect when it is present). Table 9A.3-1 shows one-sided probability levels; since the test performed here is a two-sided test, the values in Table 9A.3-1 correspond to $\alpha/2$, with values ranging from 0.047 to 0.109 (corresponding to 0.094/2 to 0.218/2).

Example for Using Table 9A.3-1

Assume $T^+ = 4$, $n^* = 9$, and $\alpha = 0.10$ (i.e., 90 percent confidence level). The value of $t(\alpha_2,n^*) = t(0.049,9) = 37$ from Table 9A.3-1, the closest value corresponding to $\alpha = 0.10/2$ in the column for $n^* = 9$. In this case, $t(\alpha_1,n^*) = t(\alpha_2,n^*)$. Thus, the two critical values are 37 and 8 [= 9 × (9 + 1)/2 – 37 = 45 – 37 = 8]. Since $T^+ = 4 < 8$, the conclusion would be that the treatment was statistically significant (i.e., effective) at the 90.2 percent confidence level [where 90.2 = 1 – 2 × 0.049] based on Equation 9A.3-7.

Table 9A.3-1. Upper Tail Probabilities for the Wilcoxon Signed Rank T^+ Statistic ($n^* = 4$ to 10)[a] (8)

	Number of sites (n^*)						
x	**4**	**5**	**6**	**7**	**8**	**9**	**10**
10	0.062						
13		0.094					
14		0.062					
17			0.109				
18			0.078				
19			0.047				
22				0.109			
23				0.078			
24				0.055			
28					0.098		
29					0.074		
30					0.055		
34						0.102	
35						0.082	
36						0.064	
37						0.049	
41							0.097
42							0.080
43							0.065
44							0.053

[a] For a given n^*, the table entry for the point x is $P(T^+ \geq x)$. Thus if x is such that $P(T^+ \geq x) = \alpha$, then $t(\alpha,n^*) = x$.

Table 9A.3-1 (Continued). Upper Tail Probabilities for the Wilcoxon Signed Rank T^+ Statistic (n^* = 11 to 15)[a] (8)

	Number of sites (n^*)				
x	**11**	**12**	**13**	**14**	**15**
48	0.103				
49	0.087				
50	0.074				
51	0.062				
52	0.051				
56		0.102			
57		0.088			
58		0.076			
59		0.065			
60		0.055			
64			0.108		
65			0.095		
66			0.084		
67			0.073		
68			0.064		
69			0.055		
70			0.047		
73				0.108	
74				0.097	
75				0.086	
76				0.077	
77				0.068	
78				0.059	
79				0.052	
83					0.104
84					0.094
85					0.084
86					0.076
87					0.068
88					0.060
89					0.053
90					0.047

[a] For a given n^*, the table entry for the point x is $P(T^+ \geq x)$. Thus if x is such that $P(T^+ \geq x) = \alpha$, then $t(\alpha, n^*) = x$.

Large Sample Approximation (n* > 15)

Table 9A.3-1 provides critical values for T^+ for values of n^* = 4 to 15 in increments of 1. Thus a minimum n^* of 4 sites is required to perform this test. In those cases where n^* exceeds 15, a large sample approximation is used to test the significance of T^+. The following steps show the approach to making a large sample approximation (8):

Step 9—Calculate the quantity *T as follows:**

$$T^* \le \frac{T^+ - E_0(T^+)}{\sqrt{Var_0(T^+)}} \tag{9A.3-8}$$

Where:

$$E_0\left(T^+\right) = \frac{n^*\left(n^*+1\right)}{4} \tag{9A.3-9}$$

and

$$Var_0(T^+) = \frac{\left[n^*(n^*+1)(2n^*+1) - \frac{1}{2}\sum_{j=1}^{g} t_j(t_j-1)(t_j+1)\right]}{24} \tag{9A.3-10}$$

Where:

g = number of tied groups, and

t_j = size of tied group j.

Step 10—For the large-sample approximation procedure, assess the statistical significance of *T using a two-sided test at the α level of significance as follows:**

- Conclude that the treatment is statistically significant if:

$$T^* \ge z_{(\alpha/2)} \text{ or } T^* \le -z_{(\alpha/2)} \tag{9A.3-11}$$

Where:

$z_{(\alpha/2)}$ = the upper tail probability for the standard normal distribution.

Selected values of $z_{(\alpha/2)}$ are as follows:

α	$z_{(\alpha/2)}$
0.05	1.960
0.10	1.645
0.15	1.440
0.20	1.282

- Otherwise, conclude that the treatment is not statistically significant.