Recommended Protocols for Developing Crash Modification Factors

Requested by:

American Association of State Highway
and Transportation Officials (AASHTO)
Standing Committee on Highway Traffic Safety

Prepared by:

Daniel Carter, UNC Highway Safety Research Center
Raghavan Srinivasan, UNC Highway Safety Research Center
Frank Gross, Vanasse Hangen Brustlin, Inc.
Forrest Council, Vanasse Hangen Brustlin, Inc.

February 2012

The information contained in this report was prepared as part of NCHRP Project 20-07, Task 314, National Cooperative Highway Research Program, Transportation Research Board.

Acknowledgements

This study was requested by the American Association of State Highway and Transportation Officials (AASHTO), and conducted as part of National Cooperative Highway Research Program (NCHRP) Project 20-07. The NCHRP is supported by annual voluntary contributions from the state Departments of Transportation. Project 20-07 is intended to fund quick response studies on behalf of the AASHTO Standing Committee on Highways. The report was prepared by Daniel Carter, Raghavan Srinivasan, Frank Gross, and Forrest Council. The work was guided by a task group which included Michael Curtit, Karen Dixon, Scott Jones, John C. Milton, David L. Piper, Stephen W. Read, Karen Yunk, Kelly Kennedy Hardy, and Richard Pain. The project was managed by Charles W. Niessner, NCHRP Senior Program Officer.

Disclaimer

The opinions and conclusions expressed or implied are those of the research agency that performed the research and are not necessarily those of the Transportation Research Board or its sponsors. The information contained in this document was taken directly from the submission of the author(s). This document is not a report of the Transportation Research Board or of the National Research Council.

Contents

Introduction 5
Advances in the Field of Highway Safety 5
Need for High Quality CMFs 5
Layout of this Document 6
Intended Audience 7
Scope 7
Prescriptive or Not? 8
Prior Guidance on CMF Development 9
Repositories of CMFs 10
Knowledge 13
Cross-Sectional Studies 13
Before-After Studies 22
Issues That Affect All Studies 30
Selection of Appropriate Study Type and Design 34
Documentation Requirements 35
Why is Documentation Important? 35
General Documentation Requirements 37
CMF and Countermeasure 37
Site Characteristics 38
Crash Characteristics 42
Study Details 43
Biases Documentation 45
Cross-sectional Studies 45
Before-After Studies 45
Examples of Biases Documentation 46
Conclusion 48
References 49
Appendix A. Summary of Literature by Study Design 55
Before-After Study Designs 55

Cross-Sectional Study Designs 63
Appendix B. Method Correction Factor Tables from Highway Safety Manual Inclusion Process 70
Appendix C. Summary of CMF Protocols Documentation Requirements 71

INTRODUCTION

Advances in the Field of Highway Safety

The topic of highway safety is an increasingly important issue in the field of transportation engineering. In past years, state and local transportation agencies would delegate safety responsibilities on an informal basis to various staff members. However, now it is not uncommon to find clearly defined safety units in these agencies. Whereas the primary focus used to be on moving vehicles through the roadway network most efficiently, now many departments are recognizing the need to address safety at multiple levels – planning, design, operations, and maintenance. Funding and influence from the federal level spurred the development of state Strategic Highway Safety Plans (SHSP), the strategies of which are carried out through efforts such as the Highway Safety Improvement Program (HSIP). Recently, state agencies through AASHTO coordinated with the Federal Highway Administration to decide on a national highway safety vision of zero deaths (http://safety.fhwa.dot.gov/tzd).

Not only is safety becoming recognized as a critical issue, indeed the *science* of highway safety is moving forward. In past decades, engineers working to improve safety on their highways had to rely on engineering judgment alone to decide where to install countermeasures, which countermeasures were most appropriate, and how effective each countermeasure would be in reducing crashes. Resources to guide them were rare, as were quality research studies on safety issues and the effectiveness of countermeasures.

Since then, the field of highway safety has progressed and expanded. There is a recognized need for safety decisions to be driven by good quality data. Studies from universities, research centers, and departments of transportation have begun producing a multitude of quantitative estimates of crash effectiveness of countermeasures. Research topics have been developed and ushered into funded studies through groups such as the Transportation Research Board committees. The science of highway safety took a major step forward in 2010 with the release of the first edition of the AASHTO Highway Safety Manual (HSM), which documented the state of knowledge regarding road safety management, predictive crash modeling, and countermeasure effectiveness in the form of crash modification factors (CMFs). Another resource came to the field in 2009 with the launch of the FHWA Crash Modification Factors Clearinghouse, which presents an online repository of CMFs.

Need for High Quality CMFs

With the growth of the science of highway safety and the creation of resources such as the HSM and the CMF Clearinghouse, there comes the need to ensure that there is good quality research to include in such resources. The focus of this document is on the quality of studies producing crash modification factors (CMFs).

A CMF is a measure of the estimated effectiveness of a safety countermeasure. Specifically, it is a multiplicative factor used to compute the expected number of crashes at a location after implementing a specific countermeasure. The CMF might be presented in terms of a single value (point estimate) or a

function that takes into account relevant site characteristics. The primary user groups for CMFs include highway safety engineers, traffic engineers, highway designers, transportation planners, transportation researchers, and managers and administrators. CMFs can be used to:

- Capture the greatest safety gain with limited funds.
- Compare safety consequences among various alternatives and locations.
- Identify cost-effective strategies and locations.
- Check reasonableness of evaluations (i.e., compare new analyses with existing CMFs).
- Check validity of assumptions in cost-benefit analyses.

For example, a traffic engineer working under limited funding could use CMFs to choose between several countermeasures for improving safety on horizontal curves. Countermeasures under consideration might include installing centerline reflectors, adding advance curve warning signs, or improving the visibility of existing chevron signs. CMFs for each of these countermeasures would assist in estimating the potential benefit (crash savings) of each one, thereby leading to a countermeasure selection based on quantitative data rather than simply the personal preference of the parties involved.

CMFs are developed by research studies that evaluate and quantify the crash effects of countermeasures. Such studies can be designed in a variety of ways and can range in terms of quality, depth, and statistical rigor, leading to CMFs of varying quality and reliability. Those who conduct critical reviews of CMFs for resources such as the HSM and the CMF Clearinghouse use criteria such as study design, sample size, standard error, source of the data, and other potential biases to evaluate the quality of a CMF. Due to the wide range in CMF quality, there is a need for protocols to more clearly define what factors in a study design make a quality CMF that will attain high ratings in the CMF Clearinghouse and will be acceptable for inclusion in the HSM.

This document will lay out recommended protocols that should be used in the development of CMFs and will hereafter be referred to as the CMF Protocols. The major goal of these protocols will be to describe what pieces of the research study should be documented by the study authors and how various potential biases should be addressed.

Layout of this Document

This document is intended both to provide basic knowledge of CMF research for those unfamiliar with the process and to dictate what and how details of the study should be documented. The main sections of this document are as follows:

- Knowledge section – basic knowledge on each of the study types that can be used to develop CMFs and other basic issues related to the development of CMFs. It features a description of potential issues that can bias the study results and recommendations as to how these biases should be addressed.
- General documentation – list of the general details on the research study that can be used to determine where it is appropriate to apply the CMF. These items are recommended for documentation by any research study that develops CMFs.

- Biases documentation – list of the potential biases for each study design. These items are recommended for documentation by any research study that develops CMFs. This section refers heavily to the Knowledge section for description of each bias.

Appendices

A. Summary of Literature. The purpose of this appendix is to identify and briefly summarize relevant literature related to each study design. The papers are grouped according to the various study designs (e.g., before-after, cross-sectional, case-control, etc.). This summary of literature is intended to be a resource for those desiring to know more about specific study types and see examples of their use.
B. Method Correction Factor Tables from HSM Inclusion Process. This appendix presents the tables showing how HSM reviewers determined which method correction factor would be used for each CMF study that was reviewed for inclusion in the HSM.
C. Summary of the CMF Protocols Documentation Requirements. This appendix provides the list of all documentation requirements presented in this CMF Protocols document.

Intended Audience

The content of the CMF Protocols is intended for two main audiences – researchers and research sponsors. Researchers who develop CMFs are a varied group, from those with long-standing experience in highway safety research, to graduate students embarking on their first experiences with crash-based research, to personnel within a state DOT who are developing CMFs based on experiences within their state. Whatever their level of experience, the CMF Protocols will demonstrate an established standard against which their research will be judged. They will be aware of which pieces of their studies are critical and must be well documented. They will be aware of the potential biases that must be addressed in their study and what are the methods to address the influence of each potential bias.

Sponsors of research are typically found either on the national scale, such as the Federal Highway Administration (FHWA) and the National Cooperative Highway Research Program (NCHRP), or on a smaller scale such as individual state departments of transportation. Those who manage and oversee the execution of research projects might have extensive experience in crash-based research and CMF development, or perhaps more commonly, they may be somewhat unfamiliar with the process and key issues of conducting this type of research. It is often the job of these contract managers or oversight panel members to write research statements of work, issue requests for proposals, judge proposals received, and ultimately review the final product of the research. The content of the CMF Protocols can aid these individuals in stipulating what issues must be addressed in their sponsored research and will provide clear information that will help them judge more accurately proposals received and work conducted.

Scope

The CMF Protocols are intended to provide guidance for the development and documentation of research studies that develop CMFs. However, there are some boundaries of this guidance:

- It is intended only for evaluations that use crashes as the measure of effectiveness, as opposed to other safety-related measures, such as near-misses, speed reductions, or vehicle path deviations.
- It is intended only for evaluations of countermeasures that address infrastructure-related treatments, such as signs, signals, markings, alignments, and barriers. It is not intended for countermeasures or safety strategies on behavioral, enforcement, or policy actions, such as safety-related media campaigns, laws regarding driver behaviors, or enforcement efforts. These types of evaluations may use different types of study methodology that would have their own unique set of potential biases. However, many of the topics presented in this document, especially the general documentation requirements, may be useful to consider when conducting these types of research studies.

Prescriptive or Not?

When creating guidance such as the CMF Protocols, a key issue is how prescriptive should be the requirements for researchers. It is tempting to develop protocols that are very prescriptive and dictate that CMF research should be conducted with a particular study methodology and should produce a particular statistical reliability. For instance, the empirical Bayes before-after study design is widely recognized as the state-of-the-art methodology for CMF development, and the CMF Protocols could stipulate that a CMF should be developed with this method in order to be considered acceptable. However, that level of prescriptiveness is likely to interfere in a negative way with the research being conducted. The prescribed study methodology may not be appropriate for the countermeasure being studied, or it might not even be possible to do, depending on the nature of the countermeasure. It may also unnecessarily restrict the use of more innovative and advanced methods (e.g., full Bayes before-after) that may provide more insight.

For example, consider the case where an empirical Bayes before-after study was prescribed, but the research topic was evaluating the safety effect of increasing median width. It would be practically infeasible to conduct the prescribed before-after study on this countermeasure, since a transportation agency would rarely undertake the expense of expanding and realigning an entire roadway solely for the sake of widening the median. The most feasible and appropriate study design for this evaluation may be a cross-sectional regression study, where roadways with wide medians are compared to roadways with narrow medians.

Prescribing a level of statistical confidence, such as a maximum standard error, has many ramifications as well. The standard error is affected by the size of the sample and the variability of the data. The sample size used in the study is often dictated by the amount of funding available, the time period of the study, and the availability of sites that underwent the safety treatment. The variability of the data is out of the hands of the researcher. Thus, prescribing a level of statistical confidence that should be required from any CMF would put the researcher in a difficult position.

So, if being too prescriptive in this guidance is overly constraining to the practice of research, the question remains: What should be required of researchers who develop CMFs? The authors of the CMF Protocols believe the answer lies in examining the intent behind prescribing these sorts of requirements

(i.e., study method, sample size, etc.). The purpose in requiring a researcher to conduct CMF research in a specific way would be to make sure that the CMF is reliable and that all potential biases are addressed. Thus, the researcher should be required to document how he or she addressed each potential bias with the research study design that was used.

It is possible that a potential bias may be adequately addressed in more than one way, depending on the situation. For example, since treatment sites are often chosen based on high crash frequencies or rates, a before-after study is often affected by the regression-to-the-mean bias. This bias may be addressed through the use of a reference group and an empirical Bayes methodology. However, this bias may not be a concern in a situation where the safety treatment was applied to all sites within a city; there would be no selection bias in this instance. So in this case, the regression-to-the-mean bias is addressed innately in the study design.

For this reason, the CMF Protocols stress the importance of documenting whether each potential bias was relevant in the particular study, and if so, how it was addressed. This is the kind of information that is necessary to confidently rate the quality of any CMFs developed from the study.

Prior Guidance on CMF Development

In 2010, FHWA undertook an effort to develop more specific guidance for agencies interested in developing CMFs. The resulting document, *A Guide to Developing Quality Crash Modification Factors*, discusses the process for selecting an appropriate evaluation methodology and the many issues and data considerations related to various methodologies (Gross et al., 2010). The guide opens with a background of CMFs, including the definition of CMFs and related terms, purpose and application, and general issues related to CMFs. It then introduces various methods for developing CMFs. Discussion of these methods was not intended to provide step-by-step instruction for application. Rather, the guide discusses study designs and methods for developing CMFs, including an overview of each method, sample size considerations, and strengths and weaknesses. A resources section is provided to help users identify an appropriate method for developing CMFs based on the available data and characteristics of the treatment in question. The resources section also includes a discussion of considerations for improving the completeness and consistency in CMF reporting.

A Guide to Developing Quality Crash Modification Factors is very much related to the CMF Protocols and can be considered a companion document. Being familiar with the content of this guide would be very beneficial for readers of the CMF Protocols, however it is not crucial to have both documents in order to be able to follow the documentation requirements laid out here. The CMF Protocols provide an overview of each study method type, similar to the guide but in a more condensed form. Readers may refer to the guide if they desire to see more detail about a particular study type or examples of its use. A short section at the end of the guide lays out some brief guidance on what should be documented in a research study. In contrast, the focus of the CMF Protocols is on the topic of documenting details of the study and how all potential biases were addressed.

Repositories of CMFs

As mentioned previously, there are currently two main resources for CMFs, the AASHTO Highway Safety Manual and the FHWA Crash Modification Factors Clearinghouse. Both resources have review processes to determine the quality of CMFs that will be included in the resource. The intent of this document is to lay out protocols for CMF development and documentation that will guide researchers to produce CMFs that would attain high marks of quality in these two resources.

Highway Safety Manual

The Highway Safety Manual (HSM) is a document that presents science-based knowledge, tools, and resources to highway safety practitioners. The first edition of the HSM was released in 2010. The first three parts of the HSM, Parts A through C, present information on the fundamentals of highway safety, methods on road safety management, and predictive models for estimating the expected average crash frequency of a site or network.

Part D of the HSM presents over 200 crash modification factors and functions. There was a defined process to select CMFs for consideration for the Manual and evaluate them according to their reliability. This inclusion process is laid out in detail in TRB E-Circular E-C142: *Methodology for the Development and Inclusion of Crash Modification Factors in the First Edition of the HSM* (Bahar, 2010). The following provides an overview of that process.

First, a literature review procedure was conducted to identify potential CMFs from the previous decades of published studies. For each study that was examined, the following process was performed:

- *Step 1. Determine estimate of safety effect of treatment as documented in respective evaluation study publication*. This step involved determining the CMF based on the information provided in the study, whether the authors explicitly presented the CMF or whether the CMF could be derived based on information provided in the study.
- *Step 2. Adjust estimate of safety effect to account for potential bias from regression-to-the-mean (RTM) and changes in traffic volume*. If the study appeared to have been affected by a significant bias, the reviewers would adjust the CMF value to account for the bias effect.
- *Step 3. Determine ideal standard error of safety effect*. The reviewers would calculate the ideal standard error if not presented in the study. The standard error gives an indication of the relative reliability of the CMF estimate, taking into account factors such as sample size and variance inherent in the data.
- *Step 4. Apply method correction factor (MCF) to ideal standard error, based on evaluation study characteristics*. An MCF would be determined by the reviewers based on the reliability of the study, with consideration of factors such as study methodology, treatment of possible biases, treatment of potential confounding factors, and selection of appropriate functional form for regression models. Higher MCF values indicated a less reliable quality of study. It should be noted that the highest MCF values were assigned to those CMFs that had a severe lack of information published regarding study data and findings. The tables showing the relation of study characteristics to MCF values can be found in Appendix B.

- *Step 5. Adjust corrected standard error to account for bias from RTM and changes in traffic volume*. Once the MCF was applied to the standard error, the reviewers would adjust it to account for any significant biases.
- *Step 6. Combine CMFs when specific criteria are met*. In situations where multiple studies evaluated the same treatment, the reviewers would, where appropriate, combine the CMFs to produce the single, best value for that treatment.

After the literature review process was completed, the HSM reviewers filtered the list of CMFs to include only those CMFs with an adjusted standard error of 0.1 (after rounding to the first decimal) or less. In addition, CMFs with an adjusted standard error between 0.1 and 0.3 that were produced from a study that also produced CMFs with an adjusted standard error less than or equal to 0.1 were also included. CMFs in Part D are displayed in various fonts (i.e., bold, italics, etc.) to indicate their relative reliability based on the value of the adjusted standard error.

CMF Clearinghouse

The Crash Modification Factors Clearinghouse (www.cmfclearinghouse.org) is an online repository of CMFs developed and maintained by the Federal Highway Administration. Launched in December 2009, the CMF Clearinghouse provides information on thousands of CMFs and the studies that developed them. The purpose of the CMF Clearinghouse is to provide transportation professionals with a dynamic and comprehensive database of CMFs, a mechanism for sharing newly developed CMFs through a user submission form, and resources and education on the proper application of CMFs. The Clearinghouse is updated on a quarterly basis by searching sources of CMF studies such as professional journals, presentations at the Transportation Research Board annual meeting, articles indexed on the Transport Research International Documentation (TRID), state research reports, and user submitted studies.

A key component of the CMF Clearinghouse is the star quality rating that is assigned to each CMF. The ratings are from 1 to 5 stars, with 5 stars representing the best quality. Each study that is identified as a potential source of CMFs undergoes a critical review process which examines five categories to determine the star quality rating of each CMF developed in the study. The five categories are study design type, sample size, standard error, potential bias, and source of the data. The critical reviewer assigns a score of Excellent, Fair, or Poor to each category and then totals up the individual scores by means of a weighted equation to arrive at a star quality rating for the CMF. Figure 1 shows a description of the rating scores for each category.

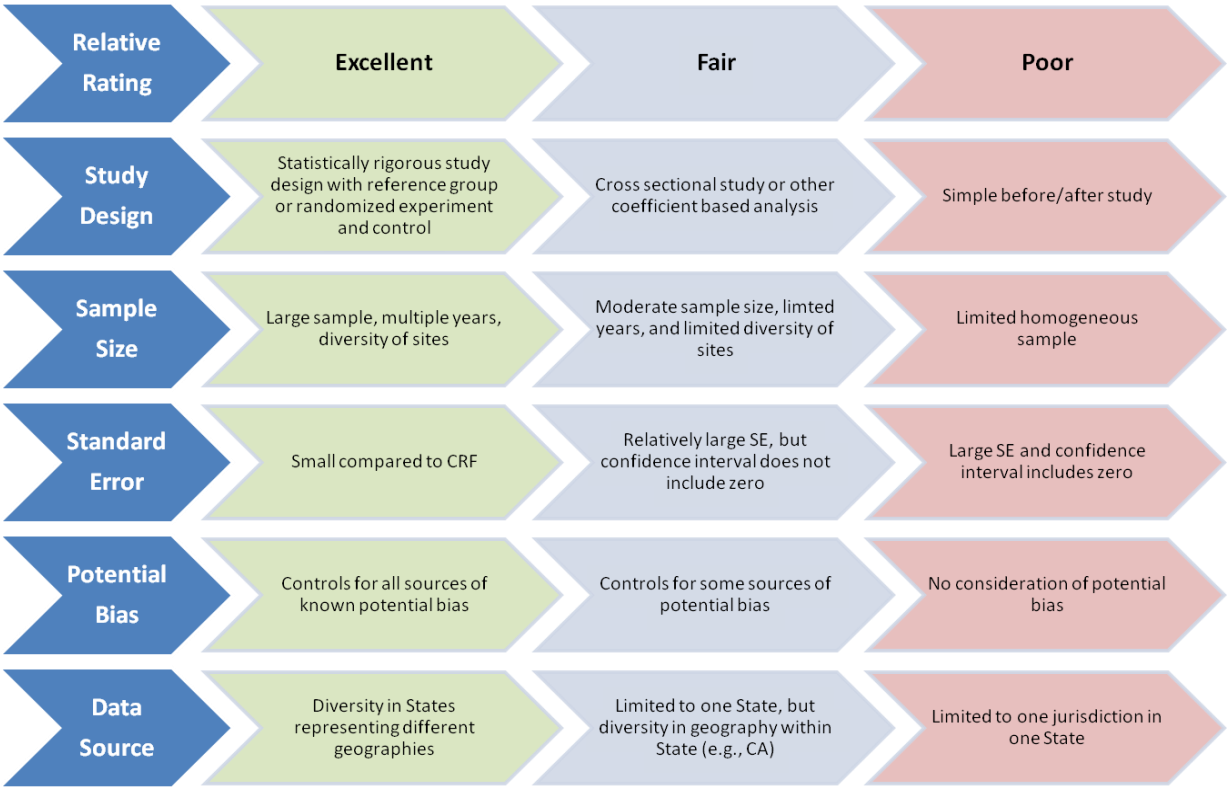


Figure 1. CMF Clearinghouse Critical Review Scoring Chart

The intent of the CMF Clearinghouse is to present all available CMFs, regardless of quality. In this manner, users of the Clearinghouse are provided with the full knowledge of the CMFs that have been published and may select which CMF is most appropriate for their situations. They are guided in their selection by the star quality rating, indicating which CMF is likely to be the most accurate and reliable, and the details about the CMF, indicating which situations are appropriate for its application. This approach differs from that of the HSM Part D, which presents on the single best CMF for any particular countermeasure. The Clearinghouse contains all the CMFs in the HSM, plus many more that did not meet the quality criteria for the HSM or were published after its release.

KNOWLEDGE

Study designs fall into two broad categories: experimental and observational. Experimental studies are planned where sites are selected at random for treatment and control. In fact, it could be argued that such experimental studies are the most rigorous way to establish causality (Elvik, 2011a). Observational studies are not planned, and typically sites are not selected as part of an experiment, but selected for other reasons including safety. Observational studies are more common because they consider safety improvements to improve the roadway system. On the other hand, experimental studies evaluate safety improvements implemented solely for the purpose of measuring their effectiveness, and are not very common partly because of potential liability considerations.

Observational studies can be broadly classified into before-after studies and cross-sectional studies. Before-after studies include "all techniques by which one may study the safety effect of some change that has been implemented on a group of entities (road sections, intersections, drivers, vehicles, neighborhoods, etc.)" (Hauer, 1997, p. 2). On the other hand, cross-sectional studies include those where "one is comparing the safety of one group of entities having some common feature (say, STOP controlled intersections) to the safety of a different group of entities not having that feature (say, YIELD controlled intersections), in order to assess the safety effect of that feature (STOP versus YIELD signs)" (Hauer, 1997, p. 2, 3). Since in a typical before-after study, we are dealing with same roadway unit located in a particular place used by probably the same users in the before and after period, these factors are less likely to confound a before-after study (Elvik , 2011a). Hauer (2005) also argues that before-after studies are less prone to confounding compared to cross-sectional studies. However, there are issues in both types of studies that need to be addressed and they are discussed below.

Following is a brief overview of different types of observational before-after and observational cross-sectional study designs and the potential issues associated with such study designs. It is important to reiterate that the scope of this discussion is limited to studies that make use of crash data to derive CMFs rather than studies that use surrogate measures.

Cross-Sectional Studies

As discussed above, CMFs from cross-sectional studies are developed by comparing the safety of a group of sites with a feature with the safety of a group of sites without that feature. The CMF can be derived by taking the ratio of the average crash frequency of sites with the feature to the average crash frequency of sites without the feature. For this method to work, the two groups of sites should be similar in their characteristics except for the feature. In practice, this is difficult to accomplish and multiple variable regression models are used. These cross-sectional models are also called safety performance functions (SPFs) or crash prediction models (CPMs). SPFs and CPMs are mathematical equations that relate crash frequency with site characteristics. The coefficients of the variables from these equations are used to estimate the CMF associated with a treatment. If the model form is log-linear (most common form), then the CMF can be obtained by taking the exponent of the coefficient (i.e., $e^{coefficient}$).

Following is an example of an SPF for non-intersection crashes on rural two-lane roads developed by Vogt and Bared (1998) using data from Washington and Minnesota. This model was estimated based on detailed information on the location of horizontal and vertical curves within each section.

$$Y = EXPO_m \exp(0.165)\exp(-0.278LW_m - 0.194SHW_m + 0.0668RHR + 0.0135DD_m + 0.139STATE)$$
$$*(\sum_i WH\{i\}\exp(0.0137DEG_m\{i\}))(\sum_j WV\{j\}\exp(0.142V_m\{j\}))(\sum_k WG\{k\}\exp(0.105GR\{k\})$$

where,

Y = predicted mean number of non-intersection accidents on the segment
$EXPO_m$ = traffic exposure in millions of vehicle kilometers
LW_m = lane width in meters
SHW_m = average of left and right shoulder widths in meters
RHR = average roadside hazard rating along segment
DD_m = driveway density in driveways per kilometer
STATE = 0 for Minnesota, 1 for Washington
$DEG_m\{i\}$ = degree of curve in degrees per hundred meters of the i-th horizontal curve that overlaps the segment
WH{i} = fraction of the total segment length occupied by the i-th horizontal curve

Since accidents are counts (i.e., non-negative integers), the Poisson regression model has been proposed as an option. However, the Poisson model restricts the mean and variance to be equal. In crash analyses, it is very common for the variance to exceed the mean, and this phenomenon is called over-dispersion. Negative binomial (NB) models are able to account for over-dispersion and have become the most common way for estimating SPFs. Typically the NB models are estimated through maximum likelihood methods (Washington et al., 2011) using an approached called generalized linear modeling. Elvik (2011a) discusses whether the coefficients from cross-sectional models represent causal relationships or non-causal statistical relationships. Elvik (2011a, quoted from p. 254-255) proposes nine operational criteria of causality:

1. Statistical association. There should be statistical association between a road safety treatment and variables measuring its effects (e.g., number of accidents, accident rate, number of injured road users).
2. Strength of association. A strong statistical association between treatment and effect is more likely to be causal than a weak statistical association.
3. Consistency of association. The statistical association between treatment and effect should be internally consistent, i.e., identical within the bounds of randomness in all subsets of data or across all similar studies.
4. Clear causal direction. The direction of causality between treatment and effect should be clear, i.e., it should be clear that treatment is (one of) the cause(s) of the effect not the other way around.

5. Control of confounders. The association between treatment and effect should exist when potentially confounding factors are accounted for (see discussion and examples below).
6. Causal mechanism. A mechanism generating the association between treatment and effect should be identified and described behaviorally or statistically.
7. Theoretical explanation. It should be possible to account for the association between treatment and effect in theoretical terms.
8. Dose-response pattern. There should a dose-response pattern in the relationship between treatment and effect (provided the treatment comes in different doses). A dose-response pattern need not be linear and may be influenced by moderating variables. [Author's note: an example of dose-response can be seen in lane width. As lane width increases (the dose), the safety benefits increase (the response), though perhaps not linearly].
9. Specificity of effect. The effect should only be found within the target group of the treatment (provided a clearly defined target group can be identified).

Among these criteria, criterion 5 that deals with control of confounders is critical. This criterion is discussed below along with other data and methodological issues with these types of models and possible ways to address them (most of this discussion is based on the content presented in Lord and Mannering (2010) and Elvik (2011a)).

Potential Issues and Biases with Cross-sectional Models

Control of confounders

A confounding variable (or confounder) is a variable that completely or partially accounts for the apparent association between an outcome and a predictor variable. Specifically, a confounder is a variable that is a significant predictor for the outcome under study, and is associated with, but not a consequence of, the predictor variable in question (Collett, 2003). For example, traffic volume is a significant predictor of crash frequency and may also be associated with several roadway characteristics such as lane width and shoulder width (i.e., roadways with higher traffic volumes are often designed to higher standards). If the safety effects of a specific design characteristic (e.g. lane or shoulder width) are in question, then the effects of traffic volume must be separated before the true effects of the variable of interest may be known (Persaud et al., 1999 and Hauer et al., 2004). This holds for many variables and emphasizes the importance of controlling for potential confounding effects when estimating the effects of the variable of interest. Reasons for confounding include lack of available data and variables that are not practical to measure or cannot be measured.

Unobserved heterogeneity and omitted variable bias

Heterogeneity refers to the differences across observations. Unobserved heterogeneity refers to differences among observations that are not accounted for in the analysis. Unobserved heterogeneity is referred to as omitted variable bias when the unobserved characteristics are correlated with a variable that is included in a model (Lord and Mannering, 2010; Elvik, 2011a). For example, if the intent is to estimate the safety effects of chevrons on horizontal curves and the curves with chevrons also have the

worst roadside hazards, but the information on the roadside hazards is not included in the model (because it was not collected), then a prediction model may incorrectly conclude that chevrons are associated with an increase in crashes. Bonnesson and Pratt (2008) discuss one way to address this problem by using matched pairs where pairs of sites are selected such that their characteristics are similar except that one site in the pair has the treatment and the site in the pair does not have the treatment. Full Bayes and hierarchical Bayes are other methods to help address issues related to unobserved heterogeneity.

Accounting for state-to-state differences if using multiple states

The use of data from multiple states in the development of cross-sectional models offers a benefit of diversifying the data. Also, it may be more possible to obtain more study sites if multiple states are used, thereby increasing the sample size. However, there may be fundamental differences between the states, such as differences in crash reporting thresholds, which can affect the results of the models. One way to account for state-to-state differences is to include indicator variables in the model to identify the sites by State.

Selection of appropriate functional form

Many studies agree that selection of an appropriate functional form and model form are critical to the estimation of reliable prediction models. For example, older studies tried to model crash rates (ratio of crash frequency to vehicle miles of traveled) while implicitly assuming that crash frequency and VMT are linearly related. It is now accepted that the relationship between crash frequency and VMT is non-linear. The current state of the art is to assume a log-linear relationship between crash frequency and site characteristics. In other words, the relationship is multiplicative. This relationship is mostly due to convenience because it allows the models to be estimated efficiently using the approach called generalized linear models (McCullagh and Nelder, 1989). Hauer (2004) argues that prediction models should include both multiplicative and additive forms, where the additive part is used for modeling point characteristics such as driveways. Full Bayes and hierarchical Bayes models allow for more complex model forms than are possible using traditional methods.

In addition to the functional form for the overall model, it is important to select the appropriate functional form for the individual variables. This includes decisions regarding whether a variable should be considered as a continuous variable instead of a categorical variable, and vice versa. Hauer (2004) discusses an approach to identify the appropriate functional form for different independent variables in a cross-sectional model. Kononov et al., (2011) also illustrate the importance of selecting the appropriate functional form for a safety performance function; their case study used data from urban freeways in Colorado and California to develop SPFs using both the traditional exponential form and sigmoid form. The sigmoid form was found to fit the data better. Xie and Zhang (2008) discussed the use generalized additive models (GAM) that provide a more flexible functional form using smoothing functions for the independent variables. However, these types of models have had limited application because they are more difficult to interpret.

A related issue with respect to the functional and model form is whether the overdispersion parameter for the negative binomial model should be a constant. Hauer (2001) argued that assuming the

overdispersion parameter to be a constant provides too much weight to shorter sections and not enough weight to longer sections, and suggested estimating the overdispersion parameter that varies based on the length of road. More recently, Cafiso et al. (2011) found the overdispersion parameter to be inversely related to segment length for rural two lane roads.

In general, a researcher should document the reason that a particular functional form was selected and why that functional form seems to be appropriate for the data.

Correlation or collinearity among the independent variables

A high degree of correlation among the independent variables makes it very difficult to come up with a reliable estimate of the effects of particular variables. For example, if horizontal curvature is correlated with the clear zone/roadside hazards, then it may be difficult to isolate the safety effect of horizontal curvature. There are no easy solutions to this problem. It may be tempting to remove one of the correlated variables, but this may lead to omitted variable bias (discussed above). Some statistical routines include tools to assess the extent of this problem. For example, one could examine the correlation matrix of the estimated parameters which will provide information about the correlation between pairs of variables. If collinearity is a concern, then the full or hierarchical Bayes method can help to reduce the issue by distinguishing heterogeneity from "noise" in the data (Orme, 2000).

Overfitting of prediction models

Overfitting of prediction models can occur when some of the relationships that appear to be statistically significant is just "noise". In other words, the model does a poor job of showing the underlying relationship. Typically, overfitting usually occurs when the model is too complex and includes too many parameters. This results in models that do not predict crashes very well and also increases the chances of correlation between the different variables in the model. One way to address this problem is using cross-validation. In cross-validation the data set is randomly divided into a two parts, where one part is used for estimating the model and the other part is used for validation. Examples of validation can be seen in two studies led by Simon Washington (Washington et al., 2001; Washington et al., 2005). Another approach is to use relative goodness of fit measures such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) that penalize models with more estimated parameters.

Low sample mean and small sample size

Since crashes are fortunately rare events, it is not uncommon to find situations where some roadway sections or intersections may have very few crashes, with many sites having zero crashes (i.e., with low sample mean). These types of data may cause problems in estimation and some models may not converge. In addition, Lord (2006) found that the overdispersion parameter that is estimated as part of a negative binomial model can be biased when the sample mean is very low. At this time, there are few solutions to this problem, except for a method suggested by Elvik (2011a) that involves "selecting a subsample with even lower mean value than the full sample and estimate model coefficients for the subsample" to see if the parameter estimates and the overdispersion parameter are stable.

Bias due to aggregation, averaging, or incompleteness in data

It is common that prediction models are based on summary data based on annual or monthly totals (i.e., aggregated data). Elvik indicates that the "use of data that are aggregated, averaged, or incomplete can

lead to biased estimates or model coefficients" (Elvik, 2011a, p. 256). One solution to this problem is to use disaggregate data, such as hourly volumes instead of AADT, but that could require the analysis to not only address the problems due to low sample mean (discussed earlier) and temporal correlation (discussed below), but also to find a source for hourly volumes. For example, Lord et al., (2005) used data on hour traffic volume from freeway sections in Montreal, Quebec, to examine the relationship between crashes, traffic flow, density, and v/c ratio. Lord et al., (2005) used the generalized estimating equations (GEE) approach to address temporal correlation. Hierarchical Bayes models can also lead to more stable parameter estimates and more accurate models with respect to both individual- and aggregate-level performance (Orme, 2000).

Temporal and spatial correlation

Temporal correlation may arise if multiple observations are used for the same roadway unit. This happens often when multiple years/months of data are used in the modeling because "many of the unobserved effects associated with a specific roadway entity will remain the same over time" Lord and Mannering (2010, p. 292). Temporal correlation can lead to incorrect estimates of the standard errors of the coefficients. Temporal correlation can be addressed using a variety of methods including generalized estimating equations (GEE) (e.g., see Lord and Persaud, 2000), random effects models (e.g., see Shankar et al., 1998), and negative multinomial models (e.g., see Hauer, 2004). Ulfarrson and Shankar (2003) in their study of median crossover crashes found that the negative multinomial model outperformed the random effects model in terms of fit, but the negative multinomial model was more difficult to estimate because of convergence problems.

Similar to temporal correlation, spatial correlation may occur because "roadway entities that are in close proximity may share unobserved effects" (Lord and Mannering, 2010). Wang and Abdel-Aty (2006) used GEE to account for spatial correlation in their examination of rear-end crashes at signalized intersections. In a study of county-level injury and fatal crashes in Pennyslvania, Aguero-Valverde and Jovanis (2006) found spatial correlation to be significant. They used the full Bayes modeling approach (discussed previously) to account for spatial correlation. Another example of the use of the full Bayes modeling approach to account for spatial correlation can be found in Guo et al., (2010).

Endogenous independent variables

There are situations when some of the independent variables may depend on the dependent variable (frequency of crashes). This is known as endogeneity. Bias due to endogeneity can lead to incorrect conclusions from a model, e.g., a model may find that a treatment is associated with an increased number of accidents, when in reality the treatment may actually reduce accidents (Elvik, 2011a). Kim and Washington (2006) show an example as part of a study to examine the safety effectiveness of left turn lanes. Since left turn lanes are likely to be implemented at intersections with high number of left-turn related crashes, a prediction model that includes the presence of left turn lanes as an independent variable is likely to suffer due to endogeneity bias. Kim and Washington (2006) illustrate one way of addressing this problem by simultaneously estimating multiple prediction equations, where in one of the equations the dependent variable was a binary variable indicating the presence/absence of a left turn lane and in the other equation, the dependent variable was crash frequency.

Misspecification of structure of systematic variation and residual terms

As mentioned earlier, negative binomial models (also called poisson-gamma models) are the most popular for modeling crash frequency. However, more recently, other types of probability models (i.e., with different assumptions on the residual terms) have been used including Poisson Log-Normal model, Zero-Inflated models, Conway-Maxwell-Poisson model, and Markov Chain Switching models (Lord and Mannering, 2010; Elvik 2011a). The Conway-Maxwell-Poisson model has been proposed as option for addressing underdispersion (i.e., the variance is less than the mean). This is not very common with crash data, but some examples have been discussed recently (e.g., Lord et al., 2010). Elvik (2011a) indicates that type of probability model does influence the coefficient estimates and consequently the CMFs that may be derived from the coefficients. Hence, care needs to be taken in selecting the appropriate option.

Correlation between crash types and injury severities

It is not uncommon for researchers to develop separate models by crash type and/or severity since the effect of a treatment can be different for different crash types/severities. However, estimating separate models independently ignores the correlation between them. Many recent papers have tried to address this issue (e.g., see El-Basyouny and Sayed, 2009) using simultaneous estimations of multiple models.

Variations of Cross-sectional Studies

The ability of cross-sectional models to adequately control confounders is questioned by some (Harwood et al., 2000). The following section discusses variations of cross sectional studies that may provide additional control.

Case-control studies

Recently, case-control studies have been used to estimate CMFs for lane and shoulder width (see e.g., Gross, 2006; Gross and Jovanis, 2007; Gross and Jovanis, 2008). Case-control studies use cross-sectional data, but sites are selected in a different way from other types of cross-sectional studies. As discussed in Gross et al., (2010, p. 30), “case-control studies select sites based on outcome status (e.g., crash or no crash) and then determine the prior treatment (or risk factor) status within each outcome group”. Another critical component of many case-control studies is the matching of cases with controls in order to control for the effect of confounding factors.

The case-control method is discussed in detail in Gross et al., (2010). Essentially, case-control methods are used to show the relative effects of treatments by using statistical techniques such as multiple logistic regression to examine the risk/benefit associated with one factor while controlling for other factors. By increasing the number of controls the probability that the test will reject a false null hypothesis is increased. On the other hand, collecting data on a larger number of controls will increase the data collection cost.

Since the number of cases and controls is predetermined, this method is very useful for studying rare events. On the other hand, traditional case-control studies do not distinguish between cases with single and multiple crashes during a specific time period. Gross (2011) recently explored this issue further and concluded that the traditional case definition underestimates the magnitude of treatment effect. The study also investigated two alternatives to address the issue. The first method is to simply create multiple cases for sites with multiple crashes (e.g., a site with three crashes during the study period would be represented as three individual cases in the database). This method would require the need to address the statistical issues due to temporal correlation that were discussed earlier. The second method defines single-crash sites separately from multiple-crash sites. In this way, a separate case-control analysis is conducted for each outcome level (e.g., sites with one crash compared to sites with no crashes, sites with two crashes compared to sites with no crashes, etc). Gross (2011) concluded that further research is necessary to determine an appropriate method to account for cases with multiple crashes in the case-control analysis. Other options to address this issue may include the use of multinomial or ordered logit models instead of binary logistic regression or to divide the time period into shorter intervals (e.g., 1 month instead of 12 months) and thereby reduce the chances of time periods with multiple crashes. However, this again would require the need to address the statistical issues due to temporal correlation that were discussed earlier.

Very few studies have compared the performance of cross-sectional regression models with case-control methods for determining CMFs except for the recent work by Gross and Donnell (2011) who found that the CMFs based on a cross-sectional regression study were similar to the CMFs from a case-control study as long as care was taken in selecting the appropriate distribution and functional form for the cross-sectional model.

The procedure for calculating the required sample size for a case-control study design is provided in page 32 of Gross et al., (2010). The sample size is a function of many factors including the following:

- Case- control ratio: number of cases divided by the number of controls.
- Magnitude of safety effect to be detected.
- Proportion of the population with the treatment.
- Significance level.
- Probability of type II error.

Cohort Studies

In Cohort studies, sites are assigned to a particular cohort based on current treatment status and followed over time to observe exposure and event frequency (Gross et al., 2010). One cohort may include the treatment and the other may be a control group without the treatment. The time to a crash in these groups is used to determine a relative risk, which is the percentage change in the probability of a crash given the treatment.

Similar to case-control studies discussed earlier, matching is sometimes used to account for potential confounding factors. There are two common types of matching: pair matching and frequency matching

(Gross et al., 2010). Since the sample is selected based on treatment status, this study type can provide valid results for rare treatments.

The equation for calculating the required sample size for a case-control study design is provided in page 37 of Gross et al., (2010). The sample size is a function of many factors including the following:

- Ratio of treatment to control group.
- Magnitude of safety effect to be detected.
- Proportion in the control group where a crash was observed.
- Significance level.
- Probability of type II error.

Sample Size Requirements in Cross-sectional Studies

Methods to estimate the required sample size for different types of cross-sectional models have been used in other fields (e.g., Signorini, 1991; Aban et al., 2008; Self and Mauritsen, 1988), however, there is very little evidence of their application in highway safety. Gross et al., (2010), indicate that there is no direct approach to determine the sample size for cross-sectional models because it depends on many factors including the number of variables included in the model, amount of variation in each variable that is included in the model, number of crashes, and the level of statistical significance required. In general, all other things being equal, a data set with a larger sample size will more likely lead to statistically significant coefficients. Hence, Gross et al., (2010) argue that one way to make a decision regarding the sample size is to see if the variables are statistically significant, i.e., "if the variables of interest are not statistically significant, then more data are required". Of course, this is not a very practical or scientific approach to the problem as the data collection would have to occur in multiple phases. As a rule of thumb, cross-sectional studies require data for many more sites than a standard before-after study. For example, a traditional before-after analysis may include 10 to 20 sites, where a cross-sectional analysis would include data for 100 or even 1,000 sites. As discussed earlier, for case-control and cohort studies, methods for estimating the required sample sizes are discussed in Gross et al., (2010).

Summary Comments for Cross-sectional Studies

Although they are not considered the most reliable and accurate method for safety evaluation, cross-sectional studies are commonly used to develop CMFs when data for other designs, such as before-after, are not available. Since "cross-section studies compare different roads, used by different road users, located at different places and subject to different weather conditions" (Elvik, 2011a, p. 263), careful attention is necessary to the modeling process and the control of confounding factors before decisions can be made regarding the reliability of the coefficients in the context of using them to estimate CMFs. The criteria outlined in Elvik (2011a) and discussed above is one way to test the reliability of the models.

Before-After Studies

As mentioned above, CMFs derived from before-after studies are based on the change in safety due to the implementation of a treatment. This section lays out the potential issues and biases that are common to before-after studies and discusses different types of before-after studies and their strengths and weaknesses.

Potential Issues and Biases with Before-After Studies

Regression-to-the-mean

Regression-to-the-mean (RTM) is the tendency of sites with abnormally high or low crash counts to return (regress) to the usual mean frequency of crashes during the following years. The bias due to RTM will arise if sites are selected for treatment based on a randomly high short-term crash count. In this case, because of RTM, crashes at the treated site may come down after the treatment is implemented due to RTM even if the treatment does not have any effect. In other words, if bias due to RTM is not addressed, a study may overestimate the safety effect of the treatment.

In many cases, sites are selected for treatment based on high crash history in the years immediately preceding the treatment. If this is the case, then crashes would be expected to decrease at some of these sites, regardless of whether any safety treatment was applied. It may be the case that these sites experienced high crashes for a few years due to the normal random fluctuation of crashes. If this is the case, then the effect of the safety treatment may be overestimated, and the resulting CMF will be too low (e.g., CMF calculated as 0.55 instead of the true value of 0.80). That is, the decrease of crashes at the treated sites in the after period may have happened anyways as a result of normal crash rate fluctuation and cannot be accurately attributed to the safety treatment.

However, if the sites to be treated are *not* selected on the basis of high crash history, then regression-to-the-mean is much less likely to bias the resulting CMF. For example, a jurisdiction-wide or "blanket" treatment would not lend itself to this bias. If all the signalized intersections in a city are converted from 8-inch signal heads to 12-inch signal heads, then there is much less likelihood of a regression-to-the-mean effect. While some treated signalized intersections may experience a decrease in crashes after treatment as a part of normal crash fluctuation, others may experience an increase after treatment as part of the same normal fluctuation, thereby effectively eliminating the regression-to-the-mean bias. Any change in crashes at the treated sites is much more likely to be due to the safety treatment (i.e., increasing signal head size), although other external effects such as weather and driver behavior changes will still need to be addressed through the use of a suitable comparison group.

RTM bias can be addressed through the use of a reference group and the empirical Bayes or Full Bayes methodology. A naïve before-after study or a before-after with only a comparison group would not properly account for the RTM bias. RTM bias could also be addressed through an experimental study design where treatment and control sites are selected randomly from the population of high crash locations.

Changes in traffic volumes

Changes in traffic volumes during the study period must be appropriately addressed. Depending on whether there is an increase or decrease in traffic volume in the after period, a study may underestimate or overestimate the safety effect of the treatment. If there are significant increases or decreases in traffic volume at the study sites during the study period, the crash rate may change. If not properly accounted for, this change may be wrongly attributed to the safety treatment. Traffic volume changes are best addressed by the use of annual traffic data to develop a crash prediction model to use in the analysis. In the empirical Bayes before-after design, this is called the safety performance function (SPF), which predicts crashes for a site based on certain site characteristics. The most crucial factor to use in developing the SPF is traffic volume. Other characteristics that might be used in the SPF are factors such as shoulder width, presence of a median, number of lanes, and presence of exclusive turning lanes. Traffic volume changes could also be addressed by the use of comparison site design.

History trends

In a before-after study it is important to account for effects due to variation in such things as weather, demography, gas prices, vehicle types, population growth, other safety treatments, or other unknown factors. Broad factors such as these could change significantly during the study period. If these factors are not accounted for, they will bias the resulting CMF. Examples of history trends are:

- Weather effects. Consider a case where a group of curves were treated with a high friction surface to decrease run-off-road crashes. If the weather during the before period consisted of several years of unusually snowy and icy winters followed by an after period with much less precipitation, then one would expect to see a decrease in the rate of run-off-road crashes, regardless of whether a high friction surface were applied.
- Vehicle types. It may be the case that the makeup of the vehicle traffic has changed within the study period. For instance, the percentage of smaller cars may have increased, leading to a potential change in the frequency and type of crashes at the study sites.
- Crash trends due to unknown factors. It may be the case that crash trends before treatment can provide some insight on the expected trends after treatment. For example, there may be a steady decrease in crashes during the before period, which could be due to a number of factors. One might expect the trend to continue in the after period regardless of the treatment, unless the underlying conditions change. In this case, it is necessary to employ a reference or comparison group to determine the general crash trends in the before and after periods. Without a reference or comparison group, it would be difficult to separate the effect of the treatment from the effect of other underlying factors.

The effect of broad factors such as these can be addressed through the use of a suitable reference or comparison group. Such a group would be composed of sites near the treatment sites, such that the selected sites are affected by the history trends in a similar manner to the treatment sites.

Other safety treatments

If other safety treatments are implemented at the treatment sites in the study period, this will affect the resulting CMF. Consider the example of evaluating the effect of high friction surface on curves. If several

of the curves were treated the next year by adding high visibility chevron signs, then this additional treatment of chevrons may assist in decreasing crashes. Any change in crashes at these sites cannot be entirely attributed to the high friction surface. Researchers must strive to ensure that no other safety treatments have been installed at the sites during the study period. This is typically done through good communication with the person in charge of the sites at the smallest level, whether that be local agency staff, district engineers, or other such positions. If there have been other safety treatments installed and the date is unknown, it is best to drop the site from the evaluation study, due to the risk of introducing bias. If there have been other safety treatments installed and the date of installation is known, it may be feasible to shorten the study time period at the affected sites to eliminate the effect of the other safety treatments on the study.

Changes in crash reporting

If there is any change in crash reporting in the jurisdiction of the treatment sites, this may lead to sharp increases or decreases in the frequencies of reported crashes. These changes would likely have nothing to do with the safety treatment of interest. Such changes may include shifting the crash reporting threshold, assigning responsibilities of crash reporting to a different agency, or a change in the way that crash reporting is performed (e.g., converting from paper crash forms to electronic submittals). The researcher should be aware of any changes in crash reporting and properly account for those changes by modifying the study period or adjusting the expected crash frequencies accordingly. The use of a reference or comparison group can also help to account for these general temporal changes.

Accounting for state-to-state differences if using multiple states

The use of data from multiple states offers a benefit of diversifying the data. Also, it may be possible to obtain more study sites if multiple states are used, thereby increasing the sample size. However, there may be fundamental differences between the states, such as differences in crash reporting thresholds, which can affect the results of the models. Selecting reference and comparison sites from the same states as the treatment sites can account for the differences between states.

Suitability of comparison or reference groups

If a before-after study makes use of a comparison or reference group, it is essential to ensure that the groups are suitable. A suitable comparison group should be a group of sites that are vulnerable to the same external effects (e.g., weather, economic changes, and changes in driver demographics) as the treatment sites, in order to correctly account for historical trends. This typically means that comparison sites will be located in the same general vicinity and jurisdiction as the treatment sites. They should also have similar crash trends in the before period as the treatment sites. Gross outlines several basic methods for evaluating the suitability of a comparison group (Gross et al., 2010, pp. 14-16).

A suitable reference group should be a group of sites that are similar to the treatment sites in terms of geometric and traffic volume characteristics but did not receive the safety treatment. A reference group is used to develop a safety performance function (SPF) for the treatment sites. The SPF is incorporated in the empirical Bayes approach to properly account for regression-to-the-mean bias. If a reference group is located in the general vicinity of the treatment sites, it may also serve as a good comparison group.

Caution should be used to make sure that the comparison or reference sites are not affected by the safety treatment of interest. The use of unsuitable reference or comparison sites may lead to under- or over-predicting crashes for the before period of the treatment sites.

Other Biases for Before-After Studies

With the exception of the naïve study design, before-after studies typically involve the development of a safety performance function (SPF) to predict expected average crashes for sites based on site characteristics. The expected crashes calculated from the SPF are then used in the derivation of the CMF. The development of the SPF is an exercise in cross-sectional modeling; therefore, the biases that are relevant to cross-sectional models are relevant to before-after studies in this respect. However, as mentioned earlier, in a before-after comparison group study or empirical Bayes before-after study, coefficients from the SPFs are not used to estimate CMFs, but the SPFs are primarily used to predict the average number of crashes for a certain AADT level and site characteristics. Harwood et al., (2000) indicate that "regression models are very accurate tools for predicting the expected total accident experience for a location or class of locations, but they have not proved satisfactory in isolating the effects of individual geometric or traffic control features". In other words, even if the coefficients from the SPFs are not reliable enough to develop CMFs, the predictions from the SPF are considered reliable enough to be used as part of a before-after comparison group or empirical Bayes study. In addition, by selecting the reference group to be similar to the treatment group in terms of the major risk factors, we can reduce the possible bias due to confounding on these predictions. Nonetheless, the researcher should be aware of these biases and ensure that they have been accounted for sufficiently. For discussion of each bias, the reader should refer to the corresponding point in the section above on cross-sectional studies.

- Selection of appropriate functional form
- Correlation or collinearity among the independent variables
- Overfitting of prediction models
- Low sample mean and small sample size
- Bias due to aggregation, averaging, or incompleteness in data
- Temporal and spatial correlation
- Endogenous independent variables
- Omitted variable bias
- Misspecification of structure of systematic variation and residual terms
- Correlation between crash types and injury severities

Types of Before-After Study Designs

The section below presents various types of before-after study designs that are used in the research field currently. The study types are described here in a general overview; for more detail, the reader is encouraged to refer to the greater detail and examples found in *A Guide to Developing Quality Crash*

Modification Factors (Gross et al, 2010) as well as the various studies referenced throughout this section and Appendix A (Summary of Literature by Study Design).

Naïve before-after study

A naïve before-after study is a simple before-after comparison where the safety effect of a treatment is assessed by directly comparing the crash frequency in the after period with the crash frequency in the before period (Hauer, 1997). The typical naïve before-after study design does not account for possible bias due to RTM and does not account for temporal effects or trends. A variation of this study design may consider the comparison of crash rates (per vehicle miles traveled (VMT)) between the before and after periods rather than crash counts. Crash rates attempt to consider the change in traffic volume between the before and after periods, but the use of rates implicitly assumes that crash frequency and traffic volume are linearly related. Many studies have shown that the relationship between crash frequency and traffic volume is often non-linear.

Before-after study with comparison group

In this study design, an untreated comparison group that is similar to the treatment group is identified to account for temporal effects and changes in traffic volume. One way to apply this method is to use the comparison group to first calculate a *comparison ratio* which is the ratio of observed crash frequency in the after period to that in the before period (Gross et al., 2010). The observed crash frequency in the before period at the treatment sites is multiplied by this comparison ratio to estimate the expected number of crashes in the after period had the treatment not occurred. The expected number of crashes in the after period (had the treatment not occurred) is then compared with the actual number of crashes in the after period to determine the safety impact of the treatment. In order for this approach to work, we have to assume that the trends in the crash counts in the treatment and comparison group are similar. Hauer (1997) proposes a test that makes use of a sequence of sample odds ratios to determine if the trends in the two groups are indeed similar. By selecting the comparison sites within the same jurisdiction as the treatment sites, it is more likely to be able to identify sites that will have similar trends as the treatment sites. However, when identifying sites from the same jurisdiction we need to be also concerned by spill-over effects, which can occur for treatments such as red-light cameras because they are usually associated with jurisdiction-wide publicity and in some applications drivers may not know which sites have the cameras (Persaud et al., 2005).

Another possible approach for applying this method is to develop safety performance functions using data from the comparison group. As mentioned earlier in this document, safety performance functions (SPFs) are mathematical equations that relate crash frequency with site characteristics including AADT. After developing SPFs, the ratio of the predicted crash frequency in the after period to the predicted crash frequency in the before period is used to estimate the comparison ratio. By estimating an SPF, the effect of traffic volume changes from the before to the after periods is accounted for. Since SPFs are cross-sectional models, many of the issues discussed earlier regarding cross-sectional models have to be addressed with some important differences. Specifically, in a before-after comparison group study or empirical Bayes before-after study (discussed below), coefficients from the SPFs are not used to estimate CMFs, but the SPFs are primarily used to predict the average number of crashes for a certain AADT level and site characteristics. Harwood et al., (2000) indicate that “regression models are very

accurate tools for predicting the expected total accident experience for a location or class of locations, but they have not proved satisfactory in isolating the effects of individual geometric or traffic control features". In other words, even if the coefficients from the SPFs are not reliable enough to develop CMFs, the predictions from the SPF are considered reliable enough to be used as part of a before-after comparison group or empirical Bayes study. The comparison group method does not explicitly account for RTM. Hence, it is a viable approach if sites are not selected based on crash history (e.g., blanket installation of a treatment) or a long before period is available reducing the possible bias due to RTM.

Chapter 9 of Hauer (1997) has a detailed discussion of sample size considerations for a before-after study with a comparison group, including the equations to be used in the calculations.

Empirical-Bayes before-after study

The empirical Bayes (EB) method is outlined in Hauer (1997). The EB method has been shown to be able to address the possible bias due to RTM, account for changes in traffic volume, and account for temporal effects. As with all before-after designs, the intent of the EB procedure is to estimate the expected number of crashes that would have occurred had there been no treatment and compare that with the number of reported crashes after the treatment was implemented. The EB procedure leads to a much more accurate estimate of the expected number without treatment.

The following steps are used to estimate the expected number of crashes that would have occurred had there been no treatment:

1. Identify a reference group of sites without the treatment, but similar to the treatment sites in terms of the major factors that affect crash risk including traffic volume and other site characteristics.
2. Using data from the reference site, estimate safety performance functions (SPFs) using data from the reference sites relating crashes to independent variables such as traffic volume and other site characteristics. As discussed earlier, since SPFs are cross-sectional models, it is important to address the issues associated with confounding. As discussed in the following steps, SPFs are used in the EB method to predict the average number of crashes based on AADT and site characteristics. By selecting the reference group to be similar to the treatment group in terms of the major risk factors, we can reduce the possible bias due to confounding on these predictions.
3. In estimating SPFs, calibrate annual SPF multipliers to account for the temporal effects (e.g., variation in weather, demography, and crash reporting) on safety. The annual SPF multiplier is the ratio of the observed crashes to the predicted crashes from the SPF. In using the annual SPF multipliers from the SPFs to account for temporal effects, it is assumed that the trends in the crash counts are similar in the treatment and reference groups. As mentioned earlier, Hauer (1997) proposes a test that makes use of a sequence of sample odds ratios to determine if the trends in the two groups are similar.
4. Use the SPFs, annual SPF multipliers, and data on traffic volumes for each year in the before period for each treatment site to estimate the number of crashes that would be expected for the before period in each site.
5. Calculate the EB estimate of the expected crashes in the before period at each treatment site as the weighted sum of the actual crashes in the before period and predicted crashes from step 4.

6. For each treatment site, estimate the product of the EB estimate of the expected crashes in the before period and the SPF predictions for the after period divided by these predictions for the before period (for each work zone site). This is the expected number of crashes that would have occurred had there been no treatment. The variance of this expected number of crashes is also estimated in this step.

The expected number of crashes without the treatment along with the variance of this parameter and the number of reported crashes after the treatment is used to estimate the index of effectiveness, which is also referred to as the crash modification factor and the variance of the crash modification factor.

There are some variations to how this method could be applied. For example, Sayed et al., (2010) argue for the use of a comparison group (independent of the reference group to develop the SPFs) to account for temporal effects. In many other studies (e.g., Srinivasan et al., 2008a, 2008b), the annual factors from the SPFs (mentioned in Step 3 above) are used to account for the temporal effects.

Currently, there are no accepted methods for determining the required sample sizes for before-after studies that employ the empirical Bayes method. One option is to use the method discussed in Chapter 9 for Hauer (1997) for before-after comparison group studies with the understanding that the sample size estimates from that method could be considered conservative. For an example, readers are referred to the section on *Study Design* in Persaud et al., (2007).

Full or Hierarchical Bayes Before-After Study

As discussed above in the case of empirical Bayes methods, a reference group is used to estimate the expected crash frequency from a SPF. These estimates are then combined with the observed crash frequency in the before period of the treatment group to estimate the long-term expected crash frequency without the treatment. In the case of the full or hierarchical Bayes approach, the distribution of likely values from the reference group is used instead of the point estimate. By using the distribution of likely values, more accurate estimates of the CMF and its variance are possible (Gross et al., 2010).

Another advantage of the full Bayes and hierarchical Bayes methods is being able to explicitly model spatial correlation and other issues related to unobserved heterogeneity. Spatial correlation can be an issue where both treatment and comparison sites are close to each other. This method allows the analyst to make use of the added information related to the spatial distribution of the sites (sites that are closer together are more likely to have similar factors than sites that are further apart). Unobserved heterogeneity is an issue when the unobserved characteristics are correlated with a variable that is included in a model. The full Bayes and hierarchical Bayes methods can help to reduce the issue by distinguishing heterogeneity from “noise” in the data (Orme, 2000). These Bayesian methods also allow the use of prior knowledge of CMFs estimated from other studies. Recently, full Bayes methods have been employed in several before-after studies (e.g., see Lan et al., 2009; El-Basyouny and Sayed, 2011).

Similar to a before-after study with the empirical Bayes method, there is no accepted method for calculating the required sample size for a before-after study employing the full Bayes method. As

discussed earlier with the before-after empirical Bayes method, a conservative estimate of the sample size can be determined assuming a before-after study design with a comparison group.

Intervention and Time Series Analysis Methods

Intervention and time series analysis methods have been used in safety for many years. One of the most common methods is the use of the Autoregressive Integrated Moving Average (ARIMA) method (Box and Jenkins, 1976) that explicitly accounts for serial correlation and seasonality. The ARIMA method assumes that the data are normally distributed and this assumption has been shown to be appropriate for aggregate data, such as the number of crashes at a county or state level. In highway safety, examples include the use of ARIMA for determining the safety effect of graduated driver licensing (Masten and Hagge, 2003), system-wide changes in speed limits (Pfefer et al., 1991), and pedestrian treatments (Zegeer et al., 2008).

When the unit of analysis is less aggregate (e.g., a roadway segment or intersection rather than a county or state), then the assumption of normal distribution may no longer be appropriate. In such situations, a count data model (e.g., Poisson, Poisson-gamma, Poisson-log-normal) is estimated along with the use of either GEE, random effects, and/or random parameters to account for the correlation. Typically, two sets of time series data (one for the sites that received the treatment and the other for comparison sites that did not receive the treatment) are included and analysed together using an equation that includes variables such traffic volume, covariates for other site characteristics, an indicator variable to represent treatment versus comparison group, an indicator variable to represent before versus after periods, and variables to account for the trend effect. This type of analysis is also sometimes called panel data analysis and it includes some of the features of both time series and cross-sectional analysis methods. El-Basyouny and Sayed (2011) presents an example where such a model was used to estimate the safety effect of treatments at intersections in Vancouver. These models typically rely on a long before period to account for possible bias due to RTM.

Methods to determine the appropriate sample size for ARIMA time series analysis are available from McLeod and Vingilis (2005, 2008). McLeod and Vingilis (2008) illustrate their method as part of a study to evaluate the impact of extended drinking hours in Ontario.

Sample Size in Before-After Studies

Knowing what sample size is necessary or desired is often an important and relevant topic for researcher sponsors and researchers alike. The goal should be to use a sample that is large enough to produce a statistically significant CMF (i.e., a CMF whose confidence interval does not include 1.0). However, each situation is unique; the necessary sample size depends on several factors. As discussed in Gross et al., (2010), the following factors can be used to determine if a statistically significant CMF can be obtained:

1. Number of crashes in the before period in the treatment group.
2. The relative duration of the before and after period.
3. The likely CMF value.
4. Number of crashes in the before and after periods in the comparison group.

Gross et al. (2010) present some numerical examples showing the effect of increasing the treatment sample or the comparison group sample. This approach is intended for studies using before-after design with comparison groups. As discussed earlier, for empirical Bayes and full Bayes before-after studies, there is not an established method for calculating sample size, but the method discussed here is often used for these study types as well.

For more detailed information on determining necessary sample sizes, the reader is referred to Chapter 9 of Observational Before-After Studies in Road Safety (Hauer, 1997), that has a detailed discussion of sample size considerations for a before-after study with a comparison group, including the equations to be used in the calculations.

For ARIMA time series analysis studies, methods discussed in McLeod and Vingilis (2005, 2008) can be used to estimate the required sample size.

Issues That Affect All Studies

Data Quality

An obvious concern for any research study is the quality of the data used to produce the results. For CMF development research, data quality can greatly affect the accuracy and reliability of the produced CMF. Researchers should take care to ensure that they are aware of the quality of the data they use and should communicate any concerns about data quality in the research report. Inspections of quality should involve an examination of any data gaps, origins of the data, and suspicious trends in the data. In particular, researchers should pay close attention to two crucial types of data that are often seen to have quality issues – crash data and traffic volume data.

Crash data

Although state agencies are typically adamant about maintaining good quality crash data, there can be issues that arise when using these data to develop CMFs. Researchers should be careful to examine the crash data for suspicious or counterintuitive trends that may bear further investigation. Researchers should also strive to obtain an understanding about the crash data they will use and what changes may have occurred in the study period that would affect CMF results. For example, the agency may have changed the reporting threshold (the dollar amount of damage under which a crash is not required to be reported) and the effect can be seen in an abrupt increase or decrease in crash frequency from one year to the next. Another example would be a switch from paper form reporting to electronic reporting, which could also cause change in the number of crashes reported. If changes like these are known, the researcher may consider changing the time period of the study to avoid these issues. Also, study designs that use nearby comparison sites would assist in mitigating this effect, if the comparison sites are affected in the same way as the treatment sites.

Traffic volume data quality

Traffic volumes (often referred to in terms of Annual Average Daily Traffic, AADT) are a major predictor of crash frequency. However, it is the unfortunate case that traffic volume is often an imprecise

measurement. Agencies do not have the funding available to constantly monitor volumes on all roads; therefore they must rely on infrequent and sometimes widely spaced point counts and extrapolate traffic volumes geographically (between count points) and temporally (over multiple years). Often this extrapolation also includes assumptions about general traffic growth in the area over the time period of interest. This problem can be exacerbated in rural areas, where lower development density and greater area of the roadway network leads to larger spacing between the count points. Similar to crash data verification, researchers should closely inspect any volume data used in the development of CMFs and investigate suspicious trends, gaps, or abrupt increases or decreases. Any known issues with the traffic volume data should be documented in the research report.

Crash Type and Severity

Although countermeasures may have an effect of generally decreasing crashes where they are installed, it is often the case that the countermeasure will have different effects on different crash types. For instance, the installation of a signal will generally decrease right angle and left turn crashes, since conflicting movements are controlled in a stop-and-go fashion. However, a signal will generally increase the incidence of rear end crashes, as more drivers are stopping, sometimes unexpectedly. The same effect can be seen in the installation of red light running cameras (Persaud et al., 2005).

Additionally, certain countermeasures are implemented to target a specific type of crash. For instance, the installation of high friction surface would be intended to decrease wet weather crashes. The installation of cable median barrier would be intended to decrease cross median head-on crashes. Countermeasures may also have a different effect on different crash severities. The cable median barrier that reduces cross median head-on crashes may cause an increase in property damage only (PDO) crashes, as more vehicles hit the barrier instead of crossing the median.

Given this, it is important for researchers to examine the effect of a countermeasure, not just on the target crash type, but on all crash types and severities. This knowledge is useful for practitioners who are considering the full impacts of installing certain countermeasures. Instead of developing one general CMF for a countermeasure, it is more useful to develop CMFs for various crash types and crash severities to fully understand and communicate the effect of the countermeasure.

Crash Modification Functions versus Crash Modification Factors

There are at least two ways of representing the effect of safety treatments on crashes. One is through crash modification factors (CMFs), which are essentially point estimates, e.g., a CMF of 0.87 for automated speed enforcement indicates that on average a 13% reduction can be expected. The second way of representing the effect of treatments is crash modification functions (CMFunctions). Often, the effect of a treatment may depend on several factors including traffic volume and site characteristics, and in such cases, crash modification functions would be more appropriate. For example, the effect of automated speed enforcement might differ based on traffic volume. Crash modification functions are not necessarily a new concept. In fact, some CMFs in the Highway Safety Manual (AASHTO, 2010) are in the form of Crash Modification Functions. For example, the safety effect of radius, length, and presence of spiral transition curves in a horizontal curve is given as follows:

$$\text{CMF} = \frac{(1.55L_c) + (\frac{80.2}{R}) - (0.012S)}{1.55L_c}$$

Where:

L_c = Length of horizontal curve including length of spiral transitions, if present (in miles)

R = Radius of curvature (in feet)

S = 1 if spiral transition curve is present, 0 if spiral transition curve is not present

This CMFunction is based on work done in the early 1990's (Zegeer et al., 1992).

As in the case of CMFs, the CMFunctions can be developed using different methods including cross-sectional methods and before-after studies. With cross-sectional models, CMFunctions may be derived directly from the multivariate model as in the case of example shown above for horizontal curves. The accuracy of the function will depend on all the potential biases applicable to cross-sectional models. With before-after empirical Bayes studies, CMFunctions can be derived by taking the CMF estimate for each site that was used in the evaluation, and developing a function through regression analysis. Such an approach was used to develop a CMFunction to determine the effect of AADT when signalized intersections were replaced by roundabouts (Srinivasan et al., 2011).

CMFunctions can also be developed when combining CMFs from multiple studies as illustrated by Elvik (2011b) during a presentation at the *International Workshop on Transferability of Crash Modification Factors* in January 2011 at the Annual Meeting of the Transportation Research Board. Elvik (2011b) used weighted regression to develop a CMFunction to determine the safety effect of the relative change in enforcement, where weights were used to account for the difference in precision (i.e., variance) between the CMFs from multiple studies.

Standard Error: A Measure of CMF Precision

The most common way of expressing the precision of a CMF is through the standard error of the mean value of the CMF (hereafter referred to as 's'). 's' is the square root of the variance of the mean value of the CMF. A relatively high value of 's' implies less certainty about the mean value of the CMF, whereas a relatively low value implies greater certainty about the mean value of the CMF. The formulas for computing the standard error of the mean value of the CMF in a before-after study are provided in Hauer (1997). Bahar (2010) discusses different ways of estimating the approximate value of the standard error of the mean value of CMFs derived from a cross-sectional model.

The standard error of the mean value of the CMF is typically used to compute a confidence interval which provides some indication of the range of values for the mean value of the CMF. If the confidence interval does not include 1.0, then it implies that the mean value of the CMF is statistically significant at the given confidence level. On the other hand, if the confidence level includes 1.0, then the mean value of the CMF is not statistically significant at the given confidence level. Part D of the Highway Safety Manual (AASHTO, 2010) recommends that such CMFs should be used with caution because the

treatment associated with that CMF could result in a reduction in crashes, no change in crashes, or an increase in crashes.

As discussed in Gross et al., (2010), the confidence interval can be computed as follows:

Confidence interval = CMF $\pm$ (Cumulative Probability*Standard Error)

For a 99% confidence interval, the cumulative probability is 2.576; for a 95% confidence interval, the cumulative probability is 1.960, and for a 90% confidence interval, the cumulative probability is 1.645.

Hauer et al., (2012) argues that in addition to estimating and reporting 's' (the one most often reported in the published literature including the Highway Safety Manual (AASHTO, 2010)), it is important to estimate the standard deviation of the *distribution of the CMF* (hereafter referred to as 'σ'). σ explicitly accounts for the fact that a CMF is a random variable which does not necessarily have a unique value, and depends on many factors including the characteristics of the sites where the treatment may have been implemented. Hauer et al., (2012) feels that σ is a better indicator of the precision of the CMF when users are trying to determine the safety effectiveness of the treatment for a future application in a different area or jurisdiction. Hauer et al., (2012) outlined an approach to estimate σ making use of the Law of Total Variance (Weiss, 2005).

Effect of Sample Size on Standard Error and Deviation

Sample size is based on the total number of crashes in the analysis. Sample size can be increased by including more sites or increasing the duration of the study period. In discussing the effect of sample size, it is important to distinguish between the effect on the standard error of the mean value of the CMF ('s') versus the effect on the standard deviation of the CMF distribution (σ). By increasing the sample size, we can reduce 's'. However, since σ more explicitly considers the variation of the CMF as a function of the circumstances (e.g., site characteristics, traffic volume, etc.), increasing the sample size will not necessarily reduce σ. σ can be reduced by an understanding of the relationship between the CMF and the circumstances based on a Crash Modification Function (Hauer et al., 2012).

Adjusting a Standard Error

The standard errors for the CMFs in Part D of the HSM are adjusted standard errors which were computed by multiplying the reported standard errors with a method correction factor (MCF) (Bahar, 2010). The MCFs were a function of several factors including the type of study design (before-after versus cross-sectional) and whether the following factors were accounted for:

- Possible bias due to regression to the mean in a before-after study
- Potential confounding factors in a cross-sectional study
- Traffic volume in any type of analysis
- The appropriateness of the functional form of the crash prediction model in a cross-sectional study

The MCFs in the HSM ranged from 1.2 to 7.0 with a lower value assigned for a better study and the higher value for a poorly designed and executed study. See Appendix B for tables showing MCF values for various study designs and characteristics.

Selection of Appropriate Study Type and Design

As discussed earlier, since cross-sectional studies are based on the comparison of roadway units with a common feature to roadway units that do not have that feature, they are more prone to confounding (Hauer, 2005; Elvik, 2011a). There are issues with before-after studies as well (e.g., bias due to regression to the mean), but there is more consensus among the research community about different ways to address the issues associated with before-after studies. Such a consensus does not yet exist to address the issues associated with cross-sectional studies. However, as discussed earlier in this document, there are some treatments for which before-after studies would simply not be feasible. One example is the safety of median width. Agencies rarely expand or realign a roadway section just for the sake of changing the median width. In such cases, cross-sectional studies would be appropriate for developing a CMF. Following is some discussion about the different factors that may be considered in selecting an appropriate study design (part of this discussion is based on information from Sections 3.8 and 4.0, as well as Figure 5 from Gross et al., 2010).

To conduct a before-after study, information is needed on a sufficient number of sites where a sufficiently similar treatment has been installed or can be installed. If there is possibility of bias due to RTM, then a reference group can be used to develop SPFs to apply the empirical Bayes method, or a sufficiently long before period should be considered to minimize the possible bias due to RTM for applying a time series or intervention analysis method. A comparison or reference group is necessary to account for temporal effects. If spatial correlation is a factor or if prior knowledge of CMF estimates need to be included in the evaluation, then a full Bayes approach would be a better approach compared to the empirical Bayes method. The full Bayes approach is also more flexible and can be applied even if the sample of reference sites is limited, or if the functional form of the SPF is complex (e.g., needs to include additive and multiplicative terms).

Cross-sectional analysis will work if it is possible to find sufficient locations with and without the treatment that are otherwise similar. In addition, data must be available for the major factors affecting crash risk. Negative binomial regression is the most common approach for cross-sectional regression. Recently, other approaches including the Poisson Log-Normal model, Zero-Inflated models, Conway-Maxwell-Poisson model, and Markov Chain Switching models, have been proposed for different situations. If the crash type being affected is rare then case-control methods can be a useful option. If the treatment is rare (i.e., limited locations), then cohort methods may be an option.

A Guide to Developing Quality Crash Modification Factors presents a useful flowchart that assists in determining the appropriate study design to use when conducting safety evaluations based on the type and amount of data available (Gross et al, 2010, p. 50).

DOCUMENTATION REQUIREMENTS

Why is Documentation Important?

Proper documentation of a study is an important issue that is often overlooked in the process of presenting research results. While results are indeed the goal of research, it is the reality of the research field that any study can produce some kind of results. Equally as important is the quality or reliability of those results. It is only through good documentation that a reader can assess the quality of a study. This can be demonstrated in an example of two studies; each one focused on developing a CMF for installing larger signal heads at signalized intersections.

- Researchers in the first study, constrained by limited funding, examined seven intersections along a single corridor in a city where larger signal heads were installed. The researchers looked only at these seven intersections and conducted a simple before-after study in which they compared three years of crash data before installation to one year of crash data after installation.
- Researchers in a second, well-funded study identified 20 signalized intersections in each of three cities, for a total number of 60 study sites that received installations of larger signal heads. These researchers also identified a reference group of 60 signalized intersections that were similar to the treatment group but did not receive larger signal heads. The researchers conducted an empirical Bayes before-after study using seven years of crash data before installation and three years of crash data after installation.

If the CMFs developed in both studies were presented with only *brief* descriptions, they may appear comparable. Each would be described as estimating the crash effect of installing larger signal heads. However, good documentation from each study would quickly reveal some key differences – namely the sample size (7 intersections vs. 60 intersections), geographic coverage (one city vs. three cities), methodology (simple before-after vs. empirical Bayes before-after with reference group), and years of crash data (4 total years vs. 10 total years). It is also likely that the standard error of the first study would be larger than that of the second study, due to the fact that only a very small sample size was used. All of these are factors that would affect the quality or reliability of the CMF.

This document presents two categories of documentation requirements. The first is *general documentation*, which indicates where the CMF may be applied, and the second is *bias documentation*, which indicates how the researchers addressed the potential biases of the study.

General documentation would focus on information about the countermeasure that was studied, the resulting CMF(s), and the applicability of the CMF(s) developed. It is important for the end user of a CMF to know in which situations it is appropriate to apply that CMF. Applicability may be related to site characteristics, such as traffic volume range, roadway class, number of lanes, and area type, or to characteristics of the crashes affected by the countermeasure, such as crash type, crash severity, or time of day. For example, if a researcher develops a CMF for a particular countermeasure and the group of sites in the study had a traffic volume range of 10,000 to 30,000 vehicles per day, it would be risky to

assume that the CMF would accurately reflect the effect on the countermeasure on a road with 1,200 vehicles per day.

Bias documentation would focus on factors affecting the quality of the CMF – how the design and execution of the study accounted for potential confounding factors. Reviewers for the HSM and the CMF Clearinghouse examine factors such as the type of study methodology, sample size used, source of the data, standard error, and influence of biases when determining the quality of a CMF. It is essential that this information be documented in the study so that reviewers will be able to accurately rate the quality and reliability of CMFs developed in the study.

The data elements that should be documented in any CMF research report are listed in the sections below. If a study produced more than one CMF (e.g., one CMF for all crashes, one CMF for fatal crashes, and one CMF for fatal and injury crashes), all data described below should be clearly presented for each CMF.

GENERAL DOCUMENTATION REQUIREMENTS

The elements listed in this section are those factors about a study which are critical for a user to know where the CMF can be appropriately applied. Each element is described and a few examples are listed. A researcher who is authoring a study should document each one of these elements in the report or paper. The examples given under each element are for illustration purposes only. The author should describe each element as it pertains to the specific facilities used in the study.

CMF and Countermeasure

Countermeasure Name and Description

Description of the countermeasure for which the CMF was developed. The description should be detailed enough to be clear to a reader exactly what was implemented at the study sites.

Example:

- "Installed three-strand high-tension cable median barrier." Simply stating "installed median barrier" would be insufficient description.
- "Installed overhead and sign post mounted flashing beacons. Beacons flashed yellow for the major road and red for the minor road." Simply stating "installed flashers" would be insufficient description.

Crash Modification Factor or Function

Crash modification factor (CMF) or function (CMFunction) produced by the study. If the study produced more than one factor or function, each should be clearly described. The value of a CMF should be greater than zero, since it is a multiplicative factor. It may be helpful to present the results in terms of crash reduction factors (CRFs) as well, to prevent any confusion. Although CMF is becoming the standard form of presentation, many engineers are accustomed to thinking in terms of percentage reduction (CRF).

Example CMFs:

- CMF = 0.79 (CRF = 21%)
- CMF = 1.08 (CRF = -8%)

When listing CMFunctions, all terms in the function should be clearly labeled. If multiple functions are developed, it may be more succinct to present the general form once and the coefficients in a table.

Example CMFunction:

- CMF for changing driveway density = $e^{0.0152(Y-X)}$ where X is the prior number of driveways per mile and Y is the new number of driveways per mile.

Measures of Precision for the CMF

Standard error of the mean value of CMF should be presented for each CMF or CMFunction developed in the study in the same manner and units. In addition, the standard deviation of the CMF distribution is also a useful measure because it provides further information regarding the variability of the CMF among the sites used in the evaluation.

Prior Conditions

Prior conditions of the site before countermeasure was implemented. The prior condition of the study sites (also referred to as the "baseline" condition) should be described in sufficient detail that a user may accurately determine if the sites to be treated with the countermeasure of interest are sufficiently similar to the prior condition of the study sites.

Example:

- If the countermeasure is "Convert stop controlled intersection to signalized intersection", prior condition of the study sites may be "minor-road stop controlled" or "All-way stop controlled".
- If the countermeasure is "Prohibit trucks from traveling in left lane", the prior condition of the study sites may be "Truck lane restriction sign posted prohibiting trucks traveling 50 mph or below from the left lane".

Site Characteristics

Roadway Class

Roadway class(es) of the study sites.

Examples:

- Interstate
- Freeway or Expressway
- Other principal arterial
- Minor arterial
- Major collector

- Minor collector
- Local

Road Division Type

Type of road division of the study sites.

Example:

- Undivided
- Divided by raised median
- Divided by depressed/traversable median
- Divided by two-way-left-turn-lane (TWLTL)
- Divided by barrier

State

The state, states, or countries in which study sites were located.

Example:

- California
- Missouri, Kentucky, and Michigan
- United Kingdom

Municipality

The municipality or municipalities in which study sites were located.

Example:

- Charlotte, North Carolina
- Phoenix, Arizona; and Cambridge, Massachusetts
- Madrid, Spain

Area Type

Area type of study sites.

Example:

- Rural

- Urban
- Suburban

Number of Through Lanes

Number of through lanes in both directions. If the study sites included intersections, the author should provide the number of through lanes for both major and minor road.

Example:

- 2 lane roads
- 4 and 6 lane roads
- 4 to 6 lane major roads and 2 lane minor roads

Speed Limit

Posted or statutory speed limit of the road or range of speed limits at study sites. If the study sites were intersections, the author should provide the speed limit for the major road at a minimum but preferably for both the major and minor roads.

Example:

- 35 to 45 mph
- 45 to 55 mph major roads and 35 mph minor roads

Traffic Volume Range

Range of traffic volumes of all the study sites. If the study sites were intersections, the author should provide the range of traffic volumes for both major and minor roads.

Example:

- 12,000 to 22,000 vehicles per day
- 18,000 to 23,000 vehicles per day for major roads and 5,000 to 8,000 vehicles per day for minor roads

Traffic Control (if applicable)

Traffic control at the study intersections.

Example:

- Signalized
- Stop control for minor road only
- Stop control for all approaches
- Stop control for minor road only with flashers
- Stop control for all approaches with flashers
- Yield control for minor road
- Roundabout
- No control

Intersection Type

Type of study intersections as defined by the types of routes that are intersecting.

Example:

- Roadway to roadway (not interchange related)
- Roadway to roadway (interchange ramp terminal)
- Roadway to pedestrian crossing (e.g., midblock crossing)
- Roadway to bicycle path or trail
- Roadway to driveway

Intersection Geometry

Geometric configuration of the study intersections. This describes the general design of the intersection, including the number of legs. If using circular intersections, describe the type of intersection, including details such as traffic control, geometrics, priority/right-of-way, and other details to distinguish between definitions of roundabouts, traffic circles, and other types of circular intersections.

Example:

- Cross intersection (4 legs)
- Cross intersection (more than 4 legs)
- T-intersection (3 legs)
- Y-intersection (3 legs)
- Circular intersection (3 legs)
- Circular intersection (4 legs)
- Circular intersection (more than 4 legs)

Other Relevant Details

The author should include other details about the study sites that may be particularly relevant to the countermeasure. For example, if the countermeasure were installing cable median barrier, it would be relevant to document the range of median widths of the study sites. If the countermeasure were installing shoulder rumble strips, it would be relevant to document the range of paved shoulder widths and lane widths of the study sites.

Crash Characteristics

Crash Type

Type of crash targeted by the CMF. Some CMFs may be more generally directed at all crashes, whereas others may be developed to estimate the effect on specific crash types (target crashes).

Example:

- All crashes
- Angle crashes
- Rear end crashes
- Wet weather crashes
- Run-off-road crashes

Crash Severity

Severity of crash targeted by the CMF. Some CMFs may be more generally directed at all crash severities, whereas others may be developed to estimate the effect on specific crash severities (target crashes).

Example:

- All crashes
- Fatal crashes
- Fatal and injury crashes (KABC)
- Fatal and injury crashes (KAB)

Time of Day

Time of day of the crashes targeted by the CMF.

Example:

- All times of day

- Day time
- Night time

Study Details

Years of Data

Beginning and ending dates of the time period of the data used in the study.

Example:

- 2002 to 2008
- 2002 to 2004, 2006 to 2008 (if 2005 data were omitted due to transition period effects)

Type of Methodology

Type of study methodology used to develop the CMF.

Example:

- Simple (naïve) before-after
- Empirical Bayes before-after
- Cross-sectional regression model
- Cross-sectional non-regression model
- Case-control

Site Selection Criteria

Criteria used in selecting sites for treatment. It is important for a researcher to document how the treated sites were chosen. The sites may have been selected on the basis of high crash history, particular characteristics, or other reasons. Sometimes all sites of a particular type in a jurisdiction are treated. Knowing how the sites were selected for treatment is important for determining how well the study methodology accounted for potential biases and for determining where (what types of locations) the CMF may be appropriately applied by the end user.

Example:

- Sites with history of high crash frequency (e.g., curves known to be high crash locations)
- Sites with a particular characteristic (e.g., curves with radius less than “X” feet)
- All sites of a certain facility type in the jurisdiction (e.g., all curves in the state)

Sample Size Used (Crashes)

Number of crashes used to develop the CMF. The information should be presented so that it is clear to the reader exactly how much data served as the basis for the development of the CMF. If an author breaks down the data in a disaggregate analysis, such as developing a CMF for all crashes and then developing individual CMFs for rear end crashes and angle crashes, it is expected that the sample size of crashes will be smaller for the individual crash types.

If a before-after study methodology is used, the author should present the "before" sample size and "after" sample size.

Example:

- 320 total crashes
- 45 target crashes (i.e., run-off-road or nighttime crashes)
- 456 crashes in the before period, 345 crashes in the after period

Sample Size Used (Sites)

Size of the sample used to develop the CMF. This may be presented in terms of number of sites, number of miles, or other units that include the time period of data availability, such as mile-years. Among these, the number of crashes is generally the most important. Whatever units are used, the information should be presented so that it is clear to the reader exactly how much data served as the basis for the development of the CMF.

If a before-after study methodology is used, the author should present the "before" sample size and "after" sample size.

Example:

- 47 intersections
- 312 miles of roadway
- 1800 mile-years (300 miles of roadway with 6 years of data for each mile)
- 2100 mile-years before; 1400 mile-years after

BIASES DOCUMENTATION

The second set of documentation requirements deals with potential biases. A research study needs to address each of the potential biases and document whether each potential biasing issue could have affected the development of the resulting CMF, and if so, how did the study design account or correct for that bias. Just as in the Knowledge section of this document, the lists below are structured according to the two basic families of study types, before-after and cross-sectional. Under each study type are listed the potential biases that are relevant to that study type. A researcher reporting on his or her research using one of these study types should present a section in the report that clearly lists each potential bias and states how the research design or the nature of the data addresses each potential bias. Description and discussion of each potential bias is provided in the respective section under the Knowledge section of this document.

Cross-sectional Studies

The following potential biases should be addressed for cross-sectional studies:

- Control of confounders
- Unobserved heterogeneity and omitted variable bias
- Accounting for state-to-state differences if using multiple states
- Selection of appropriate functional form
- Correlation or collinearity among the independent variables
- Overfitting of prediction models
- Low sample mean and small sample size
- Bias due to aggregation, averaging, or incompleteness in data
- Temporal and spatial correlation
- Endogenous independent variables
- Misspecification of structure of systematic variation and residual terms
- Correlation between crash types and injury severities

Before-After Studies

The following biases should be addressed for before-after studies:

- Regression-to-the-mean
- Changes in traffic volumes
- History trends
- Other safety treatments
- Changes in crash reporting
- Accounting for state-to-state differences if using multiple states
- Suitability of comparison or reference groups
- Other biases for before-after studies

Examples of Biases Documentation

The following are examples of how potential biases would be adequately addressed and documented. Each example provides two cases (hypothetical situations) in which the bias was adequately addressed.

Example 1.

Study: Before-after study to evaluate the effect of converting permissive left turn signals to protected left turn signals.

Potential bias: Regression-to-the-mean.

Example answers that adequately address this bias:

- Case 1: Regression-to-the-mean bias was addressed in this study through use of empirical Bayes methodology with a reference group. The reference group was composed of signalized intersections with permissive left turns which were not converted to protected left turns. The reference group also matched the treatment group on characteristic ranges such as number of lanes, number of intersection legs, and traffic volume.
- Case 2: Regression-to-the-mean bias was not a potential bias for this study. The treatment was applied to all signalized intersections in the entire city. The treatment sites were not selected on a basis of high crash history, therefore any bias due to regression-to-the-mean is assumed to be negligible.

Example 2.

Study: Before-after study to evaluate the statewide installation of curve warning signs.

Potential bias: Changes in crash reporting.

Example answers that adequately address this bias:

- Case 1: The research team determined that no changes in crash reporting took place during the time period of interest. This is based on direct questions in conversations with the state safety engineer and state law enforcement personnel.
- Case 2: The research team was informed that a policy change was made during the first year of the study period in which law enforcement changed from using paper crash forms to submitting crash reports electronically. Out of concern that this could affect the crash trends, the before period was shortened by one year to eliminate this bias.

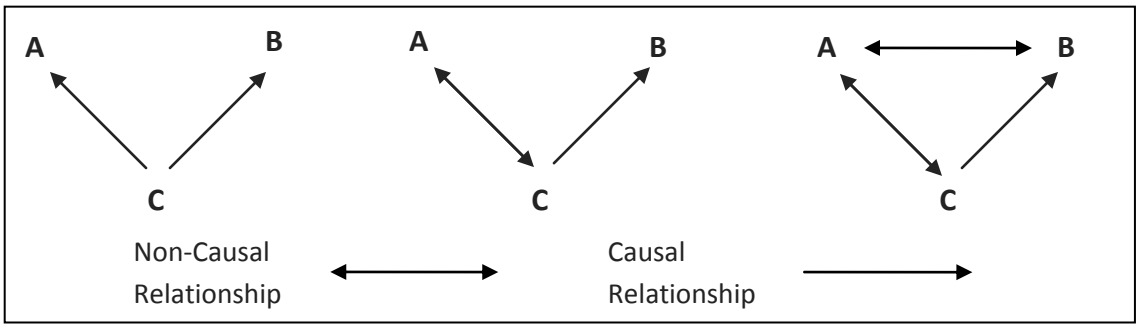
Example 3.

Study: Cross-sectional study to evaluate the effect of intersection skew angle.

Potential bias: Control of confounders.

Example answers that adequately address this bias:

- Case 1: The research team developed a list of potential confounders based on intuition and a review of previous studies related to paved shoulder width. The list included lane width, unpaved shoulder width, traffic volume, posted speed limit, roadside hazard rating, horizontal and vertical curvature, weather, and driver demographics. The following relationships were tested to identify the most likely confounding variables. Note that C is a confounder of the relationship between the explanatory variable A (paved shoulder width in this case) and the outcome B (crash frequency in this case) if any of the following relationships hold.



  Lane width, unpaved shoulder width, traffic volume, posted speed limit, and roadside hazard rating were identified as the most likely confounders while horizontal and vertical curvature were noted as less likely confounders. The research team concluded that weather and driver demographics were not likely to be confounders as there is not a clear relationship (either causal or non-causal) between these variables and paved shoulder width. Data were collected for all likely confounders and included in the multiple variable regression model to assess the safety effect of paved shoulder width.

- Case 2: Potential confounding factors were identified at the outset of the research project, similar to Case 1, and matching was employed to account for potential confounding effects. Specifically, the researchers matched segments with different paved shoulder widths by the most likely confounding factors (i.e., lane width, unpaved shoulder width, traffic volume, posted speed limit, and roadside hazard rating) to ensure proper adjustment for confounding effects and included variables that represented the less likely confounders in the model (i.e., horizontal and vertical curvature).

Example 4.

Study: Cross-sectional study to evaluate the effect of median width.

Potential bias: Correlation or collinearity among the independent variables.

Example answers that adequately address this bias:

- Case 1: Several explanatory variables were included in a model relating median width and other roadway characteristics to crash frequency. The correlation matrix was computed to investigate potential bias related to correlation. The correlation matrix indicated that lane width and paved

shoulder width are highly correlated. It was desirable to account for these variables in the analysis as they were identified as potential confounders and removing them could lead to omitted variable bias. To account for potential bias related to correlation, these variables were combined and represented as a single variable (total paved width).

- Case 2: Several explanatory variables were included in a model relating median width and other roadway characteristics to crash frequency. The correlation matrix was computed to investigate potential bias related to correlation. The correlation matrix indicated that there was no significant correlation among the independent variables and the researchers concluded that bias related to correlation was unlikely.

A summary of the documentation requirements is presented in Appendix C.

CONCLUSION

Good, data-driven decisions are essential in the field of highway safety. As engineers and other transportation professionals make decisions based on quantitative data, the role of estimates of countermeasure effectiveness such as CMFs is clearly important. As CMFs are reviewed and rated for quality and reliability, the reviewers must be sure about specifics of how the study was conducted, both in terms of applicability and potential biases. This document presents the CMF Protocols as the recommended best practice for documenting the key characteristics of the study so that both audiences, reviewers and practitioners, will be confident that they are using the CMFs with the full knowledge of their reliability and how they are best applied.

REFERENCES

AASHTO (2010), *Highway Safety Manual*, American Association of State and Highway Transportation Officials, Washington, D.C.

Aban, I.B., Cutter, G.R., and Mavinga, N. (2008), Inferences and Power Analysis Concerning Two Negative Binomial Distributions with an Application to MRI Lesion Counts Data, *Computational Statistics and Data Analysis*, Vol. 53(3), pp. 820-833.

Aul, N. and G.A. Davis (2006), Use of Propensity Score Matching Method and Hybrid Bayesian Method to Estimate Crash Modification Factors of Signal Installation. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1950, Transportation Research Board of the National Academies, Washington, DC.

Bahar, G. (2010), Methodology for the Development and Inclusion of Crash Modification Factors in the First Edition of the Highway Safety Manual, *Transportation Research Circular* E-C142, Transportation Research Board.

Bahar, G., M. Masliah, C. Mollett, and B. Persaud (2003). *Integrated Safety Management Process*. NCHRP Report 501, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC. Available online at: http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_501.pdf <Accessed May 17, 2010>

Bonneson, J.A. and M.P. Pratt (2008), Procedure for Developing Accident Modification Factors from Cross-Sectional Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2083, Transportation Research Board of the National Academies, Washington, DC.

Box, G.E.P, and G.M. Jenkins, (1976), *Time Series Analysis: Forecasting and Control*, (2nd Edition), Holden-Day, San Francisco.

Cafiso, S., Di Silvestro, G., Persaud, B., and Begum, M. (2010), Revisiting Variability of Dispersion Parameter of Safety Performance for Two-Lane Rural Roads, *Transportation Research Record* 2148, pp. 38-46.

Carlin, B.P. and T.A. Louis (2008), *Bayes and Empirical Bayes Methods for Data Analysis*, Third Edition. Chapman and Hall/CRC, New York, NY.

Carriquiry, A. and M.D. Pawlovich (2005), *From Empirical Bayes to Full Bayes: Methods for Analyzing Traffic Safety Data*. White Paper, October. Available online at: http://www.iowadot.gov/crashanalysis/pdfs/eb_fb_comparison_whitepaper_october2004.pdf <Accessed July 12, 2011>

Collett, D., 2003. *Modelling Binary Data.* Second Edition. New York: Chapman and Hall/CRC.

El-Basyouny, K. and Sayed, T. (2009), Collision Prediction Models using Multivariate Poisson-Lognormal Regression, *Accident Analysis and Prevention*, Vol. 41(4), pp. 820-828.

El-Basyouny, K. and Sayed, T. (2011), A Full Bayes Multivariate Intervention Model with Random Parameters Among Matched Pairs for Before-After Safety Evaluation, *Accident Analysis and Prevention*, Vol. 43, pp. 87-94.

Elvik, R. (2011a), Assessing Causality in Multivariate Accident Models, *Accident Analysis and Prevention*, Vol. 43, pp. 253-264.

Elvik, R. (2011b), Crash Modification Factors versus Functions, Presented at the *International Workshop on Transferability of Crash Modification Factors*, TRB 2011 Annual Meeting, January 23rd, 2011.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin (2003), *Bayesian Data Analysis*, Second Edition. Chapman and Hall/CRC, New York, NY.

Griffith, M.S. (2000), Safety Evaluation of Rolled-In Continuous Shoulder Rumble Strips Installed on Freeways, *Journal of the Institute of Transportation Engineers*, Vol. 70, Issue 6, 37-42.

Gross, F. (2006), *A Dissertation in Civil Engineering: Alternative Methods for Estimating Safety Effectiveness on Rural, Two-Lane Highways: Case-Control and Cohort Methods*, The Pennsylvania State University, December 2006.

Gross, F. and Jovanis, P.P. (2007), Estimation of the Safety Effectiveness of Lane and Shoulder Width: The Case-Control Approach, *Journal of Transportation Engineering*, Vol. 133(6).

Gross, F. and P.P. Jovanis (2008), Estimation of Safety Effectiveness of Changes in Shoulder Width using Case-Control and Cohort Methods. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2019, Transportation Research Board of the National Academies, Washington, DC.

Gross, Frank, Bhagwant Persaud, and Craig Lyon (2010), *A Guide to Developing Quality Crash Modification Factors, Federal Highway Administration*, Report FHWA-SA-10-032. Available at http://safety.fhwa.dot.gov/tools/crf/resources/fhwasa10032/. Accessed October 2011.

Gross, F. (2011), Case-Control Analysis in Highway Safety: Accounting for Sites with Multiple Crashes. *3rd International Conference on Road Safety and Simulation*, Indianapolis, IN, September 14-16, 2011.

Gross, F. and Donnell, E. (2011), Case-Control and Cross-Sectional Methods for Estimating Crash Modification Factors: Comparisons from Roadway Lighting and Lane and Shoulder Width Safety Effect Studies, *Journal of Safety Research*, Vol. 42(2), pp. 117-129.

Gross, F. (2011), Case-Control Analysis in Highway Safety: Accounting for Sites with Multiple Crashes. 3rd International Conference on Road Safety and Simulation, Indianapolis, IN, September 14-16, 2011.

Guo, F., Wang, X., and Abdel-Aty, M. (2010), Modeling Signalized Intersection Safety with Corridor-level Spatial Correlations, *Accident Analysis and Prevention*, Vol. 42(1), pp. 84-92.

Harwood, D., Council, F., Hauer, E., Hughes, W. and Vogt, A. (2000), *Prediction of the Safety Performance of Rural Two-Lane Highways*, Report FHWA-RD-99-207, Federal Highway Administration.

Harwood, D.W., K.M. Bauer, I.B. Potts, D.J. Torbic, K.R. Richard, E.R. Kohlman Rabbani, E. Hauer, and L. Elefteriadou (2002). *Safety Effectiveness of Intersection Left- and Right-Turn Lanes*. Federal Highway Administration, FHWA-RD-02-089, Washington, DC.

Hauer, E. (1997), *Observational Before-After Studies in Road Safety*, Elsevier Science, Oxford, UK.

Hauer, E. (2001), Overdispersion in Modeling Accidents on Road Sections and in Empirical Bayes Estimation, *Accident Analysis and Prevention*, Vol. 33(6), pp. 799-808.

Hauer, E., D.W. Hardwood, F.M. Council, and M.S. Griffith (2002). Estimating Safety by the Empirical Bayes Method: A Tutorial. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1784, Transportation Research Board of the National Academies, Washington, DC.

Hauer, E., F.M. Council, and Y. Mohammedshah, (2004). Safety models for urban four-lane undivided road segments. *Transportation Research Record 1897*, Transportation Research Board, National Research Council, Washington, D.C., pp. 96-105.

Hauer, E. (2004), *Speed and Crash Risk: An Opinion*. Public Policy Department, Royal Automobile Club of Victoria, Report No. 04/02, May.

Hauer, E. (2004), Statistical Road Safety Modeling. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1897, Transportation Research Board of the National Academies, Washington, DC.

Hauer, E. (2008), How many accidents are needed to show a difference? *Accident Analysis and Prevention*, Vol. 40, Issue 4, 1634-1635.

Hauer, E. (2010), Cause, Effect, and Regression in Road Safety: A Case Study. *Accident Analysis and Prevention*, Vol. 42, Issue 4, 1128-1135.

Hauer, E., Bonneson, J., Council, F., Srinivasan, R., and Zegeer, C. (2012), Crash Modification Factors: Foundational Issues, *Presented at the 91st Annual Meeting of the Transportation Research Board*, Washington, D.C.

Karlaftis, M.G. and G. Yannis (2010). Weather Effects on Daily Traffic Accidents and Fatalities: Time Series Count Data Approach. *Transportation Research Board 89th Annual Meeting Compendium of Papers*, Paper #10-0325, Washington, DC.

Kim, D. and Washington, S. (2006), The Significance of Endogeneity Problems in Crash Models: An Examination of Left-Turn Lanes in Intersection Crash Models, *Accident Analysis and Prevention*, Vol. 38(6), pp. 1094-1100.

Kononov, J., Lyon, C., and Allery, B.K. (2011), Relationship of Flow, Speed, and Density of Urban Freeways to Functional Form of a Safety Performance Function, *Transportation Research Record: Journal of the Transportation Research Board*, No. 2236 (pp. 11-19), Transportation Research Board of the National Academies, Washington, DC.

Lan, B., Persaud, B., and Lyon, C. (2009), Validation of a Full Bayes Methodology for Observational Before-After Road Safety Studies and Application of Evaluation of Rural Signal Conversions, *Accident Analysis and Prevention*, Vol. 41(3), pp. 574-580.

Lord, D. and Persaud, B. (2000), Accident Prediction Models With and Without Trend: Application of the Generalized Estimating Equations Procedure, *Transportation Research Record* 1717, pp. 102-108.

Lord, D., Manar, A., and Vizioli, A. (2005), Modeling Crash-Flow-Density and Crash-Flow-V/C Ratio Relationships for Rural and Urban Freeway Segments, *Accident Analysis and Prevention*, Vol. 37(1), pp. 185-199.

Lord, D. (2006), Modeling Motor Vehicle Crashes using Poisson-Gamma Models: Examining the Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter, *Accident Analysis and Prevention*, Vol. 38(4), pp. 751-766.

Lord, D. and Geedipally, S.R., and Guikema, S. (2010), Extension of the Application of Conway-Maxwell-Poisson Models: Analyzing Traffic Crash Data Exhibiting Under-Dispersion, *Risk Analysis*, Vol. 30(8), pp. 1268-1276.

Lord, D. and Mannering, F. (2010), The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives, *Transportation Research Part A*, Vol. 44, pp. 291-305.

McLeod, A.I. and Vingilis, E.R. (2005), Power computations for intervention analysis, *Technometrics*, 47, pp. 174-180.

McLeod, A.I. and Vingilis, E.R. (2008), Power computations in time series analysis for traffic safety interventions, *Accident Analysis and Prevention*, 40, pp. 1244=1248.

Masliah, M., G. Bahar, and E. Hauer (2007), Application of Innovative Time Series Methodology to Relationship between Retroreflectivity of Pavement Markings and Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2019, Transportation Research Board of the National Academies, Washington, DC.

Masten, S. and Hagge, R. (2003), *Evaluation of California's Graduated Driver Licensing Program*, Report CAL-DMV-RSS-03-205, California Department of Transportation, Sacramento, CA.

McCullagh, P. and Nelder, J.A. (1989), *Generalized Linear Models*, CRC Press.

Milton, J. and F. Mannering (1998), The Relationship among Highway Geometrics, Traffic-Related Elements and Motor-Vehicle Accident Frequencies. *Transportation*, Vol. 25, 395–413.

Orme, B. (2000). Hierarchical Bayes: Why All the Attention?, Sawtooth Software Research Paper Series, Sawtooth Software Inc., Sequim, WA

Pawlovich, M.D., W. Li, A. Carriquiry, and T.M. Welch (2006), Iowa's Experience with Road Diet Measures: Use of Bayesian Approach to Assess Impacts on Crash Frequencies and Crash Rates. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1953, Transportation Research Board of the National Academies, Washington, DC.

Persaud, B., C. Lyon, and T. Nguyen, (1999). Empirical Bayes Procedure for Ranking Sites for Safety Investigation by Potential for Safety Improvement. *Transportation Research Record 1665*, pp. 7-12.

Persaud, B. (2001), *Statistical Methods in Highway Safety Analysis: A Synthesis of Highway Practice*. NCHRP Synthesis 295, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC. Available online at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_295.pdf <Accessed May 17, 2010>.

Persaud, B., B. Lan, C. Lyon and R. Bhim (2010), Comparison of Empirical Bayes and Full Bayes Approaches for Before-and-After Road Safety Evaluations. *Accident Analysis and Prevention*, Vol. 42, Issue 1, 38–43.

Persaud, B. and C. Lyon (2007), Empirical Bayes Before–After Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions. *Accident Analysis and Prevention*, Vol. 39, Issue 3, 546–555.

Persaud, B., Council, F., Lyon, C., and Griffith, M. (2005), Multi-Jurisdictional Safety Evaluation of Red Light Cameras, *Transportation Research Record* 1922.

Persaud, Bhagwant; Craig Lyon; Kimberly Eccles; Nancy Lefler; and Roya Amjadi (2007), *Safety Evaluation of Increasing Retroreflectivity of STOP Signs*, Report FHWA-HRT-08-041, Federal Highway Administration.

Pfefer, R.C., Stenzel, W.W., and Lee, B.D. (1991), Safety Impact of the 65-mph Speed Limit: A Time Series Analysis (Abridgment). *Transportation Research Record: Journal of the Transportation Research Board*, No. 1318, Transportation Research Board of the National Academies, Washington, DC.

Sayed, T., deLuer, P., and Pump, J. (2010), Impact of Rumble Strips on Collision Reduction on Highways in British Columbia, Canada: Comprehensive Before-and-After Safety Study, *Transportation Research Record* 2148, pp. 9-15.

Self, S.G. and Mauritsen, R.H. (1991), Power/Sample Size Calculations for Generalized Linear Models, *Biometrics*, Vol. 44(1), pp. 79-86.

Shankar, V., F. Mannering, and W. Barfield (1995), Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies. *Accident Analysis and Prevention*, Vol. 27, Issue 3, 371-389.

Signorini, D.F. (1991), Sample Size for Poisson Regression, *Biometrika*, Vol. 78(2), pp. 446-450.

Srinivasan, R., Council, F., Lyon, C., Gross, F., Lefler, N., and Persaud, B. (2008a), Safety Effectiveness of Selected Treatments at Urban Signalized Intersections, *Transportation Research Record* 2056, pp. 70-76.

Srinivasan, R., Carter, D., Persaud, B., Eccles, K., and Lyon, C. (2008b), Safety Evaluation of Flashing Beacons at Stop-Controlled Intersections, *Transportation Research Record* 2056, pp. 77-86.

Srinivasan, R. et al. (2011), Evaluation of Safety Strategies at Signalized Intersections, *NCHRP Report* 705, Transportation Research Board.

Xie, Y. and Zhang, Y. (2008), Crash Frequency Analysis with Generalized Additive Models, *Transportation Research Recor*d 2061, pp. 39-45.

Ulfarsson, G. F. and Shankar V.N. (2003) An accident count model based on multi-year cross-sectional roadway data with serial correlation, *Transportation Research Record*, Vol. 1840, pp. 193-197.

Wang, X. and Abdel-Aty, M. (2006), Temporal and spatial analyses of rear-end crashes at signalized intersections, *Accident Analysis and Prevention*, Vol. 38(6), pp. 1137-1150.

Washington, S., M. Karlaftis, and F. Mannering (2010), *Statistical and Econometric Methods for Transportation Data Analysis*, Second Edition. Chapman and Hall/CRC, New York, NY.

Washington, S., J. Leonard, D.G. Manning, C. Roberts, B. Williams, A.R. Bacchus, A. Devanhalli, J. Ogle, and D. Melcher (2001), *Scientific Approaches to Transportation Research Volumes 1 and 2*. NCHRP Online Report 20-45, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC. Available online at: http://onlinepubs.trb.org/Onlinepubs/nchrp/cd-22/start.htm

Washington S., B. Persaud, C. Lyon, and J. Oh. (2005), *Validation of Accident Models for Intersections*. Federal Highway Administration, FHWA-RD-03-037, Washington, DC.

Weiss, N. (2005), *A Course in Probability*, Addison–Wesley.

Woodward, M. (2005), *Epidemiology: Study Design and Data Analysis*, Second Edition. Chapman and Hall/CRC, New York, NY.

Zegeer, C. V., J. R. Stewart, F. M. Council, D. W. Reinfurt, and E. Hamilton (1992), Safety Effects of Geometric Improvements on Horizontal Curves, *Transportation Research Record 1356*, Transportation Research Board.

Zegeer, C. et al., (2008), Evaluation of Miami-Dade Pedestrian Safety Demonstration Project, *Transportation Research Record* 2073, pp. 1-10.

APPENDIX A. SUMMARY OF LITERATURE BY STUDY DESIGN

The purpose of this literature review is to identify and briefly summarize relevant literature related to each study design. The following list identifies several salient pieces of literature and describes how each piece relates to the specific study design. Specifically, the summary indicates whether the document provides an overview, a detailed discussion, equations, examples, and/or case studies. The summary also notes if the document identifies or discusses potential issues and biases associated with the method.

Before-After Study Designs

Empirical Bayes

Observational Before-After Studies in Road Safety

This text provides a detailed description of the empirical Bayes before-after method. It thoroughly discusses how the study design addresses the issue of the regression-to-the-mean bias, traffic volume changes, time trends, and the suitability of reference groups.

Hauer, E. Observational Before–After Studies in Road Safety. Pergamon Press, Oxford, UK, 1997.

Estimating Safety by the Empirical Bayes Method: A Tutorial

This document provides a description of the empirical Bayes method for estimating the long-range expected crash frequency of a roadway entity. This document does not, however, demonstrate any before-after evaluations. Instead, it presents 9 examples in which the empirical Bayes method is applied to compute the expected crash frequency of a roadway entity for different scenarios. The scenarios include situations in which different roadway entities are considered (i.e., segments and intersections), different timeframes are considered (1-year versus multiple years), and different severities are considered. The description includes equations on how to compute the empirical Bayes weight for the various scenarios. This study discusses how the empirical Bayes method addresses the regression-to-the-mean bias.

Hauer, E., D.W. Hardwood, F.M. Council, and M.S. Griffith. Estimating Safety by the Empirical Bayes Method: A Tutorial. Transportation Research Record: Journal of the Transportation Research Board, No. 1784, Transportation Research Board of the National Academies, Washington, DC, 2002.

Statistical Methods in Highway Safety Analysis, A Synthesis of Highway Practice

This document briefly describes the empirical Bayes method, and it states that the method addresses the issues of regression-to-the-mean bias, traffic volume changes, and time trends. It also presents two examples in which a CMF is computed using this method (pp. 62-63).

Persaud, B. Statistical Methods in Highway Safety Analysis, A Synthesis of Highway Practice. NCHRP Synthesis 295, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2001. Available online at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_295.pdf <Accessed May 17, 2010>.

Safety Effectiveness of Intersection Left- and Right-Turn Lanes

This document provides an overview of the empirical Bayes method and contrasts it with two other types of before-after evaluations: before-after with yoked comparison and before-after with comparison group. This comparative analysis was performed by applying each of the three methods in an evaluation of adding left-turn and right-turn lanes at intersections. It succinctly explains that the empirical Bayes method accounts for regression-to-the-mean bias while other before-after methods do not (p. 134).

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. Federal Highway Administration, FHWA-RD-02-089, Washington, DC, July, 2002.

Use of Empirical Bayesian Methods to Estimate Crash Modification Factors for Daytime Versus Nighttime Work Zones

In this paper, authors describe the use of empirical Bayesian methods to develop crash modification factors for various conditions, including time of day (e.g., daytime, nighttime), work status (e.g., activity, inactivity), and temporary traffic control (e.g., lane closure, no lane closure).

Srinivasan, R., Ullman, G., Finley, M., and Council, F., Use of Empirical Bayesian Methods to Estimate Crash Modification Factors for Daytime Versus Nighttime Work Zones, Transportation Research Record 2241, 2011, pp. 29-38.

Crash Modification Factors for Changing Left Turn Phasing

In this paper, authors describe the use of empirical Bayesian and comparison group methods to develop crash modification factors for changing from permissive to protected-permissive phasing and the implementation of flashing yellow arrow. The paper also raises the importance of estimating the standard deviation of the CMF distribution in addition the more commonly used standard error of the mean value of the CMF.

Srinivasan, R., Lyon, C., Persaud, B., Baek, J., Gross, F., Smith, S., and Sundstrom, C., Crash Modification Factors for Changing Left Turn Phasing, Presented at the 2012 Annual Meeting of Transportation Research Board, Washington, D.C., January 2012.

Empirical Bayes Before–After Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions

This document first gives an overview of the empirical Bayes method and provides the equations used in this method. It then makes the case that the empirical Bayes method is superior to conventional methods for safety evaluations, because it addresses the regression-to-the-mean bias which can be substantial depending on the situation. The document also

discusses at length the importance of properly selecting a reference group when conducting an empirical Bayes analysis (p. 551-552). It warns that not doing so could drastically affect the validity of the evaluation. In addition, it discusses how important properly accounting for traffic volume changes can be and how the empirical Bayes method does so (p. 553-554).

Persaud, B. and C. Lyon. Empirical Bayes Before–After Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions. Accident Analysis and Prevention, Vol. 39, Issue 3, 546–555, 2007.

How Many Accidents are Needed to Show a Difference

This paper discusses the basics of testing for statistical significance, sample size estimates, confidence intervals, and level of significance with respect to before-after studies.

Hauer, E. How many accidents are needed to show a difference? Accident Analysis and Prevention, Vol. 40, Issue 4, 1634-1635, 2008.

Integrated Safety Management Process

This document briefly describes the empirical Bayes before-after method and states that it accounts for regression-to-the-mean bias, traffic volume changes, and time trend changes (p. D-43). It also presents an example computation with step-by-step instructions (p. C-11). In addition, it provides a brief case study of the method used in conjunction with crash surrogates (p. D-44).

Bahar, G., M. Masliah, C. Mollett, and B. Persaud. Integrated Safety Management Process. NCHRP Report 501, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2003. Available online at: http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_501.pdf <Accessed May 17, 2010>

Comparison Group

Observational Before-After Studies in Road Safety

This text provides a detailed description of the before-after with comparison group method. It thoroughly discusses the basic concept and study design considerations. It also identifies how the study design addresses the issues of traffic volume changes and time trends, but fails to account for the regression-to-the-mean bias. Relevant equations and examples are provided throughout. The text then provides a procedure for estimating sample sizes for a before-after with comparison group study.

Hauer, E. Observational Before–After Studies in Road Safety. Pergamon Press, Oxford, UK, 1997.

Statistical Methods in Highway Safety Analysis, A Synthesis of Highway Practice

This document briefly discusses the biases addressed by the before-after with comparison group method (p. 59). It briefly states that this method addresses biases related to traffic volume changes and time trends but does not address the regression-to-the-mean bias (p. 59). It also discusses the need to select a proper comparison group when using the method (p. 60).

Persaud, B. Statistical Methods in Highway Safety Analysis, A Synthesis of Highway Practice. NCHRP Synthesis 295, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2001. Available online at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_295.pdf <Accessed May 17, 2010>.

Safety Effectiveness of Intersection Left- and Right-Turn Lanes

This document compares and contrasts the before-after with comparison group method with the before-after with yoked comparison method and the empirical Bayes method (p. 134-135). This comparative analysis was performed by applying each of the three methods in an evaluation of adding left-turn and right-turn lanes at intersections. The document states that the before-after with comparison group method accounts for traffic volume changes but fails to account for the regression-to-the-mean bias (p. 134-135). It also states that suitable comparison group sites were selected by identifying sites that were as similar as possible to the treatment sites with respect to geography, intersection configuration, traffic control, geometric design, and traffic volume (p. 133).

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. Federal Highway Administration, FHWA-RD-02-089, Washington, DC, July, 2002.

Safety Evaluation of Rolled-In Continuous Shoulder Rumble Strips Installed on Freeways

This document presents a brief summary of the before-after with comparison group method. It also applies this method in an evaluation of continuous shoulder rumble strips.

Griffith, M.S. Safety Evaluation of Rolled-In Continuous Shoulder Rumble Strips Installed on Freeways Journal of the Institute of Transportation Engineers, Vol. 70, Issue 6, 37-42, 2000.

Integrated Safety Management Process

This document presents a brief description of the before-after with comparison group method (p. D-43). It also briefly discusses the need to select a suitable comparison group (p. D-43).

Bahar, G., M. Masliah, C. Mollett, and B. Persaud. Integrated Safety Management Process. NCHRP Report 501, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2003. Available online at: http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_501.pdf <Accessed May 17, 2010>

Yoked Comparison

Safety Effectiveness of Intersection Left- and Right-Turn Lanes

This document compares and contrasts the before-after with yoked comparison method with the before-after with comparison group method and the empirical Bayes method (p. 134-135). This comparative analysis was performed by applying each of the three methods in an evaluation of adding left-turn and right-turn lanes at intersections. It gives a step-by-step

overview of the before-after with yoked comparison method complete with equations (p. 134-135). It states that the before-after with yoked comparison method accounts for traffic volume changes but fails to account for the regression-to-the-mean bias (p. 134-135). It also states that a suitable comparison group was selected by identifying sites that were as similar as possible to the treatment sites with respect to geography, intersection configuration, traffic control, geometric design, and traffic volume (p. 133).

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. Federal Highway Administration, FHWA-RD-02-089, Washington, DC, July, 2002.

Safety Evaluation of Rolled-In Continuous Shoulder Rumble Strips Installed on Freeways

This document presents a brief overview of the before-after with yoked comparison method. It demonstrates this method in an evaluation of continuous shoulder rumble strips. It also states the need to select a suitable comparison group.

Griffith, M.S. Safety Evaluation of Rolled-In Continuous Shoulder Rumble Strips Installed on Freeways Journal of the Institute of Transportation Engineers, Vol. 70, Issue 6, 37-42, 2000.

Integrated Safety Management Process

This document gives a brief description of the before-after with yoked comparison and states the need to select a suitable comparison group (p. D-43).

Bahar, G., M. Masliah, C. Mollett, and B. Persaud. Integrated Safety Management Process. NCHRP Report 501, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2003. Available online at: http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_501.pdf <Accessed May 17, 2010>

Full Bayes

Integrated Safety Management Process

This report provides a brief introduction to the concept of Bayesian modeling, including the combination of prior knowledge and local conditions to estimate a CMF. The report briefly discusses two approaches for obtaining prior estimates of CMFs for use in Bayesian analysis.

Bahar, G., M. Masliah, C. Mollett, and B. Persaud. Integrated Safety Management Process. NCHRP Report 501, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2003. Available online at: http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_501.pdf <Accessed May 17, 2010>

Comparison of Empirical Bayes and Full Bayes Approaches for Before-After Road Safety Evaluations

This document begins with an overview of the full Bayes method (p. 39). It then describes an application of this method to study the effects of converting four-lane road segment to two-lane segments with two-way-left-turn lanes. The document also briefly mentions how a reference group is selected for a full Bayes analysis (p. 39).

Persaud, B., B. Lan, C. Lyon and R. Bhim. Comparison of Empirical Bayes and Full Bayes Approaches for Beforeand-After Road Safety Evaluations. Accident Analysis and Prevention, Vol. 42, Issue 1, 38–43, 2010.

Bayes and Empirical Bayes Methods for Data Analysis

While the emphasis of this text is on biomedical applications, it does provide a detailed discussion of Bayesian study designs, analysis, and implementation issues. The text incorporates several examples throughout and provides an entire chapter on case studies. Statistical codes are provided for use with common software such as WinBUGS and R.

Carlin, B.P. and T.A. Louis. Bayes and Empirical Bayes Methods for Data Analysis, Third Edition. Chapman and Hall/CRC, New York, NY, 2008.

Bayesian Data Analysis

This text provides a detailed discussion of Bayesian study design, model building, inference, and model checking. The text also discusses challenges to the Bayesian analysis and incorporates several examples throughout. An appendix is provided for computational details.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. Bayesian Data Analysis, Second Edition. Chapman and Hall/CRC, New York, NY, 2003.

Iowa's Experience with Road Diet Measures: Use of Bayesian Approach to Assess Impacts on Crash Frequencies and Crash Rates

This document provides a brief description of how full Bayes modeling was used to estimate model parameters for a time-series analysis of the conversion of four-lane, undivided roads to two-lane roads with a two-way-left-turn-lane.

Pawlovich, M.D., W. Li, A. Carriquiry, and T.M. Welch. Iowa's Experience with Road Diet Measures: Use of Bayesian Approach to Assess Impacts on Crash Frequencies and Crash Rates. Transportation Research Record: Journal of the Transportation Research Board, No. 1953, Transportation Research Board of the National Academies, Washington, DC, 2006.

From Empirical Bayes to Full Bayes: Methods for Analyzing Traffic Safety Data

This document provides a detailed description of the Full Bayes method and explains how it is distinct from the empirical Bayes method. It also presents a step-by-step demonstration of a Full Bayes analysis in which the crash rates of intersections with traffic control signals are compared to the crash rates of intersections without traffic control signals.

Carriquiry, A. and M.D. Pawlovich. From Empirical Bayes to Full Bayes: Methods for Analyzing Traffic Safety Data. White Paper, October, 2005. Available online at: http://www.iowadot.gov/crashanalysis/pdfs/eb_fb_comparison_whitepaper_october2004.pdf <Accessed July 12, 2011>

Empirical Bayes Before–After Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions

This document briefly discusses the advantages and disadvantages of using the full Bayes method instead of the empirical Bayes method (p. 554).

Persaud, B. and C. Lyon. Empirical Bayes Before–After Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions. Accident Analysis and Prevention, Vol. 39, Issue 3, 546–555, 2007.

Full Bayes Approach to Before-and-After Safety Evaluation with Matched Comparisons: Case study of stop-sign in-fill program

This paper demonstrates the use of the full bayes approach to evaluate the effectiveness of the stop-sign in-fill program of the Canadian Insurance Corporation of British Columbia.

El-Basyouny, K. and T. Sayed, Full Bayes Approach to Before-and-After Safety Evaluation with Matched Comparisons: Case study of stop-sign in-fill program, Transportation Research Record 2148, pp. 1-8, 2010.

Validation of a Full Bayes Methodology for Observational Before-After Road Safety Studies and Application to Evaluation of Rural Signal Conversions

This document provides a description of the full Bayes before-after method. This document also describes an application of the full Bayes before-after method to analyze the safety effectiveness of installing traffic signals at rural, unsignalized intersections in California. The document also discusses the regression-to-the-mean bias and its relationship to the full Bayes before-after method. It describes how this method was applied on a hypothetical dataset in which it was known a priori that there were no statistically significant changes in crashes and that the changes that were observed were due entirely to the regression-to-the-mean phenomenon. Confirming its validity, the results of the full Bayes method indicated no significant changes in crashes.

Lan, B., B. Persaud, and C. Lyon. Validation of a Full Bayes Methodology for Observational Before-After Road Safety Studies and Application to Evaluation of Rural Signal Conversions. Accident Analysis and Prevention, Vol. 41, Issue 3, Pages 574-580, 2009.

Use of Propensity Score Matching Method and Hybrid Bayesian Method to Estimate Crash Modification Factors of Signal Installation

This document describes the application of a full Bayes analysis using Markov Chain Monte Carlo techniques to estimate CMFs for the signalization of intersections in Minnesota.

Aul, N. and G.A. Davis. Use of Propensity Score Matching Method and Hybrid Bayesian Method to Estimate Crash Modification Factors of Signal Installation. Transportation Research Record: Journal of the Transportation Research Board, No. 1950, Transportation Research Board of the National Academies, Washington, DC, 2006.

Time Series Study Designs

Weather Effects on Daily Traffic Accidents and Fatalities: Time Series Count Data Approach

This study describes a time series approach to investigate the impact that weather has on crashes. The study gives an interpretation of the accuracy of the results of the analysis (p. 6-7). It also identifies traffic volume changes as a possible cause for some of the observed results (p. 7). This implies that the analysis does not properly account for traffic volume changes.

Karlaftis, M.G. and G. Yannis. Weather Effects on Daily Traffic Accidents and Fatalities: Time Series Count Data Approach. Transportation Research Board 89th Annual Meeting Compendium of Papers, Paper #10-0325, Washington, DC, 2010.

Evaluation of Miami-Dade Pedestrian Safety Demonstration Project

This paper used the multivariate ARIMA model to evaluate the safety impact of engineering, enforcement, and education treatments to reduce pedestrian crashes in different locations in the Miami-Dade area in Florida.

Zegeer, C. et al., (2008), Evaluation of Miami-Dade Pedestrian Safety Demonstration Project, Transportation Research Record 2073, pp. 1-10.

Application of Innovative Time Series Methodology to Relationship between Retroreflectivity of Pavement Markings and Crashes

This document presents a case study of a time-series analysis in which CMFs are developed for different levels of pavement retroreflectivity. The study attempts to avoid the effects of potentially confounding factors by accounting for them in the analysis. For example, CMFs are developed for different facility types and different seasons (p. 123).

Masliah, M., G. Bahar, and E. Hauer. Application of Innovative Time Series Methodology to Relationship between Retroreflectivity of Pavement Markings and Crashes. Transportation Research Record: Journal of the Transportation Research Board, No. 2019, Transportation Research Board of the National Academies, Washington, DC, 2007.

Safety Impact of the 65-mph Speed Limit: A Time Series Analysis

This study employs a time series analysis to examine the impact of the increase in speed limit in 1987 to 65 mph on rural Interstate highways in Illinois. Specifically, autoregressive integrated moving average (ARIMA) time series intervention analyses were used to examine changes in

85th-percentile speeds and speed variances as well as the changes in accident frequency, accident rates, and proportion of car-truck accidents to all accidents.

Pfefer, R.C., Stenzel,W.W., and Lee, B.D.. Safety Impact of the 65-mph Speed Limit: A Time Series Analysis (Abridgment). Transportation Research Record: Journal of the Transportation Research Board, No. 1318, Transportation Research Board of the National Academies, Washington, DC, 1991.

Cross-Sectional Study Designs

The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives

This document provides an in-depth discussion of the cross-sectional approach to developing CMFs. It discusses the various cross-sectional models that are available for analyzing crash frequency data. It also describes how correlation (e.g., temporal or spatial) in the data, if unaccounted for, can bias the results of the analysis (p. 6-8). In addition, the study briefly discusses the importance of functional form (p. 10-11).

Lord, D. and F. Mannering. The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. Transportation Research Part A: Policy and Practice, Vol. 44, Issue 5, 291-305, 2010. Also available online at: https://ceprofs.civil.tamu.edu/dlord/papers/lord-mannering_review.pdf

The Relationship among Highway Geometrics, Traffic-Related Elements and Motor-Vehicle Accident Frequencies

This document gives a detailed description of the cross-sectional method of developing CMFs. It describes the evolution of this method from its origins in linear regression to its present form which commonly uses the maximum likelihood method assuming a Negative Binomial distribution. It also provides a cross-sectional analysis of arterial roadways in Washington state and presents the resulting CMFs. The document briefly discusses how the potential bias of site history was a concern and how it was addressed. More specifically, it states that roadway segments that underwent construction changes or improvements during the study period were excluded (p. 402). The document also provides an interpretation of the accuracy of the results for each CMF produced in the cross-sectional analysis (pp. 404-410).

Milton, J. and F. Mannering. The Relationship among Highway Geometrics, Traffic-Related Elements and Motor-Vehicle Accident Frequencies. Transportation, Vol. 25, 395–413, 1998.

Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies

This document provides an overview of cross-sectional modeling of crash data accompanied by equations describing the Poisson and Negative Binomial distributions. It also presents a case

study in which a crash prediction model is constructed for a steep section of Interstate 90 in Washington State with sharp curves and severe weather. The study describes the rationale for selecting the candidate variables that were tested and which eventually defined the functional form of the crash prediction model (p. 373). It also briefly states that various interactions were tested, found to be statistically insignificant factors, and excluded from the model (pp. 379-380). This study provides an interpretation of the accuracy of each parameter in the crash prediction model (pp. 376-379, 382-385).

Shankar, V., F. Mannering, and W. Barfield. Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies. Accident Analysis and Prevention, Vol. 27, Issue 3, 371-389, 1995.

Statistical Methods in Highway Safety Analysis, A Synthesis of Highway Practice

This document gives an overview of the cross-sectional approach to developing CMFs (p. 5). It also presents a simple example of a cross-sectional model form (p. 8). It briefly discusses the issue of missing or correlated variables and how they pose a challenge to the integrity of the cross-sectional analyses (p. 9).

Persaud, B. Statistical Methods in Highway Safety Analysis, A Synthesis of Highway Practice. NCHRP Synthesis 295, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2001. Available online at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_295.pdf <Accessed May 17, 2010>.

Statistical Road Safety Modeling

This study first gives an overview of the cross-sectional method. It then presents a detailed discussion concerning how to select an appropriate model form for a cross-sectional analysis, including the need for additive forms in some cases in addition to the more conventional multiplicative form.

Hauer, E. Statistical Road Safety Modeling. Transportation Research Record: Journal of the Transportation Research Board, No. 1897, Transportation Research Board of the National Academies, Washington, DC, 2004.

Relationship of Flow, Speed, and Density of Urban Freeways to Functional Form of a Safety Performance Function

This paper raises the importance of selecting an appropriate model form for a safety performance function. Data from urban freeways in California and Colorado were used to estimate SPFs with sigmoid and the more common exponential functional forms. SPFs developed through the sigmoid functional form generally showed a better fit (based on CURE plots) indicating the need for care and thought in selecting an appropriate model form instead of relying on the default exponential form.

Kononov, J., Lyon, C., and Allery, B.K., Relationship of Flow, Speed, and Density of Urban Freeways to Functional Form of a Safety Performance Function, Transportation Research Record:

Journal of the Transportation Research Board, No. 2236 (pp. 11-19), Transportation Research Board of the National Academies, Washington, DC, 2011.

Procedure for Developing Accident Modification Factors from Cross-Sectional Data

This study presents a detailed procedure for estimating CMFs using cross-sectional analysis. It also applies the procedure in an example to develop a crash modification function for the radius of a horizontal curve. The study assesses the accuracy of the developed crash modification function by comparing it to ones found in a previous study (p. 16).

Bonneson, J.A. and M.P. Pratt, Procedure for Developing Accident Modification Factors from Cross-Sectional Data. Transportation Research Record: Journal of the Transportation Research Board, No. 2083, Transportation Research Board of the National Academies, Washington, DC, 2008.

Cause, Effect, and Regression in Road Safety: A Case Study

This document presents a detailed case study of a cross-sectional analysis of converting crossbucks to flashers at rail-highway grade crossings. It gives a lengthy interpretation of the accuracy of CMFs developed from regression models.

Hauer, E. Cause, Effect, and Regression in Road Safety: A Case Study. Accident Analysis and Prevention, Vol. 42, Issue 4, 1128-1135, 2010.

Statistical and Econometric Methods for Transportation Data Analysis

This text provides a detailed discussion of several statistical methods related to cross-sectional data. Specifically, it provides relevant equations and discusses potential biases associated with each method. Several examples and case studies are provided throughout to illustrate various methods.

Washington, S., M. Karlaftis, and F. Mannering. Statistical and Econometric Methods for Transportation Data Analysis, Second Edition. Chapman and Hall/CRC, New York, NY, 2010.

Scientific Approaches to Transportation Research Volumes 1 and 2

This document provides a brief overview of cross-sectional modeling. More specifically, it describes Poisson and Negative Binomial models. It also describes model validation, including internal and external model validation and methods to accomplish these.

Washington, S., J. Leonard, D.G. Manning, C. Roberts, B. Williams, A.R. Bacchus, A. Devanhalli, J. Ogle, and D. Melcher. Scientific Approaches to Transportation Research Volumes 1 and 2. NCHRP Online Report 20-45, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2001. Available online at:
http://onlinepubs.trb.org/Onlinepubs/nchrp/cd-22/start.htm

Validation of Accident Models for Intersections

This report describes the results of validation and calibration of crash models for rural intersections. Both the validation and calibration activities were conducted in pursuit of making improvements to an existing set of statistical models for predicting crashes at two- and four-lane rural road intersections with the primary intent to be used in the Highway Safety Manual and the Interactive Highway Safety Design Module (IHSDM). The validation and calibration included the functional form and parameter estimates as well as crash modification factors derived from the model. The transferability of the models in time and geographic location was also considered.

Washington S., B. Persaud, C. Lyon, and J. Oh. Validation of Accident Models for Intersections. Federal Highway Administration, FHWA-RD-03-037, Washington, DC, 2005.

Case-Control Study Designs

Speed and Crash Risk: An Opinion

This paper discusses the use of the case-control method in highway safety and identifies several potential strengths and weaknesses. Specifically, the paper discusses the potential issue of confounding factors and how confounders can have a significant impact on the results if not accounted for properly. The paper also provides a detailed discussion of other issues, such as that when factors associated with the cases are measured with a different precision that those factors associated with the controls.

Hauer, E. Speed and Crash Risk: An Opinion. Public Policy Department, Royal Automobile Club of Victoria, Report No. 04/02, May 2004.

Estimation of the Safety Effectiveness of Lane and Shoulder Width: Case-Control Approach

This paper provides an overview of the case-control design and its potential application in highway safety to estimate CMFs. It discusses confounding factors and how to account for potential confounders through matching. The paper also presents appropriate models to estimate the odds ratio (CMF) for a case-control design. An example is then presented using data for two-lane undivided roads in Pennsylvania to develop CMFs for lane width and shoulder width.

Gross, F. and P.P. Jovanis. Estimation of the Safety Effectiveness of Lane and Shoulder Width: Case-Control Approach. American Society of Civil Engineers, Journal of Transportation Engineering, Vol. 133, Issue 6, 2007.

A Dissertation in Civil Engineering: Alternative Methods for Estimating Safety Effectiveness on Rural, Two-Lane Highways: Case-Control and Cohort Methods

This document provides a detailed discussion of the case control method and discusses its potential to estimate CMFs. It reviews the strengths and weaknesses of the case-control method and discusses other related elements, including sample size calculations and modeling techniques. Confounding factors are identified as a potential issue in the analysis of cross-sectional data and several potential confounding factors are investigated. Empirical examples are provided using data from rural, two-lane, undivided roads in Pennsylvania and Washington. Matching is applied in the case-control design to isolate the effects of lane and shoulder width by accounting for confounding variables, such as ADT, speed limit, and segment length. Conditional logistic regression is used to account for the matching procedure and estimate CMFs.

Gross, F. A Dissertation in Civil Engineering: Alternative Methods for Estimating Safety Effectiveness on Rural, Two-Lane Highways: Case-Control and Cohort Methods. The Pennsylvania State University, December 2006.

Estimation of Safety Effectiveness of Changes in Shoulder Width using Case-Control and Cohort Methods

This paper explores the use of the case-control design to estimate CMFs for shoulder width. It provides an overview of the case-control method and discusses the associated strengths and weaknesses. It also discusses the potential issue of confounding factors and identifies two methods to account for confounding factors in a case-control design (matching and inclusion of covariates in a regression model). Conditional logistic regression models are then developed to estimate CMFs for shoulder width, using data for rural, two-lane, undivided roads in Pennsylvania.

Gross, F. and P.P. Jovanis. Estimation of Safety Effectiveness of Changes in Shoulder Width using Case-Control and Cohort Methods. Transportation Research Record: Journal of the Transportation Research Board, No. 2019, Transportation Research Board of the National Academies, Washington, DC, 2008.

Epidemiology: Study Design and Data Analysis

This text provides a detailed discussion of the case-control method, including basic design concepts, strengths and weaknesses, selection of cases and controls, methods of analysis, sample size calculations, and several exercises to apply the case-control method.

Woodward, M. Epidemiology: Study Design and Data Analysis, Second Edition. Chapman and Hall/CRC, New York, NY, 2005.

Integrated Safety Management Process

This report provides a brief overview of the case-control design and its relevance in highway safety. It briefly discusses a few strengths and limitations compared to other study designs.

Several example applications of the method are presented, but no computations or equations are provided for the method.

Bahar, G., M. Masliah, C. Mollett, and B. Persaud. Integrated Safety Management Process. NCHRP Report 501, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2003. Available online at: http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_501.pdf <Accessed May 17, 2010>

Cohort Study Designs

A Dissertation in Civil Engineering: Alternative Methods for Estimating Safety Effectiveness on Rural, Two-Lane Highways: Case-Control and Cohort Methods

This document provides a detailed discussion of the cohort method and discusses its potential to estimate CMFs. It reviews the strengths and weaknesses of the cohort method and discusses other related elements, including sample size calculations and modeling techniques. Confounding factors are identified as a potential issue in the analysis of cross-sectional data and several potential confounding factors are investigated. Empirical examples are provided using data from rural, two-lane, undivided roads in Pennsylvania and Washington. Confounding is addressed by including potential confounders as covariates in the model. Survival and count models are applied with the cohort design to estimate CMFs.

Gross, F. A Dissertation in Civil Engineering: Alternative Methods for Estimating Safety Effectiveness on Rural, Two-Lane Highways: Case-Control and Cohort Methods. The Pennsylvania State University, December 2006.

Estimation of Safety Effectiveness of Changes in Shoulder Width using Case-Control and Cohort Methods

This paper explores the use of the cohort design to estimate CMFs for shoulder width. It provides an overview of the cohort method and discusses the associated strengths and weaknesses. Cox proportional hazard models are then developed to estimate CMFs for shoulder width, using data for rural, two-lane, undivided roads in Pennsylvania.

Gross, F. and P.P. Jovanis. Estimation of Safety Effectiveness of Changes in Shoulder Width using Case-Control and Cohort Methods. Transportation Research Record: Journal of the Transportation Research Board, No. 2019, Transportation Research Board of the National Academies, Washington, DC, 2008.

Epidemiology: Study Design and Data Analysis

This text provides a detailed discussion of the cohort method, including basic design considerations, strengths and weaknesses, analytical considerations, methods of analysis, sample size calculations, and several exercises to apply the cohort method.

Woodward, M. Epidemiology: Study Design and Data Analysis, Second Edition. Chapman and Hall/CRC, New York, NY, 2005.

Integrated Safety Management Process

This report provides a brief overview of the cohort design and its relevance in highway safety. It briefly discusses a few strengths and limitations compared to other study designs. Several example applications of the method are presented, but no computations or equations are provided for the method.

Bahar, G., M. Masliah, C. Mollett, and B. Persaud. Integrated Safety Management Process. NCHRP Report 501, Transportation Research Board, National Cooperative Highway Research Program, Washington, DC, 2003. Available online at: http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_501.pdf <Accessed May 17, 2010>

APPENDIX B. METHOD CORRECTION FACTOR TABLES FROM HIGHWAY SAFETY MANUAL INCLUSION PROCESS

TABLE 2 Method Correction Factors for Before–After and Meta-Analysis Studies

Key Study Characteristic	Method Correction Factor
All potential sources of bias were properly accounted for Uses crash frequencies	1.2
Accounts for regression to the mean bias Uses crash frequencies	1.8
Regression to the mean may not be accounted for but considered to be minor, if any Uses crash frequencies or crash rates	2.2
Regression to the mean not accounted for and considered to be likely Uses crash rates	3
Severe lack of information published regarding study data and findings	5

NOTE: This table applies to empirical Bayes, simple before–after, before–after with likelihood functions, before–after with comparison group, expert panels, and meta-analysis.

TABLE 3 Method Correction Factors for Nonregression Cross-Section Studies

Key Study Characteristic	Method Correction Factor
All potential confounding factors have been accounted for by matching sites	1.2
Most potential confounding factors have been accounted for by matching sites	2
Traffic volume is the only confounding factor accounted for in the study	3
No confounding factors accounted for in the study	5
Severe lack of information published regarding study data and findings	7

TABLE 4 Method Correction Factors for Regression Cross-Section Studies

Key Study Characteristic	Method Correction Factor
All potential confounding factors have been accounted for by variables of the regression in an appropriate functional form	1.2
Most potential confounding factors have been accounted for by variables of the regression in an appropriate functional form	1.5
Several important confounding factors were accounted for, and functional form is conventional	2
Few variables used and functional form is questionable	3
Severe lack of information published regarding study data study and findings	5

Citation: Bahar, G. (2010), Methodology for the Development and Inclusion of Crash Modification Factors in the First Edition of the Highway Safety Manual, *Transportation Research Circular* E-C142, Transportation Research Board.

NOTE: The NCHRP 17-27 report indicated that the MCF values were developed internally by the experts on the project team.

APPENDIX C. SUMMARY OF CMF PROTOCOLS DOCUMENTATION REQUIREMENTS

In summary form, here are the documentation requirements detailed in the CMF Protocols:

General Documentation (knowing where the CMF can be appropriately applied)

- CMF and Countermeasure
  - Countermeasure Name and Description
  - Crash Modification Factor or Function
  - Measures of Precision for the CMF
  - Prior Conditions
- Site Characteristics
  - Roadway Class
  - Road Division Type
  - State
  - Municipality
  - Area Type
  - Number of Through Lanes
  - Speed Limit
  - Traffic Volume Range
  - Traffic Control (if applicable)
  - Intersection Type
  - Intersection Geometry
  - Other Relevant Details
- Crash Characteristics
  - Crash Type
  - Crash Severity
  - Time of Day
- Study Details
  - Years of Data
  - Type of Methodology
  - Site Selection Criteria
  - Sample Size Used (Crashes)
  - Sample Size Used (Sites)

Biases Documentation (knowing how potential biases were addressed)

- Potential Biases for Cross-sectional Studies
    - Control of confounders
    - Unobserved heterogeneity and omitted variable bias
    - Accounting for state-to-state differences if using multiple states
    - Selection of appropriate functional form
    - Correlation or collinearity among the independent variables
    - Overfitting of prediction models
    - Low sample mean and small sample size
    - Bias due to aggregation, averaging, or incompleteness in data
    - Temporal and spatial correlation
    - Endogenous independent variables
    - Misspecification of structure of systematic variation and residual terms
    - Correlation between crash types and injury severities
- Potential Biases for Before-After Studies
    - Regression-to-the-mean
    - Changes in traffic volumes
    - History trends
    - Other safety treatments
    - Changes in crash reporting
    - Accounting for state-to-state differences if using multiple states
    - Suitability of comparison or reference groups
    - Other biases for before-after studies from cross-sectional list