FULLY BAYESIAN INFEERENCE TECHNIQUES FOR TRAFFIC SAFETY TREATMENT BEFORE-AND-AFTER STUDY

by

Simon Chun-Yin Li

B.A.Sc., The University of British Columbia, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies
(Civil Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

August 2013

© Simon Chun-Yin Li, 2013
ABSTRACT

The importance of improving traffic safety is often understated, partially because it often takes a retrospective approach, garnering little public attention. Nonetheless, from both an economical and societal point of view, traffic safety presents severe and significant problems despite the sizeable benefits that advancements in transportation have brought to society. To further complicate the matter, the net results of most traffic safety interventions are not always straightforward or intuitive. This illustrates the need for sound engineering evaluation of traffic safety interventions that is grounded in statistical analysis. It should be noted that these engineering evaluations can be applied not only to location-specific safety treatments, but can also be used to test the effectiveness of traffic safety-targeted policies such as changes in BAC level or seat belt laws.

Previously, a prominent and effective methodology for conducting traffic safety intervention evaluations was known as the Empirical Bayes inference techniques. It was effective in accounting for a number of confounding factors, which threaten the validity of any claims made by simply looking at raw collision data. However, several key drawbacks have been identified, including difficulties to obtain the necessary amount of input data and the statistical discontinuity in the steps where the uncertainties around the input data are not entirely carried through to the final estimates. In theory, the recently-developed Full Bayes technique fully addresses the weaknesses of the Empirical Bayes method; however, there have been hesitations to adopt the methodology because of the increased level of complexity and the previous lack of adequate computational power. The purpose of this thesis to perform a thorough literature on methodologies for conducting traffic safety intervention models particularly with regards to Bayesian inference, devise a standardized methodology using the findings, apply the methodology on a real-world case study in Edmonton, Alberta, and summarize the results to demonstrate the strengths and the feasibility of the Full Bayes methodology. The results indicated that the treatment program was effective in reducing right-turn collisions by 39%. A standardized practical guideline was also developed using the literature review and the results and includes various provisions for flexibility and alterations.
PREFACE

This dissertation is based on the methodology developed by El-Basyouny and Sayed (2009). The data used in the research was provided by the Office of Traffic Safety under the City of Edmonton. None of the text of the dissertation is taken directly from previously published or collaborative articles. Chapters 4 and 5 included the collaboration of K. Basyouny and T. Sayed. The remaining chapters were prepared by myself.

Table 3-1, 3-2, 4-1, 5-3, and 5-4 are used with permission from applicable sources. Portions of the introductory text are used in a subsequent publication titled “Fully Bayesian Inference Technique on the Evaluation of Traffic Improvements Targeting Right-Turn Collisions in the City of Edmonton” written by T. Sayed, K. El-Basyouny and I. The paper has been accepted and is pending publication in the Transportation Research Board.
# TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................... ii

PREFACE ................................................................................................................................................ iii

TABLE OF CONTENTS .............................................................................................................................. iv

LIST OF TABLES ......................................................................................................................................... vii

LIST OF FIGURES ...................................................................................................................................... viii

ACKNOWLEDGEMENTS .......................................................................................................................... ix

1. INTRODUCTION .................................................................................................................................. 1

   1.1. Overview of Road Traffic Safety Problem .................................................................................. 1

   1.2. Background of Road Safety Initiatives in Canada .................................................................... 2

   1.3. City of Edmonton Road Traffic Safety ....................................................................................... 3

   1.4. Project Description .................................................................................................................. 4

   1.5. Thesis Objectives .................................................................................................................... 6

   1.6. Thesis Structure ....................................................................................................................... 6

2. LITERATURE REVIEW ....................................................................................................................... 7

   2.1. Road Traffic Safety ................................................................................................................ 7

   2.2. Road Section Classification .................................................................................................... 8

   2.3. Overview of Road Safety Evaluation ....................................................................................... 9

      2.3.1. Before-And-After Method .............................................................................................. 9

      2.3.2. Cross-Sectional Method ................................................................................................. 9

      2.3.3. Fully Experimental Study Method .................................................................................. 10

   2.4. Overview of Before-And-After Studies .................................................................................. 10

   2.5. Overview of BA Study Methodologies ................................................................................... 15

      2.5.1. Four Main Categories of BA Studies ............................................................................. 16

      2.5.2. Study Period Length ..................................................................................................... 18

   2.6. Before-and-After Study with Comparison ............................................................................. 20

      2.6.1. Selection of Comparison Sites ....................................................................................... 21

   2.7. Empirical Bayes Methodology ............................................................................................... 21

      2.7.1. Computation of the Odds Ratio in the EB Method ......................................................... 22
2.7.2. Empirical Bayes Refinement ................................................................. 23
2.7.3. Safety Performance Functions: Overview and Model Structure .................. 24
2.7.4. Safety Performance Functions: Distribution Specification and Resulting Models 26
2.7.5. Safety Performance Function: Modeling Approach ..................................... 28
2.7.6. Issues with Empirical Bayes Method ..................................................... 29
2.8. Full Bayes Methodology ......................................................................... 30
2.8.1. Full Bayes Models .................................................................................. 31
2.8.2. Continuous Time Intervals ..................................................................... 34
2.8.3. Spatial Trends ......................................................................................... 35
2.9. Comparison of EB and FB Method ........................................................... 36
3. REVIEW OF DATA SOURCES .................................................................... 38
3.1. Dataset Scope ............................................................................................ 38
3.2. Accident History ....................................................................................... 39
3.3. Volume Data ............................................................................................... 40
3.4. Volume Forecasting .................................................................................... 42
3.5. Summary of Available Data ....................................................................... 43
4. PRELIMINARY STUDY ............................................................................... 45
4.1. Studies not Considered .............................................................................. 45
4.2. Before-And-After Study with Comparison Groups ...................................... 45
5. MODELS AND RESULTS ......................................................................... 47
5.1. Functional Form ......................................................................................... 47
5.2. Model Specifications .................................................................................. 48
5.3. Prior Distribution Specifications ................................................................. 51
5.4. WinBUGS ................................................................................................. 51
5.5. Markov Chain Convergence ...................................................................... 52
5.6. Model Results ........................................................................................... 54
5.7. Model Comparison .................................................................................... 56
6. RESULTS DISCUSSION .......................................................................... 58
7. MONTHLY INTERVAL ANALYSIS ............................................................ 62
8. SUMMARY AND CONCLUSION ................................................................ 65
8.1. Objective 1: Demonstration of FB Methodology.................................................................65
8.2. Objective 2: Evaluation of Safety Initiative.................................................................68
8.3. Limitations of Further Research..................................................................................69
BIBLIOGRAPHY.........................................................................................................................71
APPENDIX A – USING PIVOT TABLES IN MICROSOFT EXCEL ........................................76
APPENDIX B – RUNNING WINBUGS SIMULATION..............................................................81
APPENDIX C – DEVIANCE INFORMATION CRITEREON......................................................86
APPENDIX D – DATA FILES....................................................................................................87
LIST OF TABLES

Table 2-1. Comparison of BA Methods ........................................................................................................... 18
Table 2-2. History of Crash Recording Practices in Several Traffic Jurisdictions in Australia .................. 20
Table 3-1. Treated Sites Detail ......................................................................................................................... 38
Table 3-2. Comparison Sites Detail ................................................................................................................ 39
Table 3-3. Collision Data Details .................................................................................................................... 39
Table 3-4. Traffic Data Detail ......................................................................................................................... 40
Table 3-5. Estimated 2011 AADT .................................................................................................................... 42
Table 3-6. Descriptive Statistics for Data ........................................................................................................ 44
Table 4-1. Before-And-After with Comparison Group .................................................................................... 46
Table 5-1. Common Functional Forms of Minor and Major Approach .......................................................... 48
Table 5-2. Collision Data Inclusion for Right-Turn Analysis .......................................................................... 50
Table 5-3. Summary of Results of Models ...................................................................................................... 55
Table 5-4. Summary of Results for Right-Turn Analysis ................................................................................ 55
Table 5-5. Summary of DIC ............................................................................................................................ 56
Table 6-1. Summary of Evaluation Results (MVPLNJ Model) ...................................................................... 58
Table 7-1. Summary of Monthly Analysis Contrasted with Yearly Analysis ................................................ 63
Table 8-1. Standardized Practical Guideline for the Full Bayes Inference Technique ................................... 68
Table 8-2. Summary of Program Effectiveness ............................................................................................... 69
LIST OF FIGURES

Figure 1-1. Road Traffic Safety History (Source: Edmonton) .................................................................................. 3
Figure 2-1. Traffic Accident Causes (Source: Rumar, 1985) .................................................................................. 7
Figure 2-2. Importance of Confounding Factors in BA Studies (Source: Elvik, 2002) ................................................. 14
Figure 2-3. Comparison of Controlled and Uncontrolled Estimates of Effects on 9 Measures of Safety Evaluated in Norway (Source: Elvik, 2002) ........................................................................................................... 15
Figure 3-1. Example of Pivot Table Output ........................................................................................................... 40
Figure 3-2. Sample Traffic Flow Diagram ............................................................................................................... 41
Figure 3-3. Example of Automatic Traffic Volume Detection Dataset (Source: City of Edmonton Open Data Catalogue) ........................................................................................................................................... 43
Figure 6-1. Change in Collisions under MVPLNJ Model) .......................................................................................... 59
ACKNOWLEDGEMENTS

I would like to thank you my program supervisor, Dr. Tarek Sayed, for his constant support and guidance throughout not only my thesis but my entire program. His emphasis on quality and refinement and his enthusiasm has aided my greatly as I wrote my thesis. His level of dedication and involvement in this thesis is as if it was his own.

Moreover, I would like to thank other professors who I have interacted with during my program, especially through taking other courses. I have consequently been exposed to large range of different topics, from technical areas such as reliability analysis to societal topics such as urban planning.

I would also like to thank my fellow colleagues and classmates, particularly those in the BITSAFS office. They have made my last two years extremely enjoyable and laughable.

Last but not least, I thank my parents and my siblings for their support throughout my entire life, such that I am able to achieve the things which I have achieved today.
1. INTRODUCTION

This thesis deals with the evaluation of the effectiveness of road safety improvements at a particular location using advanced statistical methodologies known as Fully Bayesian inference techniques. It abridges practice and research by employing recently developed methodologies previously reserved only for research purposes because of limitations, on a real world dataset to illustrate its advantages, feasibility, and also identify further possible areas for refinement. Concurrently, because of the various different versions of the technique, this thesis will also create a standardized approach for implementing the techniques to promote comparability and compatibility of results. Moreover, it is written in agreement with the City of Edmonton Office of Traffic Safety (OTS) as a fulfillment of a contract to perform a formal evaluation of a number of recent location-specific safety improvements.

1.1. Overview of Road Traffic Safety Problem

The introduction of motorized modes of transportation has unarguably brought about significant benefits to society through the realization of greater degrees of travel and transport of goods. Nonetheless, motorized transportation is not without its drawbacks, with one of the most prominent one being road traffic safety.

The road traffic safety problem is a global problem that transcends all levels of jurisdictions. From a societal point of view, road traffic collisions are rated as one of the top causes of deaths. A 2002 report by WHO ranked road traffic accidents as the 19th leading cause of deaths, preceded by diseases such as cardiovascular or cancer. This represents more than 2% of all deaths, amounting to 1.9 million deaths within that year, and more than 5000 deaths a day. By 2004, the number had decreased but still remained significant at 1.2 million deaths (WHO). Due to the explosion in motorization in developing nations such as China and India, this figure is expected to grow by at least 65% over the next 20 years if no interventions are made.

Despite falling rates of road traffic accidents, accident costs still represent a large portion of GDP, with the United Arab Emirates reporting 2.9%, Russia reporting 2.5%, and the United States of America reporting 2.3% (WHO, 2004). The types of costs that incur in a road safety accident are many: internal costs which
include damage repair, emergency response, administration, medical, and external costs which include pain and loss of family members, loss of productivity, and congestion caused by collisions.

1.2. Background of Road Safety Initiatives in Canada

The case study used in this thesis to develop the methodology is comprised of a number of traffic safety improvements performed in the Edmonton, AB, Canada over the past few years. A background of road safety initiatives within Canada and specific to Edmonton will give an example of a typical road safety program or approach employed by traffic safety agency, which is important to the developmental stage of the methodology.

On a national level, the Canadian Council of Motor Transportation Administrators is responsible for developing a Road Safety Vision plan. The most current vision plan is 2010 (CCTMA), and cites the following goals:

- Raise public awareness of road safety issues
- Improve communication, cooperation, and collaboration among road safety agencies
- Enhance enforcement measures
- Improve road safety data quality and collection

The 2010 Vision also has exact number targets for various approaches of improving road traffic safety, with a few of them being:

- **Intersections**: A 20% decrease in the number of road users killed or seriously injured in intersection-related crashes.
- **Vulnerable Road Users**: A 30% decrease in the number of fatally or seriously injured vulnerable road users (pedestrians, motorcyclists and cyclists).

The 2015 Vision plan is currently in the works, with some of the key changes being the flexibility to allow regional and municipal agencies to set their own targets and the shift from measuring the road safety traffic problem with total number of collisions to collision rates (either with respect to population or number of vehicles).
1.3. City of Edmonton Road Traffic Safety

Though road traffic safety affects all levels of government, issues pertaining to road traffic safety have mostly remained within the provincial and municipal levels. The City of Edmonton Office of Traffic Safety (OTS) has placed itself in an enviable position as one of the first municipal traffic safety offices across Canada.

In 2005 it developed a document known as Traffic Safety strategy for the City of Edmonton 2005-2010 (The City of Edmonton), which outlines a number of strategies and approaches for the road traffic safety problem as well as aligning its goals with that of the national Road Safety Vision 2010.

The plan highlights some of the difficulties with road safety and gives a history of the scope of the problem. Figure 1-1 below is a chart displaying the history of total collisions and collision rates for the city of Edmonton for 2005 to 2010.

![Graph showing road traffic safety history in Edmonton](image)

**Figure 1-1. Road Traffic Safety History (Source: Edmonton)**

Within the document, the city’s aim was for a 30% decrease in collisions per 1000 population, from a 31.7 average between the years 2001 to 2004 to a 22.2 average for the projected years of 2006 to 2010. The four specific targets of the city of Edmonton within the strategy are:

1. 20% reduction in the number of intersection-related collisions
2. 95% seatbelt wearing rate
3. Reduce impaired driving
4. Speed related collisions

The strategy plan also lists 18 strategies to improving the road traffic safety within the city of Edmonton. Within these 18, several of them are in line with the objectives of this evaluation.

- **Strategy 5:** Identify and improve systems for data collection and management
- **Strategy 9:** The Office of Traffic Safety will examine and support the use of leading traffic safety technologies
- **Strategy 14:** Continue traffic engineering analysis and techniques to assess and prioritize safety initiatives

In 2011, the city of Edmonton was granted the status of “Smart City” under the IBM Smart City challenge, which awards grants in the form of senior consultant time to improve urban life with respect to six major themes (VanKeeken, 2011). Its recent improvements in traffic safety and setting exact targets were recognized as some of Edmonton’s major strengths within this challenge.

### 1.4. Project Description

To meet its target of 30% reduction in number of collisions, the Office of Traffic Safety (OTS) has initiated a number of engineering safety reviews in the past few years to improve safety levels. A number of different locations were reviewed and improved along the 97th Street, 91st Street, 137th Street, and 23rd Avenue corridors. Moreover, OTS also initiated an improvement project specifically targeting right turns. A number of potential right-turn improvements were implemented at various places as a result of this project.

In accordance with the aforementioned strategies and with OTS’ focus on evidence-based decision making, evaluating the effectiveness of past traffic safety improvements is imperative for a number of reasons.

- It allows for a subsequent economic evaluation of the effectiveness of the improvement
• Produce crash reduction factors (CMF) for estimating safety effects of future traffic implementations

Post-treatment evaluations are of great importance because the magnitude (size of reduction) or differing effects (on accident severity or accident types) are not necessarily intuitive and straightforward. The results of the evaluation will have a profound impact on future decision makings and resource allocations for traffic safety improvements in terms of the selection of locations as well as the selection of specific improvement. In this report, five locations within the City of Edmonton that were improved over the years 2008, 2009, and 2010 were selected to conduct an engineering evaluation of the traffic safety effects. The data provided by the City of Edmonton is used in conjunction with the expertise of University of British Columbia (UBC) in the area of traffic safety and improvement evaluations. The project will review the existing techniques used to conduct an engineering evaluation, from simple methods to the latest-state-of-the-art techniques, and conduct an analysis on the selected sites.

Up until most recently, the latest state-of-the-art technique for conducting a road evaluation was known as Empirical Bayesian (EB) inference. The methodology has been extensively used across many traffic authorities and has many advantages in terms of dealing with the statistical difficulties of traffic data. Within the past few years, a Fully Bayesian (FB) technique has been proposed in the literature, which theoretically addresses some of the shortcomings of the EB method (Aul and Davis, 2006; Pawlovich, Li and Carriquiry, 2006; Li, Carriquiry, Pawlovich and Welch, 2008; Lan, Persaud Lyon and Bhim, 2009; El-Basyouny and Sayed, 2012). There are many different versions within the FB method, and these will be discussed further in the literature review in Chapter 2.

Lastly, the need to adopt advanced statistical techniques for conducting evaluations pertains to the characteristics of collision data. Road collisions are rare and random events. The rarity of road collisions is fortunate, but also poses difficulties in terms of statistical strength of the results. Coupled with the randomness of collision counts, the reported number of collisions may not always be indicative of the true level of safety at a particular location. In theory, considering longer study periods will help to increase sample size and dampening out random fluctuations for a particular period. However, when a traffic improvement is proposed, it is difficult to justify delay in implementation for the sake of
subsequent evaluations. Moreover, Hauer (1995) has suggested that even a 5-year study period is not sufficient in resolving all the statistical uncertainties around collision data. Consequently, it is beneficial to explore advanced statistical methods that can deal with these difficulties and estimate the safety of locations with confidence.

1.5. Thesis Objectives

The thesis objectives are as follows:

- Demonstrate the application of several FB techniques on the evaluation of several locations that were improved under the recent city of Edmonton road safety program which targeted right-turn collisions
  - Through the demonstration, assess the advantages, feasibility, limitations, and overall maturity of the methods with respect to use in practice
- Using the results, make conclusions on the effectiveness of the city of Edmonton program with respect to reducing the number of collisions.

1.6. Thesis Structure

The structure of the thesis is as follows. The current chapter (Chapter 1) is an introduction to the thesis to introduce the topics, the context, and identify the thesis objectives. Chapter 2 is an extensive literature review on the existing methodologies for conducting a road traffic safety evaluation, highlighting the strengths, weaknesses, and data requirements for various techniques. Chapter 3 will discuss the available data provided by the city of Edmonton for the demonstration, followed by a discussion of the possibilities (with respect to techniques) and limitations of the dataset. Chapter 4 is a preliminary study using the BA with Comparison Group study method, which will be used for comparison purposes. Chapter 5 will conclude on the various FB models developed for the project, as well as its results and significance. A discussion of the results will follow in Chapter 6. Chapter 7 is a separate discussion on the experimentation of using monthly intervals for evaluation (as opposed to yearly) with comments on its appropriateness. Finally, the thesis will end with a summary and conclusion in Chapter 8.
2. LITERATURE REVIEW

The purpose of this chapter is to bring the reader up to date with regards to relevant topics, further set the context of the report, and explore the various existing techniques for conducting a safety improvement evaluation. It will begin with a discussion that acts as a precursor to conducting evaluations, such as the definition of road entities. It will then move on to an overview of conducting safety evaluations, particularly focusing on the necessary components of a sound evaluation. It will then focus on the various different methodologies that are available in the literature and in practice.

2.1. Road Traffic Safety

Traffic collisions occur for a number of different reasons: speeding, distractions, obstruction of view, etc. They are typically organized into three main groups of factors, which are Driver factors, Roadway factors, and Vehicle factors. In some cases, several factors can contribute to the occurrence of a collision. For example, a driver driving off the road at a sharp horizontal curve could simultaneously be attributed to the inattention on the driver’s part and the inadequacy of the road design to accommodate off-road vehicles. Rumar (1985) studied the causes of road traffic accidents in American and British Roadways. He concluded that, including overlaps, driver, road, and vehicle factors are involved in 93%, 34%, and 12% of collisions, respectively. The diagram below summarizes the proportions found from the study.

![Figure 2-1. Traffic Accident Causes (Source: Rumar, 1985)](image)

The results suggest that the most effective approach to addressing road traffic collisions is to focus on driver related initiatives. However, targeting driver behavior is difficult because the benefits are often temporary, and because most drivers already believe they have better than average driving skills. McCormick et al. (1986) found that 80% of drivers consider themselves as having above-average driving
abilities. From a road safety jurisdiction point of view, the most effective approach for improving safety is to focus on roadway contributing factors. This can either be from a proactive approach that implements safety features in the design of new roads, or from a reactive approach that looks to mitigate and retrofit existing roads. In light of limited funds, traffic authorities will prioritize funding to improve sites that have been identified to be problematic, which may be through ranking of hazardous hot spots as within the Black Spot Methodology (BSM) or through network screenings for over representations of a particular crash type.

2.2. Road Section Classification

When measuring, improving, or evaluating the safety of a road system, it is necessary to maintain a consistent definition and database of road systems. Within road safety, two types of “sites” exist: these are intersections and a specific section of a road. This report mainly deals with evaluations for intersections, but the principles and discussion of techniques applies to sites of both types.

Intersections are typically classified as individual sites. In most cases, the classification is quite straightforward. However, in some cases, care needs to be taken in atypical intersections where multiple streams of traffic may warrant separation into multiple separate intersections. An example of this is a highway intersection with on/off ramps; if sites are defined simply at the intersection of two main roads, then the entire highway-road intersection will be considered a single intersection. However, a typical highway intersection with on/off ramps may have multiple streams of traffic at any particular time. In such cases, it may be useful to define two separate intersections to define the on/off ramps on separate sides of the highway.

For road sections, two main approaches are the clearly defined population and the sliding window method (Sorenson, 2007). Proponents of using clearly defined road elements argue that it is less resource demanding, easier to comprehend, and that subsequent analysis and treatment are more practical (Hauer et al., 2002; Andersen and Sorensen, 2004; Pedersen and Sorensen, 2007). Unfortunately, the size of the elements is a matter of debate and also viewed as a limiting shortcoming of this approach. Details are lost if the road sections are made to be too large; extensive separation into small sections will lead to difficulty in traffic collision reporting as well as possible separations of potentially hazardous sites. The sliding
window approach uses dynamic division of the road network to consider all possibilities when conducting safety reviews for identification of hazardous locations. In practice, the simplicity and robustness of the clearly defined road element approach makes it the more accepted and recommended methodology (Hauer et al., 2002).

2.3. **Overview of Road Safety Evaluation**

As mentioned previously, one of the goals of conducting road safety evaluations is to build a database of the effect of certain countermeasures so that they can be used to predict the benefits of future implementations. These are formerly known as crash reduction factors (CRF’s) or crash modification factors (CMF’s). There are two main methods for assessing the traffic safety effects of countermeasures, and they are known as the before-and-after method and the cross-sectional method.

2.3.1. Before-And-After Method

In the before-and-after (BA) method measurements of traffic safety are taken before and after a particular major change (i.e. safety countermeasure), and the difference between the levels of safety is attributed to the change. Typically, study periods of at least 2 to 3 years are used for both the before and the after period, while crashes that occurred during the construction period are usually omitted (Griffith, 1999). The advantage of the before-and-after method is that the before- and after-data are taken from the same location; given that the only major change in the location during the study period is the intervention of interest, then the effects of changes in other factors is minimized. The disadvantage is that it requires a certain degree of planning, since it relies fundamentally on accidental histories before and after the event.

2.3.2. Cross-Sectional Method

The cross-sectional method for determining traffic safety effects does not necessarily deal with countermeasures, per se. Instead, all countermeasures are treated as “features.” A large sample of roadway segments is included, and the absence or presence of various features is recorded (Shen and Gan, 2003). A snapshot of the level of traffic safety in the location is also included. Regression methods are then employed to investigate the safety effects of having or not having certain road features. The study periods are typically short, and the intersections do not experience any major changes within the period of analysis.
The advantages of the cross-sectional method are the flexibility and time of study (less planning required) and the smaller data requirements. However, the regression is only useful and accurate for factors that have been accounted for (Benekohal, 1991). Consider the example of a certain road feature “A” which has not been identified but has a profound effect on traffic safety. It is a safer assumption that the characteristic of feature A for a particular site remains the same throughout the entire study period of a BA study than to assume that the characteristic of feature A is the same for all locations in a cross-sectional study.

2.3.3. Fully Experimental Study Method

One other important evaluation methodology is the fully experimental study design. In contrast to a before-and-after study, treatment sites are selected at random, as with their consequent comparison group. The advantage of experimental study over before-and-after is that it minimizes the existent bias. In BA studies, treatment sites are selected because of a unique characteristic which warrants the countermeasure, whether it is an overrepresentation of a particular collision type or a particular high collision frequency. Unfortunately, despite the potential benefits, experimental road studies are few (Basile, 1962), mainly because it is difficult to justify randomly choosing locations to implement safety countermeasure while being ignorant to the fact that certain locations may exhibit greater effectiveness than others. This is analogous to medical studies, where the golden standard for study design is a randomized controlled experiment; however, it is often met with resistance by a principle known as intention-to-treat (ITT), particularly in experiments where the control group is not given any type of treatment (Montori and Guyatt, 2001). Similarly, ITT states that it is unjustifiable to delay treatment to a particular patient for the purpose of conducting a study. In conclusion, experimental study currently holds little role in road safety evaluation.

2.4. Overview of Before-And-After Studies

BA studies have the advantage over cross-sectional studies of being able to attribute the change in level of traffic safety to a particular intervention, given that other factors change minimally between the before-and after-periods. However, even with this advantage, naively assuming that the difference between the raw accident histories is the treatment effect itself can be fairly erroneous; several reasons exist. These
reasons can then be summarized into five main factors that are responsible for instigating observed changes in the level of safety between the before-and-after periods, with the first being the treatment effect, while the remaining four are often referred to as confounding factors.

- **Factor 1 – Treatment Effect**: Determining the magnitude and direction of this effect is the primary focus of any evaluation study.

As mentioned previously, traffic collisions are rare and random effects. Thus, the particular accident history at a location is a function of both the true level of road safety plus a degree of random fluctuations. In other words, determining the treatment effect should deal with the difference between the true (expected) safety levels rather than the observed safety levels. This distinction is closely tied with a phenomenon known as regression artifacts or regression to the mean (RTTM).

- **Factor 2 – Regression to the Mean (RTTM)**: Regardless of the approach to site selection, most safety countermeasures are implemented on sites that have an overrepresentation of collisions, whether it is total collisions or collisions of a particular type. This type of prioritization is intuitive in light of limited funding for safety projects. Trouble occurs when the observed level of safety (i.e. indicated by high collision frequency or rates) is due to random fluctuations rather than being a systemic indication of the true level of safety. Not only does this decrease the effectiveness of targeting certain locations, it causes a problem in the post treatment evaluation. RTTM concept indicates that if the extremity of a particularly high value is due to random fluctuation (rather than being systemic), then there is a high probability that the subsequent value will be closer to the true mean and be lower than the preceding value. In the context of traffic safety evaluations, this means that if sites are selected for their high collision count in the before-period, if RTTM effects are not accounted for, the after-period is likely to register a lower count, even if no countermeasures are implemented. Failure to address RTTM ultimately overestimates the benefits of safety countermeasures.

Although BA studies compares collision frequency in the before- and after- periods of the same location, when dealing with periods in the magnitude of years, it is not uncommon for certain characteristics of the intersection to change over time, whether it be with respect to traffic volume or overall traffic trends.
• **Factor 3 – Exposure Effects:** The most common measure of exposure is traffic volume, which can be represented in a number of ways. It can be represented as the total volume entering the location in a set period, be separated into major or minor entering traffic volumes, or even be separated down to the particular movement. The interaction between traffic volumes and their relationship with collision frequency can also be represented in a number of ways. Traffic volume can vary over time because of various reasons, such as increased demand of travel, population growth, or a change in the capacity of the intersection (i.e. conversion of a through-lane into a left-turn bay). Regardless of the methodology for accounting for exposure, ignoring its effects could have a significant impact on the evaluation (ITS, 2007). In some cases, measures to improve traffic congestion may actually end up attracting more traffic volume, as is the case in Persaud et al.’s (2001) study which found that installation of roundabouts at a number of intersections actually increased AADT.

• **Factor 4 – Trend Effects:** Idealistically, all other individual factors that have a significant effect on traffic safety are identified and the changes in each one from the before- to the after-period can be mapped. Unfortunately, asides from exposure, many factors are not recognized, measured, or even understood. Examples of these include changes in vehicle composition (i.e. nearby construction of highway leads to an increase in proportion of heavy traffic), or month-to-month seasonal variations.

The last one is not so much a factor, but an observed (although without universal consensus) phenomenon that arises as a result of safety countermeasures. Often, safety countermeasures target a specific location and target specific types of accidents. Accident migration theory states that following a safety intervention, a decrease in the target type of accident may be followed by a subsequent increase in other types of accidents or an increase in accidents at nearby sites.

• **Factor 5 – Crash Migration:** Crash migrations can be divided into non-geographical and geographical. In the former, collision frequency migrates within the same location across: (i) severity, and (ii) type. For example, installing traffic signals at an unsignalized intersection may decrease the frequency of cross-traffic collisions but increase the frequency of rear-end collisions. In the latter, collisions may migrate from the treatment site to nearby untreated sites; for instance, installation of speed humps on an arterial road section may be met with an increase in collisions.
in nearby road because of rerouting. The existence of crash migration was first documented by Boyle and Wright (1984); since then numerous studies and meta-analysis have attempted to confirm and determine the mechanism and prevalence of the problem (Elvik, 1997). To this date, research on crash migration remains inconclusive. For an evaluation point of view, there are several methods of accounting for crash migrations. For non-geographical crash migration, collisions and subsequent evaluations can be separated for different severities or types. In the event where crash migration seems apparent, an economical evaluation of the net effect of the countermeasure may be useful in determining the overall effectiveness. For geographical crash migration, sites near the treatment sites can be included in the evaluation. It should be noted that in such a case, it may be important to include both nearby intersections and nearby arterial road sections, since intersection changes can affect collision frequencies at upstream and downstream road sections. Finally, with regards to the scope of consideration, Mountain and Fawaz (1992) have proposed 500 m as the limit of inclusion.

The aforementioned five factors are theorized as the ones that are responsible for any observed changes in the level of safety at a particular site following a traffic safety countermeasure. The objective of a sound treatment evaluation is then to account for the four extra factors in order to isolate the treatment effect factor. To this date, it is difficult to objectively scale the quality of a particular evaluation, as evident by a comprehensive study done by Elvik (2008). Nevertheless, the study does conclude on a number of major threats that are relevant to a BA study and that should be addressed as effectively as possible.

There have been studies that have looked to the consequence of failure to account for the aforementioned factors. Figure 2-2 below (Elvik, 1997) proposes the effects of not accounting for a certain confounding factor. For a theoretical road safety countermeasure that actually has no benefits (0% Change), inability to account for any of the factors will lead to an observed reduction in collision frequency of 55%. 
Another useful example is a study performed on the data from a road safety intervention in Norway (Elvik, 2002). The study compared the difference between the observed (raw) collision frequency reductions with the adjusted reductions (accounting for confounding factors) for nine different road safety measures. The results of the analysis are shown below in Figure 2-3. For most of the countermeasures, failing to control for confounding factors led to an overestimation of the treatment effect, particularly for traffic separation and black spot treatment. For the latter, the substantial difference between the two may be attributed to the fact that black spot treatments are performed on locations that have been identified as having high collision frequencies for a period of time. This method of selection inevitably leads to greater tendencies for RTTM effects. Accounting for lane addition and medians actually saw a further 3% increase in effectiveness when controlling for confounding factors.
The discussion of this section exemplifies the objectives and goal of any BA study. An important note to make is that a study not only accounts or doesn’t account for a confounding factor; rather, there is also the extent of accounting for a particular confounding factor which is closely tied to the quality and the validity of the technique. In some cases, a study may claim to control for confounding factors, when in reality the factors were either actually not controlled for, or that the technique employed does not adequately control for the factors. For example, Ogden’s (1997) study on the safety effects of paved shoulders on rural highways claimed that it controlled for trend effects and RTTM effects. However, closer inspection of his data and methodology revealed that his sample size and techniques would not have been adequate enough to effectively account for either of the effects.

2.5. Overview of BA Study Methodologies

In the literature, there are various methods for conducting BA studies, each with its advantages and disadvantages. In addition to accounting for the aforementioned confounding factors, the decision making process for choosing a particular methodology over another is governed by:
• **Format and Availability of Data:** The importance of availability of data is held paramount for BA studies (Elvik and Mysen, 1999). This applies to various types of data, such as collision history, traffic volume, and any other data that may be relevant when used to account for confounding factors. The imperativeness of having consistent, reliable, and detailed collision history data cannot be overstated. Consistency across data ensure compatibility of the collision history both temporally (through the years) and between locations (which is usually satisfied unless the study spans multiple traffic jurisdictions which may have differing accident record practices). Reliability and detailed collision history data is dependent on a systematic process of accident recording which strikes a balance between the desired level of detail and the expendable amount of effort.

• **Intellectual and Economical Resources:** As a trade-off for improving qualities of BA study estimates through the advancement of techniques, the statistical complexity of the studies is also increasing. Moreover, certain techniques, as discussed later on, can require vast amounts of data and analysis, which may accumulate to large amounts of costs.

2.5.1. Four Main Categories of BA Studies

There are four main categories of BA studies, three of which have been used in road safety evaluations, while the last one has only recently been used in research. A quick overview of each method is given.

1. Naïve Before-and-After Study

   • The naïve before-and-after study will be mentioned simply for completeness purposes. It is the simplest technique for observational study, and the treatment effect is simply the difference between the observed collision frequency in the before- and after-period. This method does not account for any of the previously mentioned confounding factors and cannot isolate the treatment effect. It may be useful for a quick first impression, but it generally holds little value in a formal BA study evaluation because of the potential erroneous and misleading results (ITS, 2007). The data requirements and complexity of this method are low, since it only requires accident history of the treated sites, and the computations are fairly simplistic.

2. Before-And-After Study with Comparison (Yoked Comparison or Comparison Groups)
• The second methodology is a BA study whilst having a comparison group or comparison sites. Within this type of study, in addition to the treated sites, comparison sites are selected for each treated site. The criteria for comparison sites are that they are untreated and that they are similar with respect to geometry, intersection type, zoning, volume, etc, with the particular treated site (Griffin and Flowers, 1997). A single comparison site may be selected for a treated site, or a group of comparison groups may be included for a particular study, with each treated site having its own comparison groups which draws from the pool of comparison sites and overlap is typically allowed. The advantage of this methodology is that the comparison site or groups are used to account for trend effects, one of the confounding factors. Depending on the data availability, traffic exposure effects can also be included into the study.

3. Empirical Bayes Methodology

• The Empirical Bayes (EB) methodology is typically recognized as the most popular methodology for conducting traffic evaluations in practice. It combines various techniques of collision modeling and statistical Bayesian inference to estimate the expected level of safety for a treated site had no treatment taken place. This is then contrasted with the observed level of safety after treatment to determine the treatment effect. A more in depth discussion on the EB method will ensue, but the advantages of the EB method is that it can account for long-term trend effects, exposure effects, as well as RTTM effects, which was unaccounted for in the previous methodologies. The disadvantages include the substantial increased complexity and the high data requirements.

4. Full Bayes (Fully Bayesian) Methodology

• The Full Bayes (FB) methodology has only recently been developed and applied to traffic safety evaluations. Firstly, it should be noted that the EB methodology is actually an abbreviated version of the FB methodology which simplifies certain steps, greatly reducing the computational demands. Moreover, the FB method addresses several drawbacks of the EB method, particularly on the EB method’s heavy reliance
of large datasets and its inabilities to carry statistical uncertainties through to the final estimates.

The latter three techniques will be discussed more in depths in the following sections. Error! Reference source not found. below is a comparison of the four technique categories with respect to their ability to control for confounding factors and their data and computation requirements.

<table>
<thead>
<tr>
<th>Technique Category</th>
<th>Controlling for Confounding Factors</th>
<th>Minimum Data Requirements</th>
<th>Complexity and Computational Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exposure Effects</td>
<td>Long-Term Trend Effects</td>
<td>Regression-To-The-Mean Effects</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Naïve Before-and-After</td>
<td>(Maybe)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Before-and-After with Comparison Methodology</td>
<td>Yes</td>
<td>No</td>
<td>(Maybe)</td>
</tr>
<tr>
<td>Empirical Bayes Methodology</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Full Bayes Methodology</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2-1. Comparison of BA Methods

2.5.2. Study Period Length

A topic that warrants revisiting is the selection of the study period, which is imperative in addition to selecting a particular methodology for BA study. As mentioned before, the fact that collisions are rare and random events causes several difficulties because of the lack of statistical stability around raw collision data, particularly when the study periods are short. The typical study period is about 3 years, typically to reduce the amount of random fluctuations as well as fluctuations that may arise due to inter-year seasonal changes. This is important since the effects of seasonal changes on collision prevalence are not always easy to interpret. For example, the wet road conditions due to raining weather may increase the
frequency of collisions because of decreased vision, but may also decrease the associated severity because of lower driving speeds.

Having a long study period is statistically beneficial but causes three practical problems. As mentioned before, the first one is that it is difficult to justify delay treatment for the purpose of conducting post-treatment evaluations; even if the implications are understood by traffic safety authorities, the idea may not be well accepted by the general public.

The second is with respect to the confounding factors, particularly exposure effects and long-term trend effects. While increasing the length of the study period may decrease the potential RTTM effects by reducing the influence of a particular extreme point, it also increases the chance of substantial changes in the traffic exposure and traffic trends of the sites. While changes may occur even for short study periods, it is much easier to account for smaller changes because the estimation of effect is much more accurate for smaller changes than for larger ones. For example, the exact relationship between traffic volume and collision frequency has not been established. Nevertheless, a linear approximation between the two may be fairly accurate for small changes in exposure, but perform poorly for larger volume changes.

The third problem relates to the format of the data itself, especially with respect to consistency. Traffic safety programs may span across several traffic jurisdictions, each with their own collision recording practice. Moreover, even with a particular traffic jurisdiction, recording practice changes throughout the year to reflect either changes in available resources or the desired level of detail. These two factors severely threaten the compatibility of the data, in some cases invalidating the evaluation. To illustrate a point, several significant changes in collision recording practice in Australia are listed below in Table 2-2 (BTE, 2001), each of which would invalidate the evaluation if any of the changes took place within a particular study period.
### Traffic Jurisdiction

<table>
<thead>
<tr>
<th>Traffic Jurisdiction</th>
<th>Year</th>
<th>Change in Recording Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>North South Wales</td>
<td>1988</td>
<td>No longer need to record all collisions, only when property damage exceeds $500</td>
</tr>
<tr>
<td>Western Australia</td>
<td>1980</td>
<td>Property damage threshold (below which it is not necessary to record the collision at all) increased from $100 to $300</td>
</tr>
<tr>
<td></td>
<td>1988</td>
<td>Further increased from $300 to $1000</td>
</tr>
<tr>
<td>Queensland</td>
<td>1991</td>
<td>Property damage threshold increased from $1000 to $2500</td>
</tr>
<tr>
<td></td>
<td>1989</td>
<td>Responsibility for maintaining official crash data transferred from ABS to Statistics Unit of Queensland: Department of Transport</td>
</tr>
<tr>
<td>Victoria</td>
<td>1988</td>
<td>Adopted a 3-Tier Severity System (Fatal, Injury, PDO) from the previous 4-Tier System (Killed, Hospitalized, Treated, and First-Aid)</td>
</tr>
</tbody>
</table>

Table 2-2. History of Crash Recording Practices in Several Traffic Jurisdictions in Australia

### 2.6. Before-and-After Study with Comparison

As mentioned before, a BA study with Comparison can either be performed on a single comparison site or a comparison group for each treated site. The advantage of the latter is the greater statistical stability for representing long-term trends because of the averaging of effect from several sites. The theory behind the methodology is fairly straightforward, and introduces the term odds ratio. Odds ratio can sometimes be understood by the relationship where \((1 - \text{Odds Ratio})\) represents the treatment effect, and can be summarized in the following equation for a particular treatment site \(i\).

\[
(O, R.)_i = \frac{B}{D} \frac{C_i}{\sum_{j=1}^{n} A_j}
\]

Where:
- \(A_i\) = Accident count of comparison site \(j\) in the before-period
- \(B_i\) = Accident count of treat site \(i\) in the before-period
- \(C_i\) = Accident count of comparison site \(j\) in the after-period
- \(D_i\) = Accident count of treated site \(i\) in the after-period

BA Study with Comparison may be useful for evaluating traffic implementation programs where sites were selected at random, in which case the evaluation is similar to an experimental design, and RTTM effects are supposedly minimal since sites were not selected on the basis of high accident counts. For example, the Insurance Corporation of British Columbia (ICBC) launched a program in 1998 that would ensure that there was a stop sign at every second intersection along any particular road. For the evaluation, 133 untreated locations (sites where a stop sign was not installed) were selected to represent...
the pool of comparison sites for 380 treated sites, and a subsequent study was performed using the said technique (Sayed et al., 2006). Because RTTM bias was minimal (due to random selection), and exposure effects were controlled under the assumption that the changes exhibited in the comparison sites reflected regional changes in traffic volume, the resulting 52.8% reduction was attributed directly to the treatment effect.

2.6.1. Selection of Comparison Sites

The selection of comparison sites is not necessarily trivial. The overall approach is that comparison sites should be selected based on its similarity and proximity to the treatment site. What constitutes similarity is of much debate; the relationship between various location features and traits and level of traffic safety is not fully understood yet. As such, it is even more difficult to determine which traits should be used as a basis of comparison. Some common characteristics used include the volume of traffic, overall intersection configuration, or even signalization. The choice of a comparison site based on proximity is akin to a two-edged sword. Constricting the allowed distance limits the number of comparison sites available; expanding the limits too much risks including comparison sites which are not entirely relevant. This is further complicated by the fact that even when logical comparison sites are selected, the actual results can be significantly altered depending on the chosen set. This is shown in Scopatz’s (1998) re-analysis of the study done by Hingson et al. (1996) on the effects of the probability of a fatal collision caused by a driver with blood alcohol content (BAC) over 0.08% after reducing the limit from 0.10% using a different but similarly logical set of comparison sites. Scopatz’s results found a 5% reduction in the probability, while Hingson et al.’s study showed a staggering 16% reduction.

2.7. Empirical Bayes Methodology

The introduction of the EB methodology marks the introduction of Bayesian analysis into BA studies for road safety evaluations, which is relevant for both the EB and the FB method. It should be noted that for a black spot treatment program which identifies, ranks, and prioritizes hazardous locations for countermeasures, Bayesian analysis is used as early as the identification stage to estimate the “expected” collision frequencies, which is a more accurate representation of the level of safety. A few key points distinguish Bayesian analysis from conventional methods.
1. With conventional methods, the level of safety at a certain location is inferred by either of two clues: (i) the accident history, as seen with the first two BA techniques, or (ii) with its traits, which is similar to the cross-sectional study design. Bayesian analysis uses statistical methods to combine these two clues together (Hauer, 1971).

2. Whether the parameter of interest is collision rate or frequency, conventional methods are interested in solving for point estimates as indicators of level of safety. On the other hand, Bayesian analysis treats the unknown parameters as random variables with specific probability distribution, resulting in parameters that represent both the mean value as well as the spread (variance) of the data.

3. Conventional methods rely on data either from the treatment site itself or comparison sites. Bayesian analysis introduces a 3rd group of locations, known as reference population. Reference sites (along with their accident history and traffic volume data), are used to develop Safety Performance Functions (SPF’s) or Collision Prediction Models which predict the collision frequencies based on traits and traffic volume. A reference population often needs to be fairly large (particularly in the EB method), but the criteria for including a reference site is not as stringent, although similar, to that of selecting a comparison site. It should be noted, however, that there are debates as to what dictates a reference population (Wright et al., 1988; Elvik, 1988; Mountain and Fawaz, 1989). In the FB methodology, the reference population and comparison population are in fact, the same, although this will be further discussed later on.

The specific steps for implementing an EB study is fairly well demonstrated by Estimating Safety by the Empirical Bayes Method: A Tutorial by Hauer (1992). To provide a comparison in methodology with the BA Study with Comparison, the procedures for an EB study will be presented in a backwards fashion, beginning with the computation of the odds ratio, the EB refinement, and the development of SPF’s.

2.7.1. Computation of the Odds Ratio in the EB Method

The equation for the odds ratio in the EB method is identical to that of the previous method. The following is the computation of the odds ratio in the event where supplemental data is available for:

- Reference population accident history (single year)
Comparison sites/population accident history (both before- and after- periods)

\[(O. R.)_i = \frac{B}{D} \frac{\sum_{j=1}^n C_j}{A_j}\]

The only difference is how the equation deals with the second term \(B_i\), where:

\(A_j\) = Accident count of comparison site \(j\) in the before-period

\(B_i\) = EB Estimate of count of treat site \(i\) in the after-period had no treatment taken place

\(C_j\) = Accident count of comparison site \(j\) in the after-period

\(D_i\) = Accident count of treated site \(i\) in the after-period

An alternative to the above version is when before- and after-period accident histories are available for the reference population, in which case the reference population can also take the role of the comparison group (Sayed and de Leur, 2004), where:

\(A_j\) = Accident count of reference site \(j\) in the before-period

\(B_i\) = EB Estimate of count of treat site \(i\) in the after-period had no treatment taken place

\(C_j\) = Accident count of reference site \(j\) in the after-period

\(D_i\) = Accident count of treated site \(i\) in the after-period

The EB Estimate is a product of the Empirical Bayes Refinement process, which combines the two clues of accident history and predicted collision frequency to estimate the level of safety at a site in the after-period had not treatment taken place.

2.7.2. Empirical Bayes Refinement

The following is the process with which the EB method combines two clues to estimate the level of safety for a location (Hauer, 1992). The EB Estimate is computed by:

\[EB_{Estimate} = \alpha \times E(A) + (1 - \alpha) \times C\]
Where: $\alpha = \frac{1}{1 + \frac{\text{Var}(E(\Lambda))}{E(\Lambda)}}$

$E(\Lambda)$ = Predicted collision frequency, estimated by SPF

$C$ = Observed collision frequency in study period

$\text{Var}(E(\Lambda))$ = Variance of SPF prediction

The terms $E(\Lambda)$ and $C$ represent the two clues that were mentioned previously. $E(\Lambda)$ is the predicted collision frequency, obtained using SPF’s. It represents the expected collision frequency at a particular site, given its various traits, which may include its traffic volume and other road features. $C$ is the observed collision frequency in the study period, which also represents the clue of the site’s accident history.

The remaining two terms $\alpha$ and $\text{Var}(E(\Lambda))$ are related to the weight that is assigned to either of the two clues. Since SPF’s are developed using empirical data, there is a certain degree of uncertainty around its predictions, which is represented by $\text{Var}(E(\Lambda))$. The term $\alpha$ takes into account the value of the prediction itself and its variance to decide how much weight should be placed on the prediction when computing the EB estimate. To illustrate, if the variance of the SPF is large, then $\alpha$ will be small, meaning there is little confidence in the prediction of the SPF. This is then reflected in the weight that is placed upon the prediction $E(\Lambda)$ and consequently places greater weight on the accident history and vice versa.

The theory behind this process is Bayesian inference, which is used to update the probability estimate for a hypothesis as additional evidence is learned. For the case of the EB method is road safety, the prior distribution is developed using data from the reference group. As the accident history data “becomes available”, the prior distribution is updated to produce the posterior distribution of accident counts.

2.7.3. Safety Performance Functions: Overview and Model Structure

In the most general sense, a safety performance function (or collision prediction model) predicts the collision frequency (or rate) of a location based on site characteristics. Consider the vector $Y = (Y_1, \ldots, Y_n)$ which represents a collection of $n$ number random parameters $Y_i$, which corresponds to the collision frequency at site $i$ within a study period of $T_i$. In this case, it is implied that theoretically, the study period of each site can vary across the analysis. However, in practical use, a single analysis is rarely performed on varying study periods. Thus, in most cases, $T_i$ can be considered to be a constant (Miranda-Moreno
The mean value of $Y_i$ is then a function of various attributes and variables that pertain to site $i$, such that:

$$Y_i = f(\theta_i)$$

Where:

- $\theta_i = f(F, x; \beta)$
- $F$ = Vector of traffic volume variables
- $x$ = Vector of site specific attributes
- $\beta$ = Vector of regression parameters for the variables, estimated from the data

Given the aforementioned structure, the process of developing an SPF is then to:

- Select the model structure, which dictates how the variables relate to one another
- Specify the distributions of each of the variables

**Model Structure**

For illustration purposes, below are two common forms of the function $\theta_i$ (Hauer, 1997; Miaou and Lord, 2003):

$$\theta_i = (F_{i1} + F_{i2})^{\beta_1} \exp(\beta_0 + \beta_3 x_{i3} + \cdots + \beta_k x_{ik})$$

$$\theta_i = F_{i1}^{\beta_1} \cdot F_{i2}^{\beta_2} \exp(\beta_0 + \beta_3 x_{i3} + \cdots + \beta_k x_{ik})$$

Where:

- $\theta_i$ = Collection of site specific attributes for site $i$
- $F_{i1}$ = Traffic on major approach of intersection for site $i$
- $F_{i2}$ = Traffic on minor approach of intersection for site $i$

The rest of the variables represent the aforementioned covariates and the regression coefficients, which are estimated with the reference population data. The particular forms listed above satisfy the intuition that no volumes entering an intersection should lead to no collisions (Sawalha and Sayed, 2006). The
main difference is that while the first equation assumes that the exposure of an intersection to collisions is related to the sum (or total) volume entering into a location, the second equation assumes that it is represented by the product of the opposing traffics, which means that the exposure is represented by the total number of possible conflicts caused by opposing vehicles. Some other model structures are discussed in Chapter 4. Li et al. (2008) emphasizes that choosing the right structure is as important as choosing the right methodologies. Turner and Nicholson (1998) developed a complex form which disaggregated volumes into pairs of conflicting vehicle movements (instead of direction of travel). Such a form requires data on detailed turning flows and crash movement types.

2.7.4. Safety Performance Functions: Distribution Specification and Resulting Models

One of the first models for collision frequency specified that collision frequencies at an intersection for a certain time duration T was represented by a Poisson distribution on the grounds that collisions are rare and random events. This led to what is known as the Poisson regression model, where:

\[ Y_i|\theta_i \sim \text{Poisson}(\theta_i) \]

The nature of a Poisson distribution is that the mean and the variance are the same, which leads to:

\[ E[Y_i|\theta_i] = Var[Y_i|\theta_i] = \theta_i \]

While this model was accurate and reasonable in the occurrence of collisions, it also fixed the variance to be equal to the mean. Past work with collision data indicates that collision frequency for a particular site is usually over-dispersed, meaning that the variance is greater than the mean (Winkelmann, 2003). This over-dispersion has been attributed to various reasons, with one of them being the unobserved heterogeneity within the data (Hauer, 1997). Unfortunately, even with improving methods of collection, certain variables may remain immeasurable, so it is more intuitive to deal with the problem of over-dispersion rather than waiting for the identification and measurement of all safety related variables.

Two-Stage Poisson Model
The two-stage Poisson model is developed specifically to address the issues of over-dispersion. It involves adding a multiplicative random effect term to the Poisson regression model to relax the constraint where the variance must equal to the mean. The revised form of the equation becomes:

$$\theta_i = \mu_i \exp(u_i)$$

Where:

- $\mu_i = f(F, x_i; \beta)$
- $\exp(u_i)$ = Multiplicative term which relieves the constraint

The selection of the distribution of this term determines the type of two-stage Poisson model.

**Negative Binomial Regression**

This is the most common and widely accepted form of the two-stage Poisson model, and also one of the most important ones because of its simplicity in computation. It is obtained by specifying that the multiplicative term has a gamma distribution.

$$\exp(u_i) | k \sim \text{Gamma}(k, k)$$

Where: $k$ = Represents the parameters of the gamma distribution

With respect to the Empirical Bayes Refinement, $\alpha$ is also known as the dispersion parameter, and within the negative binomial regression approach, it can also be represented by $1/k$. Under this structure, the parameters for the random variable $\Lambda_i$ then becomes

$$E(\Lambda_i) = \mu_i$$

$$\text{Var}(E(\Lambda_i)) = \mu_i + \frac{\mu_i^2}{k}$$

Conveniently, when the collision data is not over-dispersed, $k$ approaches infinity, making $\text{Var}(E(\Lambda_i))$ decompose into the mean, which is identical to the Poisson regression model. The negative binomial regression is currently used extensively because of its straightforward calculations with respect to the
posterior analysis, giving merit to the fact that the gamma distribution is a conjugate prior to the Prior distribution.

Although the introduction of the multiplicative term has relaxed the constraint of equivalence between the mean and the variance, some researchers have further suggested that the dispersion parameters should vary across different sites as a function of certain covariates (Heydecker and Wu, 2001; Miaou and Lord, 2003). In this regard, the dispersion factor $\alpha_i$ is made a function of certain covariates, leading to what is known as the generalized negative binomial model.

**Poisson-Lognormal Regression Model**

While the gamma distribution is the most prominent and popular form of the two-stage Poisson model, another common distribution the term $u$ is a normal distribution, leading to the multiplicative term to have a log-normal distribution.

$$\exp(u_i) \sim 0, \sigma^2 \sim \text{Lognormal}(0, \sigma^2)$$

Some researchers have found that the Poisson-Lognormal regression model produces better fit depending on the particular dataset, attributing to the fact that the lognormal tails are asymptotically heavier than those of the Gamma distribution (Kim et al., 2002). However, the computational efforts for the Poisson-Lognormal Regression model are significantly higher than for the negative binomial model (Rao, 2003).

2.7.5. Safety Performance Function: Modeling Approach

Given the aforementioned discussion, following the selection of the model form and the specification of the variable distributions, the last related topic is the process of estimating the regression coefficients from the data. Given that most model structures are non-linear and the non-normal error distributions, simple linear regression has been recognized as been inadequate for modeling collision data (Hauer et al., 1988).

Instead, most safety performance functions (consequently its regression coefficients) are developed using generalized linear regression modeling (GLIM), which is available in various statistical software packages.
The advantages of GLIM are that it can take on a range of different probability distributions, and software packages are able to convert non-linear model into linear forms.

2.7.6. Issues with Empirical Bayes Method

It is important to recognize that despite being the current state-of-the-art in practice, the EB method has several drawbacks. This also suggests that despite the EB methods ability to account for all four confounding factors, there is still potential for further refinement and improvements.

- **Inability to Conduct Multivariate Analysis**: El-Basyouny and Sayed (2010) suggested that crashes of different severity and crashes of different types can be strongly correlated. Within the EB method, these separate estimates are modelled explicitly and independent of one another. It has been suggested that in light of the strong correlation, multivariate modeling can lead to more accurate and precise estimations.

- **Cost of Developing SPF’s**: Most of the effort within an EB method is placed on developing the SPF’s (Carriquiry, 2004). This is often the most cited shortcoming of the EB method. Developing SPF’s require large amounts of data, and the transferability of developed SFP’s from one region or one type of analysis to another is low due to the compatibility of data formats. While a sufficient number of comparison sites are usually available, a obtaining a large reference population is difficult, and often sites may be chosen substantially far away from treatment sites, possibly from different jurisdictions. The use of inter-jurisdiction sites for reference population should be taken with caution. Even if site features are similar, differences in jurisdiction specific policies of driving can have a profound effect on the level of safety. This includes legal BAC limits, driving education program, speed limit, etc.

- **Uncertainties Not Carried Over**: After SPF’s have been developed, values computed from the SPF’s are accepted as true values (Park et al., 2009). That is to say that the EB method uses the results of the SPF’s in such a way that assumes it is the ground truth. However, any model developed with empirical data will have an associated degree of uncertainty. In the EB method, the uncertainty does not propagate through the steps. Fundamentally, the development of SPF’s is performed externally from the actual analysis and any information or characteristics of the reference population data are not transferred to the latter steps. The result is overconfidence in the parameter estimates and consequently in the variable estimates.
• **Inability to Deal with Multi-Level Data:** It has been proposed that aside from being correlated across different severities and different types, collision data exhibit a multi-level structure that the EB method is incapable of accounting for. Accounting for correlation within groups at different levels can strengthen the parameter estimations. Helai and Abdel-Aty (2009) proposed a hierarchy known as the “5 x St” hierarchy, which states that collision data can be organized into a triangular prism consisting of five levels: geographical region, traffic site, traffic crash, driver-vehicle unit, and occupants. A sixth dimension applies to all levels of the prism, and that is the Spatio-Temporal level.

2.8. **Full Bayes Methodology**

One of the main differences of the FB method from the EB method is that the prior information and all available data are integrated and included into the posterior distribution. Recall in the EB method that SPF estimates are taken as true value, and any uncertainty of the SPF itself is not accounted for in the final estimate. The FB methodology on the other hand carries these uncertainties through the steps. This is achieved by the fact that the FB methodology combines the development of SFP’s and the computation of the odds ratio into a single step.

From a statistical point of view, recall that in the EB method, the distribution parameters of the prior were estimated from the empirical data. The FB method differs in that the parameters of the prior now have their own priors (hyper-priors) which belong to a distribution rather than a point estimate. From a hierarchical standpoint, the FB method adds another level of uncertainty and distribution. In the EB method, parameters are estimated from the auxiliary data, while in the FB method, these parameters are drawn from hyper-prior distributions that depend on a higher-level set of parameters. A typical model specification for the FB method is as follows:

\[ y_{ij} \sim \text{Poisson} (\theta_{ij}) \]

\[ \theta_{ij} = \lambda_i \cdot e_{ij} \]

\[ \log(\lambda_i) = \beta_1 \text{signal}_i + c_i \]
\begin{align*}
    c_i &\sim \text{Normal}(0, \sigma_b^2) \\
    \beta_i &\sim \text{Normal}(m, \tau^2) \\
    \sigma_b^2 &\sim \text{Inv} - \text{Gamma}(\delta, \gamma)
\end{align*}

The specification of the parameters of the higher level of distribution, known as hyper-parameters, is what is considered the most difficult part of the FB method. This leads to two types of specifications, known as informative or uninformative. Informative priors are used if there is prior knowledge about the particular road safety evaluation and past experience. On the other hand, uninformative priors take on specifications simply to capture as wide of a variance as possible since nothing is known beforehand. Some have criticized that blindly specifying hyper-parameters is analogous to the flaw in the use of SPF’s in the EB method. However, studies have shown that the results evaluations that used uninformative priors were comparable to ones that used informative priors (Mirando-Moreno, 2009; Li et al., 2009). Moreover, Carriquiry and Pawlovich (2004) have shown that the FB method is fairly robust to poor specification of the priors. Commonly used priors include the diffused normal distributions (zero mean and large variance) for regression parameters while Gamma distributions are often used for inverse variance parameters (Persaud, 2010).

It should be noted that the FB method was made possible because of recent computational developments, more specifically with respect to the Markov Chain Monte Carlo methods (Carlin and Louis, 2000; Gelman et al., 2004). This is because the complexity of the calculations involved as a result of an extra level of uncertainty made it difficult to carry out with previous techniques. An analysis involving only 20 different locations becomes a 22-dimensional probability distribution (Carriquiry and Pawlovich, 2004).

2.8.1. Full Bayes Models

As mentioned before, in the FB method, the development of SPF’s is incorporated into the computation of the odds ratio. This also means that the FB method does not require a separate reference population to calibrate SPF’s. Instead, the pool of treated sites and comparison sites act as the reference population, and the SPF’s are developed such that they also include the effect of the treatment itself. This is contrasted
against the SPF’s in the EB method which are developed using a large reference population. Another difference is that while the EB method typically deals with the entire study period as a single data point (either total or calculated as per year), the FB method treats each time period as an individual data point. That is, if the time period selected for the analysis is by month, then each month of the year represents a separate data point in the analysis. This allows for two things: (i) the ability to account for seasonal changes throughout the year, and (ii) the ability to look for changes in treatment effects with respect to time.

Thus, the models for the FB method is similar to that of the EB method, but is developed for a single time period (for a single month or year), and includes a number of binary variables to specify where the data point belongs (i.e. treatment or comparison, before or after implementation). Another key advantage of the FB method is the increased flexibility in model specification. Under computational limitations, the EB methodology is constrained to one of few model specifications. However, the FB method is able to take on any model specification which is reasonable for collision prediction (Carriquiry and Pawlovich, 2004; Li et al., 2008).

Univariate Poisson-Lognormal Intervention (PLNI) Model

One of the simplest model forms is the univariate Poisson-lognormal intervention model. The specifications are as follows (El-Basyouny and Sayed, 2010):

\[ Y_{it} \sim \text{Poisson}(\theta_{it}) \]

\[ \ln(\theta_{it}) = \ln(\mu_{it}) + \varepsilon_i \]

\[ \ln(\mu_{it}) = \beta_0 + \beta_1 T_i + \beta_2 t + \beta_3 (t - t_{0i}) I_{it} + \beta_4 T_i t + \beta_5 T_i (t - t_{0i}) I_{it} + \beta_6 \ln(V_{1it}) + \beta_7 \ln(V_{2it}) + \beta_8 X_{8it} + \ldots \]

\[ + \beta_9 X_{9it} \]

Where:

\( T_i \) = The treatment indicator (1 for treated site; 0 for comparison site)

\( t_{0it} \) = Intervention year for the ith treated site

\( I_{it} \) = Time indicator (0 for before; 1 for after-period)

\( V_{1it}, V_{2it} \) = Traffic volume on major and minor approaches, respectively
Multivariate Poisson-Lognormal Intervention (MVPLNI) Model

As mentioned before, one of the advantages of the FB method is the ability to account for multivariate correlations. Park et al. (2009) found that the correlations exist not only between crash severity levels, but also for repeated crash observations. The model specification is similar to that of the PLNI, but with the added subscript to denote each variable, as well as the addition of a covariance matrix (El-Basyouny and Sayed, 2010). A comparison between univariate and multivariate approaches found that the multivariate approach produces slightly smaller uncertainty estimates (Park et al., 2010). While severity is the most typical separation used for multivariate analysis, various other categories can be used. Park et al. (2010) differentiated collisions based on a two condition categorization which included the number of vehicles involved and the total cost of the damages.

Multivariate Poisson-Lognormal Intervention with Random Parameters (MVPLNI-RP) Model

Li et al. (2008) suggested the existence of correlation within matched comparison groups, given the selection of comparison groups to be similar to the treated sites. This led to the development of a model that allows for the regression coefficients to vary for each comparison “group.” The model is as follows (El-Basyouny and Sayed, 2010):

\[
\ln(\mu_{it}) = \beta_{p(i)0} + \beta_{p(i)1} T_i + \beta_{p(i)2} t + \beta_{p(i)3}(t - t_{0i}) I_{it} + \beta_{p(i)4} T_i t + \beta_{p(i)5}(t - t_{0i}) I_{it} + \beta_{p(i)6} \ln(V_{1it}) \\
+ \beta_{p(i)7} \ln(V_{2it}) + \beta_{p(i)8} X_{bi} + \ldots + \beta_{p(i)j} X_{ji}
\]

This equation is identical for that of the previous two models, except for the addition of the subscript \(p(i)\) which denotes which pair or comparison “group” for the \(i\)th treated site. Several other researchers have also looked into the use of random parameters (Shankar et al., 1998; Milton et al., 2008).
2.8.2. Continuous Time Intervals

In other methods (i.e. EB), the total collision frequency or the average collision rate throughout a study period (either before- or after-) are used. However, collision frequencies at a location from one time period to another have been shown to have a certain degree of correlation (El-Basyouny, 2012). Aggregating collision frequency across the entire study period hinders the ability to study trend effects in collision frequencies. The implementation of the FB method allows for each individual year (or month) to be modeled as a separate time interval. Under this form, time trend parameters may be implemented into the models to account for the effects and improve the fit. Three types of collision trends can be identified and can be studied using the example model forms listed above.

- **General Collision Trends**: Under the EB methodology, the before- and after-periods are modeled as a single state of traffic safety. However, it is recognized that there are trend effects that occur with the periods themselves, such as the effects of an ageing population. The continuous time interval modelling approach allows for individual collision trends within both periods.

- **Intervention Time Trends**: It has been recognized that traffic safety intervention effects do not occur instantaneously but rather incur over a period of time (Li et al., 2008). Early researchers of this phenomenon assumed a linear accumulation of effect over time (Li et al., 2008; Park et al., 2010). However, recently a non-linear profile of treatment effect has been proposed (El-Basyouny, 2012).

- **Seasonal Cyclic Trends**: The use of continuous time intervals has allowed for the investigation of another type of temporal trends: season cyclic trends. Li et al et. (2008) proposed using a piecewise variable \( S_t \) to indicate the season to which a month belongs, and then include a sinusoidal term that is a function of \( S_t \) to the safety prediction function.

\[
S_t = \begin{cases} 
1 & \text{if } t \text{ is a winter month (December to February)} \\
2 & \text{if } t \text{ is a spring month (March to May)} \\
3 & \text{if } t \text{ is a summer month (June to August)} \\
4 & \text{if } t \text{ is a fall month (September to } \end{cases}
\]

\[
x_{2t} = \cos\left(\frac{2\pi * S_t}{4}\right) \\
x_{3t} = \cos\left(\frac{4\pi * S_t}{4}\right) \\
x_{4t} = \cos\left(\frac{2\pi * S_t}{4}\right)
\]
Nofal and Saeed (1997) not only look at empirical evidence but also further researches on how seasonal variations can affect road traffic safety in Riyadh City. For cities that experience snowfall or increased precipitation during the winter and fall period, it is assumed that collision rates are highest during those months because of decrease in visibility and road traction because of the accumulation of water and snow on the road. However, Riyadh City, Saudi Arabia, the summer months saw the highest accident rates. Nofal and Saeed looked at measures of temperature, visibility, and humidity and in the case of Riyadh City, increased stress and decrease in performance of intellectual tasks as a result of high temperature and high humidity increased the likelihood of collisions. In fact, despite having virtually zero precipitation, the summer months of July to September registered 32% of the annual collisions, with only 21% of them occurring in the winter months. A note to make is that although observations can be made about varying accident rates during different months using a single year of data, to model the effects of seasonal changes in a safety performance function amidst other variables requires at least many years of monthly collision data, which is often difficult in a BA study.

2.8.3. Spatial Trends

As mentioned previously, another key advantage of the FB methodology is the ability to study spatial trends and effects. Although it is intuitive to assume that locations close to one another or along the same corridor will exhibit similar collision trends, there has actually only be limited research on spatial correlation (Levine et al., 1995). This can be due to the complexity in attempting to account for spatial effects. In comparison, accounting for temporal trends is more straightforward, given the one-dimensional aspect of time. However, in spatial correlation, not only is there physical 2 dimensions to the spatial position of one location from another, there are also more complex organizations of intersections. Several possible approaches to accounting for spatial trends include identifying intersections along the same corridor, intersections that are of high proximity to one another, intersections which fall within a major commuting route, or intersections that fall within a particular zone.

Wang and Abdel-Aty (2006) proposed three different correlation structures and concluded that the correlation increases as the gap between intersection decreases. More importantly, they also concluded
that amongst collision types, rear-end collisions exhibited the highest spatial correlation. Currently, one of the most common methodologies for accounting for spatial correlation is proposed by Miaou et al. (2003), who used an area-based approach and a Conditional auto-regressive (CAR) model.

2.9. Comparison of EB and FB Method

As mentioned before, the FB method displays several key advantages over the EB method.

- Combining SPF Development and OR Computation into 1 Step: Whether it is EB or FB, both methods rely on the development of safety performance function to predict what would have happened had no intervention taken place. Fundamentally, the main difference is that for the EB method, the development of the safety performance function is a separate step, while in the FB method, the development is performed at the same time as the actual analysis.

This has many implications and consequences. Within the EB methodology, once the SPFs have been developed, its results are taken as “absolute”, even when it is well known that the development of any model with empirical data is at best an estimate. This is because the distributions and uncertainties of the function development cannot be carried over to the next step. Calculation the Odds Ratio within EB methodology uses SPFs to estimate expected accident counts, which relies only on the regression coefficients given from the development. Obtaining a large reference population for the development of SPFs is typically the effort used to compensate for this known shortcoming; by including more data points into the development, the uncertainties are minimized. However, this causes problems with respect to both cost and the selection of relevant locations. Depending on the form of the function chosen, the required data for a location for a safety performance function development can include accident counts (separated either by year or by month), volume (aggregated AADT, separated by minor and major approaches), and other site characteristics that may be deemed relevant (width of lane, existence of certain features, etc). While characteristics such as width of lane do not change much over time, data such as volume counts per month are costly to collect. In some cases, even with adequate funds, it may be impossible to find enough similar intersections to develop SPF with adequate confidence to take the results as point estimates, particularly in smaller traffic authorities. In such cases, SPF’s developed by others or in different areas may be borrowed in order to complete the analysis, although the transferability of SPF from one region to another is low. One proposed short coming is that while the inclusion of uncertainty of the SPF’s would produce more realistic and statistically valid estimates, it
would also increase the standard deviation compared to the EB method, which assumes the SPF’s are absolute (Carriquiry and Pawlovich, 2005). However, Persaud (2010) conducted a direct comparison between the EB and FB methodology using the dataset, and found that the standard errors from the FB method were actually smaller, suggesting that incorporating the data uncertainties into the final estimates will not threaten the statistical significance of results. Given the aforementioned reasons, incorporating the development of SPF’s in a single step allows for the final estimates to account for the uncertainties around the regression coefficients, absolving the need to attain absolute certainty of the function because the variability are reflected in the final results.

- Ability to Incorporate Multi-Level and Hierarchical Data: Prediction models in the EB method are developed under the assumption that accident counts from year to year and from site to site are independent of one another. However, it is well known that both of these assumptions are incorrect. In the literature, this is often referred to as the “multilevel” or “hierarchical” nature of traffic data (Huang, 2010). Sites close to one another are known to exhibit correlation in their accident counts, whether it is because they belong to the same geographical regions, or because they share similar traits that are unaccounted for in the prediction models. Year to year accident counts within a site are known to correlate with one another. Huang proposes a 5 x ST hierarchy of accident data, and these are geographical region level, traffic site level, traffic crash level, driver vehicle unit level, and occupant level, while all of these are correlated with the spatiotemporal level. Accounting for hierarchical data structure has been shown to lead to improvements in model fitting (Huang, 2010).
3. REVIEW OF DATA SOURCES

This chapter will review the data that was made available by the City of Edmonton and the Office of Traffic Safety (OTS). A comprehensive review of the datasets and its characteristics is crucial in determining the available techniques for conducting a BA study.

3.1. Dataset Scope

Five treated locations were selected for the study, and their countermeasures were implemented in the years 2008, 2009, and 2010. The details for the five locations are summarized in the table below.

<table>
<thead>
<tr>
<th>ID</th>
<th>Group</th>
<th>Primary Road</th>
<th>Secondary Road</th>
<th>Year of Completion</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Ellerslie Road</td>
<td>91st Street</td>
<td>2008</td>
<td>Modification of entry angle</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Yellowhead Trail</td>
<td>Victoria Trail</td>
<td>2009</td>
<td>Island RT to intersection RT</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>87th Avenue</td>
<td>170th Street</td>
<td>2010</td>
<td>Modification of entry angle</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>Yellowhead Trail</td>
<td>St. Albert Trail (WB Right Turns)</td>
<td>2010</td>
<td>Island RT to intersection RT</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>Yellowhead Trail</td>
<td>St. Albert Trail (EB Right Turns)</td>
<td>2010</td>
<td>Island RT to intersection RT</td>
</tr>
</tbody>
</table>

Table 3-1. Treated Sites Detail

As mentioned before, all the implementations target collisions that arise due to the right-turn movement. Each treated site represents a separate treatment “group” or “matched pairing.” The ID of the sites is not continuous because a number of comparison sites were selected for each treated site. A summary of the comparison sites is shown in the table below.

<table>
<thead>
<tr>
<th>ID</th>
<th>Group</th>
<th>Primary Road</th>
<th>Secondary Road</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>Ellerslie Road</td>
<td>66th Street</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Ellerslie Road</td>
<td>111st Street</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Ellerslie Road</td>
<td>50th Street</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>Yellowhead Trail</td>
<td>50th Street (WB Off Ramp)</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>Yellowhead Trail</td>
<td>Fort Road (WB Off Ramp)</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>87th Avenue</td>
<td>178th Street</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>95th Avenue</td>
<td>170th Street</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>Yellowhead Trail</td>
<td>156th Street (WB)</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>Yellowhead Trail</td>
<td>150th Street (WB)</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>Yellowhead Trail</td>
<td>156th Street (WB)</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>Yellowhead Trail</td>
<td>150th Street (EB)</td>
</tr>
</tbody>
</table>
Site 12 and Site 15 are identical, since the same comparison site was applicable to both treated sites 11 and 14. The comparison sites were selected based on similarity and proximity to the treated site.

### 3.2. Accident History

Collision records were available for all the treated and comparison sites from January 2005 to December 2011. In this study, the selected periods are yearly, and the year of implementation is excluded from the analysis. This leads to a range of before-periods of 3 to 5 years and after-periods of 1 to 3 years. The accident history is recorded in a per-incident manner, such that each reported collision has a separate entry within the data. For each collision, a number of details are recorded and they are summarized in Table 3-3.

<table>
<thead>
<tr>
<th>Detail</th>
<th>Description</th>
<th>Range of Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collision Location Name</td>
<td>Denotes the location at which the collision occurred</td>
<td>Sites [1-16]</td>
</tr>
<tr>
<td>Collision Report Year</td>
<td>Denotes the year in which the collision occurred</td>
<td>[2005 – 2011]</td>
</tr>
<tr>
<td>Collision Report Month</td>
<td>Denotes the month (within the year) in which the collision occurred</td>
<td>[1-12]</td>
</tr>
<tr>
<td>Collision Type Name</td>
<td>Indicates the severity of the collision</td>
<td>[Property Damage, Injury, Fatal]</td>
</tr>
<tr>
<td>Collision Cause Name</td>
<td>Indicates the primary type of collision</td>
<td>[Followed too closely, left turn crossed path, obstruction of view of traffic sign, improper turn, ran off road, improper lane change, backed unsafely, yield sign violation, stop sign violation, etc]</td>
</tr>
<tr>
<td>Travel Direction Primary Object</td>
<td>Indicates the primary direction of travel prior to the collision</td>
<td>[North, East, South, West]</td>
</tr>
<tr>
<td>Driving Lane Primary Object</td>
<td>Indicates the closest lane at which the collision occurred</td>
<td>[Second from curb, third from curb, fourth from curb, right turn bay, left turn bay, unspecified]</td>
</tr>
<tr>
<td># of Collisions</td>
<td>Indicates how many separate collisions occurred in that one incident</td>
<td>[-]</td>
</tr>
</tbody>
</table>

Table 3-3. Collision Data Details
The study makes use of pivot tables, a common tool in spreadsheet software packages, to consolidate the data to the desired level of aggregation. For this study, the initial data is consolidated such that the total number of collisions for each location and for each year is reported, separated by collision severity.

![Figure 3-1. Example of Pivot Table Output](image)

Figure 3-1 above illustrates a typical pivot table output that will consequently be used as the accident history inputs for the subsequent models.

### 3.3. Volume Data

For each of the 16 sites, traffic volume data was also available. A single day (6:30 AM to 6:30 PM) of traffic volume was collected for each site, and the collection was performed between May 26th and June 15th of 2011. The volume data was presented in a format which reports the total volume that passed through the intersection for a given time period. The details (which represents the level of separation) for each recorded volume of traffic are summarized in the table below.

<table>
<thead>
<tr>
<th>Detail</th>
<th>Description</th>
<th>Range of Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time of Day</td>
<td>Denotes the start of the 5-minute interval where the volume was observed</td>
<td>[6:30 AM – 6:30 PM]</td>
</tr>
<tr>
<td>Approach</td>
<td>Denotes the direction of approach for that volume</td>
<td>[North, East, South, West]</td>
</tr>
<tr>
<td>Movement Type</td>
<td>Denotes the specific movement of that volume</td>
<td>[Left turn, straight, right turn]</td>
</tr>
<tr>
<td>Vehicle Type</td>
<td>Denotes the specific vehicle type of that volume</td>
<td>[Bus, car, truck]</td>
</tr>
</tbody>
</table>

Table 3-4. Traffic Data Detail
Moreover, the data is partially aggregated to display the traffic volume during the AM peak hour, PM peak hour, and an off-peak hour. The estimated 24-hour average annual daily traffic (AADT) was also extrapolated using the data. A traffic flow summary diagram, with an example shown in Figure 3-2 below, was generated for each of these aggregations.

![Figure 3-2. Sample Traffic Flow Diagram](image)

The study is mainly concerned with the major and minor approach AADT for each site. Approach volume is defined as *the total volume of traffic coming out of an intersection for a particular direction (both ways)*. That is, for each four-way intersection there are two approach volumes, one representing all traffic coming out of the intersection and travelling either North- or Southbound, and the other for traffic travelling East- or Westbound. The major approach is denoted as the approach with the larger volume. Figure 3-2 is the traffic flow diagram for the AADT of Site 1 (Ellerslie Road and 91st Street). In this case, EW is the major approach. The estimated AADT’s and the major approach for each of the 16 sites are shown in Table 3-5 below.
### Table 3-5. Estimated 2011 AADT

<table>
<thead>
<tr>
<th>ID</th>
<th>Volume E-W</th>
<th>Volume N-S</th>
<th>Major Approach Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21805</td>
<td>12032</td>
<td>E-W</td>
</tr>
<tr>
<td>2</td>
<td>13877</td>
<td>6734</td>
<td>E-W</td>
</tr>
<tr>
<td>3</td>
<td>18621</td>
<td>18610</td>
<td>E-W</td>
</tr>
<tr>
<td>4</td>
<td>5816</td>
<td>18759</td>
<td>N-S</td>
</tr>
<tr>
<td>5</td>
<td>3524</td>
<td>16009</td>
<td>N-S</td>
</tr>
<tr>
<td>6</td>
<td>6631</td>
<td>32312</td>
<td>N-S</td>
</tr>
<tr>
<td>7</td>
<td>16906</td>
<td>42403</td>
<td>N-S</td>
</tr>
<tr>
<td>8</td>
<td>24415</td>
<td>39227</td>
<td>N-S</td>
</tr>
<tr>
<td>9</td>
<td>18501</td>
<td>36758</td>
<td>N-S</td>
</tr>
<tr>
<td>10</td>
<td>22223</td>
<td>42312</td>
<td>N-S</td>
</tr>
<tr>
<td>11</td>
<td>4981</td>
<td>77571</td>
<td>N-S</td>
</tr>
<tr>
<td>12</td>
<td>7029</td>
<td>63321</td>
<td>N-S</td>
</tr>
<tr>
<td>13</td>
<td>6154</td>
<td>67613</td>
<td>N-S</td>
</tr>
<tr>
<td>14</td>
<td>8134</td>
<td>59941</td>
<td>N-S</td>
</tr>
<tr>
<td>15</td>
<td>7029</td>
<td>63321</td>
<td>N-S</td>
</tr>
<tr>
<td>16</td>
<td>15309</td>
<td>86601</td>
<td>N-S</td>
</tr>
</tbody>
</table>

### 3.4. Volume Forecasting

Idealistically, traffic volumes for each intersection are regularly recorded so that there is a corresponding AADT estimate for each of the 16 sites for each of the study years. However, regular detailed intersection volume collection is extremely costly because of the reliance on manual observation counting methods. Arterial traffic volumes (along a corridor) can be collected automatically; however, for intersection volume where the total volume for particular movement is of interest, currently only manual methods are available. In scenarios where there are missing AADT volumes, forecasting techniques are used to extrapolate the data using the available existing data.

Three methods of traffic forecasting were considered for this study, although only the last one was used for the analysis. They are:

- Forecasting using city population growth
- Forecasting using registered vehicle growth
- Forecasting using local traffic volume growth

The first two methods rely on municipal data and work under the assumption that the yearly changes in AADT for the particular site is proportional to the population or registered vehicle growth (or decline) of the surrounding municipality. The drawbacks of these methods is that the assumption may not
necessarily hold true since traffic volumes are also influenced largely by local changes in demand for travel. For example, the construction of a new supermall in one locality may greatly affect the traffic patterns in the immediate vicinities.

As a result, the forecasting will instead be based off local traffic volumes. Although the City of Edmonton does not have regular intersection volume counts, it has 1496 automation traffic recorders installed for arterial sections of roads. The data is available at the City of Edmonton Open Data Catalogue’s *Average Daily Street Traffic Volumes from 2004 – 2009*. However, within the dataset, there are various missing data. This approach of forecasting will make use of the automatic data collection points near each of the study site to forecast the AADT for the missing years using the 2011 baseline. An example of the dataset is shown in the Figure below.

![Figure 3-3. Example of Automatic Traffic Volume Detection Dataset (Source: City of Edmonton Open Data Catalogue)](image)

In the forecasting, the 16 study locations will be labelled as *Primary Sites*, and 4 *Secondary Sites* from 1496 automatic traffic recorders will be selected for each primary site. The average growth factor throughout 2004 to 2009 for each of the secondary sites will be computed, and then averaged along with other secondary sites from the primary pairing. The resulting growth factor is used to backward forecast the traffic volumes for 2005, 2006, 2007, 2008, 2009, and 2010 of the 16 study sites. Within this forecasting method, it is assumed that arterial traffic volume changes near the intersection are proportion to the approach traffic volumes of the intersection.

### 3.5. Summary of Available Data

The study is comprised of 5 treated sites and 11 comparison sites (1 duplicate). For each of the $n = 16$ sites, accident history is available for the years January 2005 to December 2011. The accident histories can be segregated based on various attributes, such as collision cause, severity, etc. Volume data is available in
the form of AADT for the major and minor approaches for each intersection. The 2011 AADT were combined with estimated growth factors to forecast AADT for the remaining years. Thus, for each of the locations, there were \( m = 7 \) periods of accident history and volume data, leading to a total of \( n \times m = 117 \) data points.

Referring to Table 2-1, this dataset allows for naïve BA study, BA study with comparison, and FB methodology. The absence of a reference population excludes the EB methodology from the possible techniques. It should be noted that this dataset represents the minimum data requirements for conducting an FB methodology. However, the FB methodology is actually capable of utilizing much more detailed and complex datasets, whether it is in terms of the hierarchical structure of the data, or data pertaining to speeds, geometric features, etc. A description of the statistics of the data is shown in Table 3-6 shown below.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treated sites during the before period (5 sites)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major volume</td>
<td>15246</td>
<td>73874</td>
<td>40866.04</td>
<td>24623.86</td>
</tr>
<tr>
<td>Minor volume</td>
<td>3356</td>
<td>23251.7</td>
<td>10111.32</td>
<td>7976.88</td>
</tr>
<tr>
<td>Collisions(Avg/Year)</td>
<td>10.7</td>
<td>60.6</td>
<td>27.9</td>
<td>19.62</td>
</tr>
<tr>
<td><strong>Treated sites during the after period (5 sites)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major volume</td>
<td>18759</td>
<td>77571</td>
<td>43460.6</td>
<td>25160.39</td>
</tr>
<tr>
<td>Minor volume</td>
<td>3524</td>
<td>24415</td>
<td>10617.2</td>
<td>8375.97</td>
</tr>
<tr>
<td>Collisions(Avg/Year)</td>
<td>13</td>
<td>42</td>
<td>21.4</td>
<td>11.67</td>
</tr>
<tr>
<td><strong>Comparison sites during the before period (11 sites)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major volume</td>
<td>13215</td>
<td>82474</td>
<td>42067.86</td>
<td>22338.2</td>
</tr>
<tr>
<td>Minor volume</td>
<td>5538</td>
<td>21164</td>
<td>11336.64</td>
<td>6051.58</td>
</tr>
<tr>
<td>Collisions(Avg/Year)</td>
<td>3</td>
<td>56</td>
<td>22.49</td>
<td>19.04</td>
</tr>
<tr>
<td><strong>Comparison sites during the after period (11 sites)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major volume</td>
<td>13877</td>
<td>86601</td>
<td>44172.55</td>
<td>23455.8</td>
</tr>
<tr>
<td>Minor volume</td>
<td>5816</td>
<td>22223</td>
<td>11903.82</td>
<td>6354.35</td>
</tr>
<tr>
<td>Collisions(Avg/Year)</td>
<td>5.5</td>
<td>53</td>
<td>20.36</td>
<td>15.62</td>
</tr>
</tbody>
</table>

Table 3-6. Descriptive Statistics for Data
4. PRELIMINARY STUDY

Since the Full Bayes methodology is a relatively new innovation, it is useful to compare its results with that of other methodologies.

4.1. Studies not Considered

The naïve before-and-after study will not be considered in this thesis because it is universally agreed that its results are too unreliable to be of any use. The EB methodology, although being the current practice state-of-the-art methodology, cannot be performed within this study because of the absence of a reference population, which is necessary for developing the SPF’s.

4.2. Before-And-After Study with Comparison Groups

The comparison will be made against the BA Study with comparison groups, which will make use of both the accident history and volume data. The study will mainly rely on the following formula for the Odds Ratio.

\[
O.R.i = \frac{\left(\sum_{j=1}^{N} A_{ij} / C_{ij}\right) / N}{B_i / D_i}
\]

Where:

- \(O.R.i\) = The odds ratio for treated site \(i\)
- \(A_{ij}\) = Accident rate for before-period for comparison for comparison site \(j\) of treated site group \(i\)
- \(B_i\) = Accident rate for before-period for treated site \(i\)
- \(C_{ij}\) = Accident rate for after-period for comparison for comparison site \(j\) of treated site group \(i\)
- \(D_i\) = Accident rate for after-period for treated site \(i\)

The average of the \(C_{ij}/A_{ij}\) amongst each treated site group \(i\) represents the comparison group adjustment for the corresponding treated site. A summary of the results of the study is summarized in Table 4-1 below.
Treated sites 1 and 14 registered an increase in accidents, while sites 5, 8, and 11 showed a reduction in accident rates. To reiterate, the use of comparison groups accounts for long term trend effects. The use of accident rates instead of frequency accounts for exposure effects. However, the validity of this claim may be challenged because the use of accident rates assumes the linearity between accident frequency and volume, which itself is a topic of heavy debate in traffic safety. Factors that are unaccounted for in this methodology are regression-to-the-mean bias, and crash migration effects.

Often, traffic safety is considered not only with the treatment effect of a specific location, but also on an entire improvement program. Averaging out the effects across the 5 locations produces an overall reduction in accident rate of 8.74%.
5. MODELS AND RESULTS

This chapter will explore various versions of the FB methodology and apply them to the dataset to compare their effectiveness as well as to conduct a formal evaluation of the OTS’ right-turn target program. It will start with a discussion on the selection of functional form and models, move on to the implementation of the FB methodology, and conclude with the results of the analysis. As mentioned before, the data will be analyzed under a year-to-year time frame, and all collisions will be separated by severity into two levels: property-damage only, and severe (injury and fatal).

With respect to controlling for confounding factors, all models within this study will account for:

- **Exposure Effects**: The inclusion of yearly AADT traffic volume will account for the fluctuation in traffic volume and their effects on collision prevalence.
- **Long-Term Trend Effects**: The inclusion of the comparison sites for each treated site will account for the long-term trend effects that are observed near the treated site.
- **RTTM Effects**: The use of Bayesian inference to solve for expected collision frequency will account for the RTTM bias.
- **Crash Migration Effects**: The selection of site does not allow for the control of geographical crash migration effects. However, the separation of collisions by severity allows for the investigation of non-geographical crash migration with respect to severity.

5.1. Functional Form

The functional form of a collision model represents the hypothesized relationship between various site specific attributes and the observed collision frequency. Chapter 2 presented two common functional forms, one which multiplies the major and minor approach volumes, and one which sums them. Miaou and Lord (2003) have conveniently summarized a number of common functional forms. They are shown in the table below; F1 and F2 are the major and minor traffic volume, respectively, and $\beta$ represents the associated coefficients.
For this study, the chosen functional form for accounting for exposure effects for No. 2. There are three advantages of this form, with the first being that it is fairly simple and well understood. The second is that for vehicle-vehicle collisions, it satisfies the logic that zero volume on either approach should lead to zero collisions because of the absence of potential intersections of vehicle paths. Moreover, it also expresses the intuition that the number of potential conflicts is the multiplicative to the volumes on either of the approaches, not scalar. The third is that the exponential coefficients for each of the volumes allows for a large degree of freedom with respect to relationship, allowing for both linear and non-linear representations (Sawalha and Sayed, 2006).

However, it should be noted that there are some drawbacks to the functional form. The multiplicative relationship is effective for expressing potential number of conflicts between intersecting vehicles. Unfortunately, it may lack of the ability to predict collisions involving a vehicle: (i) colliding with another vehicle traveling on the same approach, (ii) colliding with a non-vehicle user (pedestrian or cyclist), or (iii) colliding with a stationary object (i.e. traffic fixtures). Under this functional form, the condition of zero volume in either of the approaches does not allow for any of the abovementioned scenarios. Fortunately, in practice, zero volumes in either approaches is not normally observed.

### 5.2. Model Specifications

As mentioned before, this chapter will test and compare several versions of the FB methodology with respect to models. The specification of the error structure in all models will be the Poisson-lognormal distribution. All models except one will be multivariate across the two crash severities. It will explore the benefits of the addition of the two “features” of allowing for random parameters amongst matched pairs, and allowing for a jump effect. The jump effect was first investigated by Li et al. (2008), who concluded

<table>
<thead>
<tr>
<th>No.</th>
<th>Functional Form for Exposure Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \beta_{0.1}(F_{1,lt} + F_{2,lt})^{\beta_1} )</td>
</tr>
<tr>
<td>2</td>
<td>( \beta_{0.1}F_{1,lt}^{\beta_1}F_{2,lt}^{\beta_2} )</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_{0.1}(F_{1,lt}F_{2,lt})^{\beta_1} )</td>
</tr>
<tr>
<td>4</td>
<td>( \beta_{0.1}(F_{1,lt} + F_{2,lt})^{\beta_1}(F_{2,lt}^{\beta_2}/F_{1,lt})^{\beta_2} )</td>
</tr>
<tr>
<td>5</td>
<td>( \beta_{0.1}F_{1,lt}^{\beta_1}F_{2,lt}^{\beta_2}\exp(\beta_3F_{2,lt}) )</td>
</tr>
<tr>
<td>6</td>
<td>( F_{1,lt}\exp(\beta_{0.1} + \beta_1F_{2,lt}) + F_{2,lt}\exp(\beta_{0.1} + \beta_2F_{1,lt}) )</td>
</tr>
</tbody>
</table>

Table 5-1. Common Functional Forms of Minor and Major Approach
that while the effects of a safety countermeasure may incur over time after the implementation, the implementation may also lead to an immediate effect, termed the \textit{jump effect}.

Thus, for the multivariate analysis, two yearly models are specified.

- MVPLNI – MultiVariate Poisson-LogNormal Intervention by Year
- MVPLNIJ – MultiVariate Poisson-LogNormal Intervention by Year w/ Jump

The collision model for the MVPLNI model is as follows:

\[
\ln(\mu_{it}^k) = \alpha_1^k + \alpha_2^k T_i + \alpha_3^k (t - t_{0i}) I_{it} + \alpha_4^k T_i t + \alpha_5^k T_i (t - t_{0i}) I_{it} + \gamma_1^k \ln(V_{1it}) + \gamma_2^k \ln(V_{2it})
\]

Where:
- \(\mu_{it}\) = Mean accident rate for location \(I\) and year \(t\)
- \(T\) = Treatment indicator (1 for treated sites and 0 for comparison sites)
- \(t_{0i}\) = Intervention time period
- \(V_{1it}, V_{2it}\) = AADT at major and minor approaches, respectively, for year \(t\)
- \(\alpha_1, \ldots, \gamma_2\) = Regression coefficients estimated from simulation
- \(k\) = Crash severity for multivariate analysis (0 for PDO and 1 for severe)

The collision model for MVPLNIJ model is nearly identical, save for the addition of an additional term \(\alpha_6 T_i t\) to account for the jump effect.

\[
\ln(\mu_{it}^k) = \alpha_1^k + \alpha_2^k T_i + \alpha_3^k t + \alpha_3^k (t - t_{0i}) I_{it} + \alpha_4^k T_i t + \alpha_5^k T_i (t - t_{0i}) I_{it} + \alpha_6^k T_i t I_{it} + \gamma_1^k \ln(V_{1it}) + \gamma_2^k \ln(V_{2it})
\]

The computation of the odds ratio for the \(i\)th treated site is given by the following equation:

\[
OR_i = \frac{\mu_{TBI}}{\mu_{CAI}}
\]

Where:
- \(\mu_{TBI}\) = Average collision count for the \(i\)th treated site averaged over after-period intervals
- \(\mu_{CAI}\) = Average collision count for the \(i\)th treated site averaged over before-period intervals
\( \mu_{CBI} \) = Average collision count for the comparison sites corresponding to the \( i \)th treated site, averaged over after-period intervals.

\( \mu_{CAi} \) = Average collision count for the comparison sites corresponding to the \( i \)th treated site, averaged over before-periods intervals.

An additional model is developed to test for treatment effectiveness particularly in targeting right-turn related crashes. This is a univariate model on right-turn related crash history. With respect to the data inputs, pivot tables specifications are used to exclude unrelated crashes. While the first two models make use of the entire accident history data set, the data of the right-turn analysis is summarized in the table below. In order to maintain an adequate sample size, the right-turn collision analysis will be conducted on collisions of all severities.

<table>
<thead>
<tr>
<th>Detail</th>
<th>Range of Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collision Type Name</td>
<td>[Property Damage, Injury, Fatal]</td>
</tr>
<tr>
<td>Collision Cause Name</td>
<td>[Followed too closely, yield sign violation, stop sign violation]</td>
</tr>
<tr>
<td>Driving Lane Primary Object</td>
<td>[Right turn bay, right turn lane, unspecified]</td>
</tr>
</tbody>
</table>

Table 5-2. Collision Data Inclusion for Right-Turn Analysis

The selection of the inclusion criteria reflects the primary cause and the lane which are relevant to a right-turn collision. The model for this analysis is univariate, so it will not be separated for different severities. However, this model will test using random parameters for each pairing. This means that the regression coefficients are allowed to vary for each treatment group. The model, denoted as PLNJ-RP (RT), is as follows.

\[
\ln(\mu_{it}) = \beta_{p(i),0} + \beta_{p(i),1} T_i + \beta_{p(i),2} t + \beta_{p(i),3}(t - t_0) I_{it} + \beta_{p(i),4} T_i t + \beta_{p(i),5} I_{it} + \beta_{p(i),6} \ln(V_{1it}) + \beta_{p(i),7} \ln(V_{2it})
\]

The only difference in the model is the addition of the \( p(i) \) subscript which denotes which treatment group the regression coefficient belongs to.
5.3. Prior Distribution Specifications

As mentioned previously, one of the crucial steps of the FB methodology is the specification of the prior distributions when estimating for unknown parameters. As mentioned previously, the most common used priors are diffused normal distributions (with zero mean and a large variance) for regression parameters and a Wishart($P, R$) prior for $\Sigma^{-1}$, where $P$ and $r \geq K$ represents the prior guess at the order and the magnitude of the precision matrix $\Sigma^{-1}$ and the degrees of freedom, respectively. Small values of $r$ correspond to vaguer prior distributions and it is recommended to choose $r = K$ (Li, Carriquiry et al., 2008, El-Basyouny and Sayed 2011). In this study, the following priors were used:

- $\beta_j \sim N(0, 100^2)$ for Regression Parameters
- $\Sigma^{-1} \sim \text{Wishart}(I, 2)$ for the Precision matrix, where $I$ is the 2x2 identify matrix

5.4. WinBUGS

The FB method will be implemented in one step using the WinBUGS statistical software package, which stands for Bayesian Inference Using Gibbs Sampling in WINdows. It employs a Markov Chain Monte Carlo method for simulation. It is an open-source development freeware developed by a team of researchers at MRC Biostatistics Unit at Cambridge University London. The latest version of the software was released in 2007 and it is no longer under development. However, it is still prominently and extensively used in research for various different studies, including transportation road evaluations and disease mapping.

Several key steps are required before simulation can be performed in WinBUGS. A more detailed example of the necessary steps for the implementation can be found in Appendix B.

1. The first step is to import the particular model form into the software’s integrated text editor. It is written in a computer language which is unique to the WinBUGS software. An in depth tutorial on coding in WinBUGS can be found in Spiegelhalter’s WinBUGS User Manual (2003).

2. WinBUGS can run codes from a given text file and also load the required data from the same text file. The accident history data and traffic volume data are also imported into software. The dataset, formatted for importing to the WinBUGS software, can be found in the Excel file “Excel Template_Yearly_SevPDO.xlsx” in Appendix A. After the data is imported, the text
file is saved as an “.odc” file, which is a file format that is readable by WinBUGS for further use. A compilation of the “.odc” file can be found in Appendix B.

3. The next step is to read components of the text file into the software itself so that simulation can be performed. This includes instructing the software about where to read the model, the data, as well as the initial values. This is found under the “Specification Tools Window.”

4. WinBUGS will provide the outputs of the odds ratio and regression coefficients in a seamless integration. The desired outputs of the simulation are specified beforehand within the “Sample Monitor Tool Window.”

5. At this point, the program is ready for simulation, and this is performed in the “Update Tool Window.” It should be noted that a quality simulation will first allow for the program to burn-in before taking samples. Typical studies using WinBUGS uses a minimum of 10 000 iterations, with half of them used for burn-in purposes. However, with current levels of technology, 100 000 burn-in iterations and 100 000 simulations should take less than an hour. This is the specification of iterations for this study. The simulations are run separately so that the DIC statistic, a measure of goodness of fit for such models, can be reset before the program begins to take samples.

6. After the burn-in and simulation iterations have been performed, the specified outputs of the program are extracted and saved.

5.5. Markov Chain Convergence

With simulation-based Bayesian inference techniques, one of the important steps before making inferences about the results is to check the simulation results have converged; in other words, whether the Markov Chain has reached a stationary state. This should be performed for all parameters, not only the ones of interest. There are usually two approaches to testing for convergence. One is the visual analysis of the trace plots which are generated from the simulation. Consider the following three example trace plots. Figure 5-1 below is an example of a perfect convergence for the simulation of parameter \textit{gamma}. The key is to look for a consistent average value with relatively small fluctuations.
Figure 5-1. Example of Trace Plot of Perfect Convergence (Source: SAS, 2012)

Figure 5-2 below illustrates a simulation where the initial value began remotely far away from the average, but converges to a stable value after a certain number of iterations. This differs from the previous example where the simulation began relatively near the final value. In simulation-based Bayesian Inference, this is precisely the reason why burn-in iterations are necessary.

Figure 5-2. Example Trace Plot of Convergence but Remote Initial State (Source: SAS, 2012)
Finally, Figure 5-3 is a trace plot exemplifying no evidence of convergence. At this point, it is important to consider either reparamaterizing the model or to run the simulation for a much larger number of iterations.

![Figure 5-3. Example Trace Plot of Non Convergence (Source: SAS, 2012)](image)

A number of statistical tests are also available for assessment of convergence (Cowles and Carlin, 1996; Brooks and Roberts, 1999). It should be noted, however, that while there are various diagnostic tests, none of them are foolproof; each of them only suggest the possibility of convergence. A commonly-used test is known as the Brooks-Gelman-Rubin (BGR) diagnostic, where a BGR statistic of less than 1.2 indicates convergence. Another common test is monitoring the ratio of the Monte Carlo error relative to standard deviation. The rule of thumb is that a ratio of less than 1-5% is highly suggestive of convergence.

### 5.6. Model Results

The primary outputs of the models are the odds ratios, both for the entire sample set and for each individual treatment site. Instead of reporting a point estimate as in conventional methods, Bayesian inference methods report mean values along with certain percentile values that reflect the distribution of the results. The level of confidence for this study is set at 95%; the specification of a confidence level reflects the fact that statistical inferences are estimates, and that the outputs are irrelevant if the required level of statistical confidence needed to accept or reject the results is not set beforehand.
To easily understand the concept of confidence interval, consider an imaginary evaluation study which reports an odds ratio with a mean of 0.8. The value itself indicates that there is a reduction effect. However, without considering the spread the data, the value of 0.8 cannot be taken as significant. In research, a 95% confidence level is fairly significant. In this case, the 2.5th percentile and the 97.5th percentile of the OR are also reported, to represent 95% of the data between these two outer limits. If the range between these two values include 1 (i.e. 0.7 – 1.1), then the range of 0.8 is concluded as insignificant at the 95% confidence level. Conversely, if the range does not include 1 (i.e. 0.7-0.9), then the value is accepted as significant. Choosing a higher confidence level may lead to higher confidence in results, but at the expense of excluding results due to overly stringent criteria.

The results of the first two models are summarized in the table below. Only the overall OR, which represents the average across the five treated locations, are displayed in the table.

<table>
<thead>
<tr>
<th>Model</th>
<th>Severity</th>
<th>OR Mean</th>
<th>Treatment Effect</th>
<th>OR 2.5th Percentile</th>
<th>OR 97.5th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVPLNI</td>
<td>1</td>
<td>0.7234</td>
<td>-28%</td>
<td>0.5804</td>
<td>0.8866</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.9484</td>
<td>-5%</td>
<td>0.6136</td>
<td>1.3960</td>
</tr>
<tr>
<td>MVPLNIJ</td>
<td>1</td>
<td>0.6392</td>
<td>-36%</td>
<td>0.5020</td>
<td>0.7965</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.8416</td>
<td>-15%</td>
<td>0.5112</td>
<td>1.2880</td>
</tr>
</tbody>
</table>

Table 5-3. Summary of Results of Models

The OR for PDO collisions was significant for both models. On the other hand, the OR for Sev collisions was not found to be specific for either model. A more in depth discussion of the results will ensue in the following Chapter. The results of the univariate right-turn only model are summarized in the table below.

<table>
<thead>
<tr>
<th>Treatment Location</th>
<th>OR Mean</th>
<th>Treatment Effect (%)</th>
<th>OR 2.5th Percentile</th>
<th>OR 97.5th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.078</td>
<td>+8%</td>
<td>0.518</td>
<td>2.028</td>
</tr>
<tr>
<td>2</td>
<td>0.509</td>
<td>-49%</td>
<td>0.262</td>
<td>0.868</td>
</tr>
<tr>
<td>3</td>
<td>0.476</td>
<td>-52%</td>
<td>0.262</td>
<td>0.777</td>
</tr>
<tr>
<td>4</td>
<td>0.479</td>
<td>-52%</td>
<td>0.154</td>
<td>0.959</td>
</tr>
<tr>
<td>5</td>
<td>0.852</td>
<td>-15%</td>
<td>0.325</td>
<td>1.738</td>
</tr>
<tr>
<td>Overall</td>
<td>0.609</td>
<td>-39%</td>
<td>0.408</td>
<td>0.851</td>
</tr>
</tbody>
</table>

Table 5-4. Summary of Results for Right-Turn Analysis
The right turn analysis found a greater treatment effect on right turn collisions compared to total collisions, which indicated a degree of success for targeting specific crash types. Reductions were found for four of the locations, while an insignificant increase was found for one of them.

As mentioned before, other outputs of the simulation include the individual odds ratio for each location for both crash severity, as well as the various regression coefficients which can be used to reconstruct the SPF’s which were developed by the simulation. These can be found in Appendix D. It should be noted that within the procedures for the simulation, the two (major and minor) traffic volumes were standardized (by subtracting the mean) to speed convergence. The standardization would affect the intercepts, but not the regression parameter estimates. The intercepts can easily be computed by reversing the standardization (adding the mean).

It should be noted that the estimates within the model all reached convergence through: (i) visual inspection of the trace plots, (ii) diagnosis of the BGR being less than 1.2, and (iii) diagnosis of the Monte Carlo error to SD ratio being less than 0.05. This step is critical in determining whether the inference results are useful.

5.7. Model Comparison

The models will mainly be compared with one another mainly based on their goodness of fit to the data. The particular statistic used for comparison is known as the Deviance Information Criterion (DIC). A more in-depth discussion of the DIC can be found in the Appendix C. With the DIC, the lower value indicates a better fit to the data. The DIC is fairly sensitive, so that a difference in DIC between models of 5 to 10 reflects a significant difference in goodness of fit, while a difference of greater than 10 will most definitely rule out the model with the higher DIC. The DIC for the three models are given in the table below.

<table>
<thead>
<tr>
<th>Model</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVPLNI</td>
<td>1261</td>
</tr>
<tr>
<td>MVPLNIJ</td>
<td>1252</td>
</tr>
<tr>
<td>PLNIJ-RP (RT)</td>
<td>625</td>
</tr>
</tbody>
</table>

Table 5-5. Summary of DIC
The inclusion of the jump factor significantly improved the fit of the total collisions model, reflected by a reduction of DIC of 9. The difference is large enough that the MVPLNIJ may confidently be considered superior to MVPLNI with respect to fit. The DIC of the PLNIJ-RP (RT) model cannot be directly compared to that of the other two models, but the value does suggest that despite a smaller sample size, the PLNIJ-RP (RT) is quite useful for modeling specific accident types.
6. RESULTS DISCUSSION

In this chapter, the results of the models will be discussed in greater depth. The implications of the results on the success of the program will also be illustrated.

Given the better fit, the results of the MVPLNJ model will be used for discussion of the treatment evaluations. The adjusted percentage change for PDO and Severe collisions for each of the five treated locations are shown in Table 6-1 and Figure 6-1.

<table>
<thead>
<tr>
<th>ID</th>
<th>Group</th>
<th>Primary Road</th>
<th>Secondary Road</th>
<th>% Change in PDO (Total Collisions Types)</th>
<th>% Change in Sev (Total Collisions Types)</th>
<th>% Change in RT-Collisions (All Severities)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Ellerslie Road</td>
<td>91st Street</td>
<td>-19%</td>
<td>+1%</td>
<td>+8%</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Yellowhead Trail</td>
<td>Victoria Trail</td>
<td>-33%*</td>
<td>-14%</td>
<td>-49%*</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>87th Avenue</td>
<td>170th Street</td>
<td>-44%*</td>
<td>-25%</td>
<td>-52%*</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>Yellowhead Trail</td>
<td>St. Albert Trail (WB Right Turns)</td>
<td>-43%*</td>
<td>-25%</td>
<td>-52%*</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>Yellowhead Trail</td>
<td>St. Albert Trail (EB Right Turns)</td>
<td>-37%*</td>
<td>-8%</td>
<td>-15%*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Overall</td>
<td>-36%*</td>
<td>-15%</td>
<td>-39%*</td>
</tr>
</tbody>
</table>

Table 6-1. Summary of Evaluation Results (MVPLNJ Model)

* indicates change is significant at 95% confidence level
The following treatment effects were observed from the analysis:

- Change in PDO collisions ranging from a reduction of 19% to 44%
- Change in RT collisions ranging from a reduction 8% of 52%
- Overall reduction in PDO, Sev, and RT collisions of 36%, 15%, and 39%, respectively

These results can be contrasted to the preliminary study results from Chapter 4. In the BA with Comparison Group study, two of the treated sites actually registered an increase in severe and right-turn collisions. With respect to accounting for confounding factors, the main difference between the FB method and the aforementioned method is a more extensive assessment of exposure effects (through the use of collision models) and the compensation of regression-to-the-mean bias with the use of statistical inference and probability analysis. In comparison to the results in Chapter 4, these two key differences in methodology reported a reduction in collision frequency across all 5 treated sites, as opposed to just three.

Overall, the safety program has demonstrated significant reductions in collisions, particularly for property damage only collisions (for total collisions) and right-turn collisions. Statistically, the reductions for the property damage only and right-turn collisions were found to be significant at the 95% confidence
level. Site 5, 8 and 11 saw the greatest effect of treatment, with PDO-collision reductions of over 35% and RT collisions of over 40%. Site 14, which is the intersection of the Yellowhead Trail and St. Albert Trail (EB Right Turns Only) saw a modest reduction in collisions in all three areas.

Site 1, the intersection of Ellerslie Road and 91st Street, saw a small (19%) reduction, which was also not found to be significant at the 95% confidence level. It also saw an insignificant increase in severe and right-turn collisions. Overall, statistically speaking, no significant change in collision frequency due to treatment was observed at this site. The results of this site can be compared to that of Site 8, which also underwent the same countermeasure. A detailed analysis of the implementation at the site is warranted to determine possible causes of the lack of response to treatment.

Several potential explanations were identified. First of all, the particular implementation at this location was not selected based on a particularly high collision rate, but rather on the opportunity to showcase a new right turn design. While this reduces the likelihood of RTTM bias, it may also reduce the relative effectiveness of the treatment given that the location was not particularly hazardous to begin with. A look at the traffic collision data indicate that from 2006 to 2008 (within the before-period for this intersection), only two rear-end collisions occurred, which is the primary target of this implementation (Southbound Right-Turn modification), indicating that the location was already fairly safe to begin with. However, it should be noted that the traffic volume adds a slightly different perspective. For this study, detailed intersection traffic volume was only available for 2011, with the exception of Ellerslie Road and 91st Street. The Southbound right-turning and through traffic nearly doubled between 2008 and 2011, likely attributed to development in the surrounding areas. In comparison, within the methodology, the total volume and forecasting analysis for traffic volume only computed a difference of 10% in total approach volumes between the two years. This underestimation of the traffic volume growth can have a profound effect on the computation of the treatment effectiveness; the increase in exposure for the right-turn collision was not adequately accounted for, possibly reducing the apparent significance and effect of any improvement.

These observations suggest two ideas. The first is the importance of maximizing available traffic volume in order to accurately capture changes in exposure. It indicates that traffic volume can be very sensitive to
very localized changes, such that broad estimation and forecasting techniques may not always be favorable. The second is the recognition that when looking at traffic safety improvements of a specific type, it may be more useful to separate traffic volumes based on movements that are applicable to the targeted collisions rather than simply aggregates of approach volumes or total volumes. The idea is that changes in applicable traffic volumes should be more greatly emphasized.


7. MONTHLY INTERVAL ANALYSIS

In simpler methods such as naïve BA or BA with comparison study, BA studies are carried out by comparing either total collision frequency or average collision frequency across multiple years in the before- and after- period. For example, 3 years of before-period accident history may be consolidated into either total collision frequency or average collision rates.

The introduction of the FB methodology and the ability to model for temporal or time-related trends opens new avenues for different study periods. The models introduced in this thesis allows for separate time-trends in both the before- and after- period, as well as a time-trend for the treatment effect. The inclusion of time-trend capabilities recognizes that traffic collisions exhibit trends, and that the information stored within these trends can be utilized to obtain better data fit and develop more refined models.

The downside of looking at smaller time intervals is that collision totals for shorter time intervals are much more susceptible to random fluctuations. The random fluctuations will severely compromise the statistical validity of the results. As evident in literature, researchers have attained a fair amount of success with yearly intervals when employing FB methodologies. The purpose of this chapter is to test whether it is, one, feasible and, two, beneficial (statistically) to use monthly intervals. The potential benefit of using monthly intervals is the increase in individual periods, such that a greater depth of collision trends can be modeled for greater fit.

The two models considered for this study is identical to that for the yearly analysis, namely the MVPLNI and MVPLNIJ model. The only difference is that the data is inputted in the algorithm in a month by month format. With this change, for each location, 7 years of collision data then begins 84 individual months of data. The dataset used for the yearly analysis can be found in the “Excel Template Yearly_SevPDO.xlsx” file in the Appendix. It should quickly be noted that from a computational point of view, the monthly interval analysis takes a substantially longer time for WinBUGS to run because of the excess of data points. The results of the study are shown in the table below; the results of the yearly analysis are also included for comparison purposes. Only the overall program effectiveness is reported; the detailed statistical output can be found in the file “Outputs.xlsx”.


Table 7-1. Summary of Monthly Analysis Contrasted with Yearly Analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>Severity</th>
<th>OR Mean</th>
<th>Effect</th>
<th>OR 2.5&lt;sup&gt;th&lt;/sup&gt; Perc.</th>
<th>OR 97.5&lt;sup&gt;th&lt;/sup&gt; Perc.</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVPLNII, Monthly</td>
<td>1</td>
<td>0.8829</td>
<td>-12%</td>
<td>0.7351</td>
<td>1.051</td>
<td>5853</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.132</td>
<td>+13%</td>
<td>0.7942</td>
<td>1.561</td>
<td></td>
</tr>
<tr>
<td>MVPLNIIJ, Monthly</td>
<td>1</td>
<td>0.7978</td>
<td>-20%</td>
<td>0.6468</td>
<td>0.9704</td>
<td>5852</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.191</td>
<td>+19%</td>
<td>0.8013</td>
<td>1.691</td>
<td></td>
</tr>
<tr>
<td>MVPLNII, Yearly</td>
<td>1</td>
<td>0.7234</td>
<td>-28%</td>
<td>0.5804</td>
<td>0.8866</td>
<td>1261</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.9424</td>
<td>-5%</td>
<td>0.6136</td>
<td>1.396</td>
<td></td>
</tr>
<tr>
<td>MVPLNIIJ, Yearly</td>
<td>1</td>
<td>0.6392</td>
<td>-36%</td>
<td>0.502</td>
<td>0.7965</td>
<td>1252</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.8416</td>
<td>-15%</td>
<td>0.5112</td>
<td>1.288</td>
<td></td>
</tr>
</tbody>
</table>

The DIC of the monthly model is significantly larger than that of the yearly, indicating a poorer fit. At the 95% confidence level (indicated by the 2.5<sup>th</sup> and the 97.5<sup>th</sup> percentile of the OR), the results of the two monthly models did not see a significant reduction in either the PDO or Sev collisions, with the exception of PDO collisions under the MVPLNII model.

The results suggest that with the current models, monthly interval analysis displays no advantages over yearly analysis from either a data fit point of view or results point of view. Two potential shortcomings are identified:

- **Short Time Interval Leads to Low Collision Counts:** The short time interval led to severely low collision counts in some locations for particular months, as evident in the dataset. Many intervals did not report collisions at all. This makes the data much more susceptible to random fluctuations.

- **Lack of Additional Model Parameters to Account for Monthly Analysis:** One of the purported benefits of utilizing monthly analysis was to extract information on monthly trends in order to improve fit. The current model allows for linear time trends throughout the entire study periods. This form of time trends may work well for yearly analysis. However, fluctuations in monthly collision frequencies may be cyclical throughout the year (i.e. seasonal effects), as opposed to linear throughout the study period. In this light, it may be more beneficial to add additional binary parameters to indicate the season which each data point belongs to, or sinusoidal parameters that are able to model for smooth cyclical trends throughout the year.
Given the aforementioned shortcomings and results, it appears that with the current complexity of models, yearly intervals are still favorable because of greater data stability (i.e. less fluctuations) in time interval totals.
8. SUMMARy AND CONCLUSION

8.1. Objective 1: Demonstration of FB Methodology

The state-of-the-art Full Bayes methodology, which was previously reserved mainly for research purposes, was successfully demonstrated on the dataset provided by the City of Edmonton to evaluate the treatment effectiveness of five treated sites. The favorable results have led to the proposal of a standardized methodology for future traffic evaluations, shown in the table below. Lists of steps for implementing Full Bayes methodology are available in literature, but they focus only on very high-level tasks, do not address the data collection and preparation stages, and also do not include provisions for differing types of evaluation or data availability (Park and Park and Lomax, 2010). It should be noted that Steps I and II do not actually fall within the evaluation process, but are crucial and necessary. Considerations and recommendations for each step are given.

<table>
<thead>
<tr>
<th>STEP</th>
<th>TASK</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>• Select treatment sites&lt;br&gt; • Select comparison sites</td>
<td>• Treatment sites are typically selected either based on high collision history, through a formal BSM process, or through manual selection. One approach is to choose comparison sites specifically for its similarity to a particular site. Another approach is to simply have a large pool of comparison sites, and one comparison site may fall within the multiple treatment groups if they are applicable.</td>
</tr>
<tr>
<td>II</td>
<td>• Collect collision data&lt;br&gt; • Collect traffic volume data&lt;br&gt;   o Regular interval (monthly/annually)&lt;br&gt;   o One interval in both the before- and after- period&lt;br&gt;   o One single interval within the study period&lt;br&gt; • Collect location trait data</td>
<td>• The collection of data occurs both before the implementation and after the implementation. Typically the minimum duration for both periods is at least 2 years. Collision data should be collected to the degree of detail that is of interest (e.g. If one wishes to evaluate changes in collisions of different severities, the distinction must be made at the data collection stage). There are various practices of traffic volume data collection, so a number of the options are listed in descending order to favourability.</td>
</tr>
<tr>
<td>I</td>
<td>• Select the length of the intervals within the study period</td>
<td>• The length of the intervals is typically annually, but in recent literature there have been researchers who have proposed the use of monthly intervals. However, with low collision counts and the lack of additional information such as seasonal effect, annual time intervals are still recommended.</td>
</tr>
</tbody>
</table>

Table 8-1. Standardized Practical Guideline for the Full Bayes Inference Technique
<table>
<thead>
<tr>
<th>STEP</th>
<th>TASK</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>• Organize collision data</td>
<td>• Most collision recording practice involves recording each collision as a separate entry. The collisions need to be consolidated for each time interval. Separation and filtering of collisions may be needed depend on the level of detail or type of collisions of interest.</td>
</tr>
<tr>
<td>3</td>
<td>• Select measures of volume data</td>
<td>• Before processing the volume data, the first step is to determine what measures of volume data (i.e. measures of exposure) is desired. It should be emphasized that at this step, only the measure of exposure is selected, not the functional form of the exposure to level of safety relationship. Typically volume data/exposure is measured in AADT; for volume data where less than a day’s worth of volume is collected, the results are usually extrapolated.</td>
</tr>
<tr>
<td></td>
<td>o Total volume AADT</td>
<td>• The selection of the measures of exposure is dependent on the available level of detail, the type of collisions of interest, and other considerations.</td>
</tr>
<tr>
<td></td>
<td>o Approach volume (major/minor)</td>
<td>• This methodology requires exposure measures for each time interval. Ideally, actual traffic data for each time interval is available. Missing data should be extrapolated using any of the prescribed method. For treatments that are known to significantly alter exposure (e.g. conversion of through lane to turning lane), at least one instance of traffic volume before and after the treatment is necessary.</td>
</tr>
<tr>
<td></td>
<td>o Specific movement volumes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Etc</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Organize volume data (extrapolate as necessary)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Extrapolate based on municipal population changes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Extrapolate based on changes in number of registered vehicle</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Extrapolate based on changes in local traffic volumes at nearby sites</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>• Select functional form and specify model</td>
<td>• A list of common exposure functional forms is listed in Table 5-1. The specification of the model depends on other location-specific traits that are to be included in the model, such as the presence or absence of particular road features. A list of other common features or inclusions to the model specific is shown. A multivariate approach is recommended because separation of collisions by severities allows for a more comprehensive subsequent economical evaluation and the multivariate process accounts for correlation and improves parameter estimate accuracy over a univariate approach. Jump feature is also recommended because of the suggested instantaneous change in level of safety following a treatment. Treatment-time effects and general collision trend effect features may be included if the study period is long enough to warrant such an inclusion.</td>
</tr>
<tr>
<td></td>
<td>o Univariate vs. Multivariate (separated by collision severity)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Random parameters</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Jump effect feature</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Treatment-time effect (linear or non-linear)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o General collision trend effects</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o Spatial Correlation</td>
<td></td>
</tr>
</tbody>
</table>

Table 8-1. Standardized Practical Guideline for the Full Bayes Inference Technique
<table>
<thead>
<tr>
<th>STEP</th>
<th>TASK</th>
<th>COMMENTS</th>
</tr>
</thead>
</table>
| 5    | • Prior specifications | • The priors for the FB methodology have to be specified before the simulation. A more in depth discussion can be found in Section 5.3.  
• Uninformative priors (zero mean with large variance) may be used in the absence of any prior knowledge. |
| 6    | • Specify level of confidence | • The level of confidence (statistical) needs to be specified before the simulation. This is because any statistical inference has a level of uncertainty, and the results are of little use if the required level of statistical significance is not selected beforehand.  
• Typically, a confidence level of at least 95% (significance level less than 5%) is required. |
| 7    | • Perform FB simulation to estimate parameters  
  ○ Enter the specified model form into the software  
  ○ Input collision, traffic, and trait data in a form that is readable by the software | • FB simulation can be performed using a number of statistical software and packages. One commonly used one is WINBugs, which was also the software used for this study. The number of iterations required is dependent on the computational power available and the complexity of the model; with modern technology, even with modest model complexity, iterations in the order of magnitude of 100 000’s should not be difficult.  
• A certain portion (sometimes half) of the initial iterations should be discarded as burn-in. Both the parameters and the measures of goodness-of-fit should be estimated from the remaining iterations. |
| 8    | • Visual inspection of trace plots for convergence  
  • Statistical tests for convergence  
  ○ $\text{BGR} < 1.2$  
  ○ $\frac{\text{MCEr}}{\text{S.D.}} < 0.05$ | • Test for convergence is required for all variables, not only those of interests (e.g. odds ratio). Although no method is conclusive, a visual inspect of trace plots and two statistical tests are sufficient to suggest convergence. Two commonly used statistical tests are indicated.  
• If convergence is not met, the number of simulation iterations may need to be increased to reach convergence, or the model parameters may need to be changed. |

Table 8-I. Standardized Practical Guideline for the Full Bayes Inference Technique
<table>
<thead>
<tr>
<th>STEP</th>
<th>TASK</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>• Test for significance of the odds ratios</td>
<td>• Test for significance is performed by extracting certain percentile values of the parameter estimate distributions. Typically a two-tailed test for significance is used. For a confidence level of 95%, the 2.5th and the 97.5th percentiles of each odds ratio estimate is generated. An odds ratio is deemed significant if the range between the two percentile values do not overlap with 1, and vice versa. It should be noted that failure of the significance test does not conclude that the resulting change is non-existent.</td>
</tr>
<tr>
<td>10</td>
<td>• Extract the mean value of the odds ratios (and other parameters of interest)</td>
<td>• The mean value of the odds ratio and other parameters of interest are extracted from the simulation results. The effective reduction in collisions for any particular site is simply computed as ((1 - OR)). The SPF’s can also be regenerated using the original model specification and the SPF regression estimates.</td>
</tr>
<tr>
<td>X</td>
<td>• Subsequent analysis</td>
<td>• For locations that did not exhibit strong (or significant results), further detailed analysis may be warranted to determine the cause. • The results can be added to a database to compile CMF (Collision Modification Factor) estimates to predict the effects of future treatments of similar type on similar sites. • The results can also be used for a subsequent economical evaluation of the treatments or treatment program. Each agency should have estimates of cost for collisions of different severities. When supplemented with treatment costs, various economical evaluations (eg. B/C ratio, Net present value, etc) may be performed.</td>
</tr>
</tbody>
</table>

Table 8-1. Standardized Practical Guideline for the Full Bayes Inference Technique

The above proposed methodology can effectively account for the previously mentioned confounding factors in traffic treatment evaluations. It is also flexible in data requirements, inclusion and exclusion of various traits, as well as overall model specifications.

8.2. Objective 2: Evaluation of Safety Initiative

The results of the analysis have concluded that the recent safety initiative to target right-turn collisions has been successful in reducing collision prevalence significantly. These results arise from the implemented Full Bayes methodology utilizing the MVPLNIJ model. Reductions of property damage
only collisions were as high as 44% and as high as 52% for right-turn collisions. These reductions also withstand technical scrutiny with their statistical significance at a 95% confidence level. The table below illustrates the overall effectiveness of the program.

<table>
<thead>
<tr>
<th>% Change in PDO (Total Collision Types)</th>
<th>% Change in Sev (Total Collisions Types)</th>
<th>% Change in RT (All Severities)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-36%</td>
<td>-15%</td>
<td>-39%</td>
</tr>
</tbody>
</table>

Table 8-2. Summary of Program Effectiveness

8.3. Limitations of Further Research

It is acknowledged that this study has several limitations and also large potential for future research. The limitations are such:

- **Lack of Comparison to EB Methodology:** The EB method represents the current state-of-the-art method that is used in practice. This report abridges research and practice to introduce the FB method as the next state-of-the-art methodology. As such, it is desirable to have a direct comparison between the two methods using the dataset. However, because of the absence of a reference population, an EB analysis cannot be performed.

- **Lack of After-Period Intervals:** One of the potential weaknesses of this analysis was the lack of adequate after-period intervals, which can threaten the validity of the analysis. Fortunately, despite the limitation, the results withstood technical scrutiny and have found significant results at a 95% confidence level.

Areas for further research include:

- **Exploring Temporal Effects of Traffic Safety Benefits:** The possibilities of the FB method are certainly not limited to what has been demonstrated in this report. Multiple after-period intervals could not only develop a time-dependent profile for treatment effects, but it could also explore a recent hypothesis that traffic safety benefits incur over time in a non-linear manner (El-Basyouny and Sayed, 2012).

- **Crash Migration:** A multivariate analysis across different crash types could further investigate on the concept of internal crash migration.
• **Movement volumes for Specific Crash Types**: Given the high level of detail available within the traffic data, when evaluating for changes in the prevalence of a particular crash type, the exposure could be represented with the specific relevant movement volumes.

In conclusion, this thesis has successfully demonstrated that the FB methodology is ready and feasible to be adopted into standard practice. It displays clear advantages over the current state-of-the-art, the EB methodology. To name a few, the FB methodology produces more realistic estimates of treatment effects while reducing the cost of carrying out evaluations. The developed methodology was applied onto a traffic improvement program in the city of Edmonton, and reported an overall success in reducing traffic collision frequencies. Finally, the thesis was able to identify the strengths of certain features of characteristics, such that a standardized approach for implementing the FB methodology has been outlined.
BIBLIOGRAPHY


APPENDIX A – USING PIVOT TABLES IN MICROSOFT EXCEL

Pivot tables are powerful tools for data summarization, and the particular version of it within Microsoft Excel is fairly simple to use. This is an abbreviated tutorial for some of the main steps necessary to get the tool working for the purpose of this report.

In the figure above, the typical format of the raw data is shown. As evident, each collision is a single entry with various characteristics. Pivot table will allow for simple calculations or organizations of different combinations of the variables. This tutorial will demonstrate how to sum up the total number of collisions occurring for each month, in each location, and further separated by two severities (Property Damage Only & Injury or Fatality).

In Excel 2007 and 2010, the Pivot Table function can be found in the “Insert Tab” at the top. By default, clicking the button should automatically highlight the data; accepting the data range should bring up the PivotTable Field List in a separate spreadsheet.
The “PivotTable Field List” is the core of the tool and is shown in the figure above; initially the Column Labels and Row Labels field are empty. The first field box outlines the available variables for separation. For this study, we wish to separate the collision data on a monthly basis by columns, with the locations on the rows. To accomplish this, we simply drag the variables from the first field into the Column Labels and Row Labels fields.

For the columns, first we want to draw the variable Collision Report Year into the columns field. This will separate the data by collision years. Since we want to further subdivide the data by months, we also drag the variable Collision Report Month into the field, **underneath** the first variable. This will nest the collision reported month under its corresponding year.

For the rows, first we want to separate the data based on their locations. We drag the variable Collision Location Name into the rows field. We want to further subdivide the collisions based on the severity, so we also drag the variable Collision Type **underneath** the first variable.
The last step is to drag a variable into the “Values” field in order to count up the total numbers. Since we are only concerned with the number of entries (counts) and not with the sum or average of any particular value, we can simply drag any unused variable into the box. In this example, we dragged the variable **Count of Collision Cause Name**. Right-clicking and selecting “Value Field Settings” will bring up the window shown in the figure below.

![Value Field Settings Window](image)

This box allows the users to specify what is the actual value that he or she wants to see. For example, it can allow the users to easily see the sum or average of all the values that fall within any particular location and time categorization (i.e. Location 1 for Month 4 of year 3). In this study we are only concerned with the count, so selecting “Count” should suffice.
The figure above is a sample of the output of the pivot table. The columns are first separated by the reported year, and then by the reported month. The rows are separated first by the locations, and then by the collision severity. This is all the necessary steps for using a PivotTable for this study. The last thing to add is after the PivotTable is created, the fields can be changed at any time; that means that variables can be added or removed from any field. Moreover, it is possible to hide certain portions of the data. For example, in the figure above, the collision severities of property damage, injury, and fatal are all shown. If the user wishes to only look at fatal collisions, he or she can select the button at the top of the column, which will bring up the menu shown in the figure below.
Simply deselecting the any of the severity will hide it; this selection is reversible, meaning that one can hide or unhide any portion of the data at any time.
APPENDIX B – RUNNING WINBUGS SIMULATION

Assuming that the code has already been compiled, the first step is to open up the “.odc” file.

The first step is to complete the specifications for running the model. Under the Model menu, select “Specifications” to pull down the Specifications Tool windows. At each step of the specification, one can refer to the text line at the bottom of the program to ensure that there are no errors in the code.

Highlighting the word “model”, the next step within the code, press the button “check model.” Then, indicate the number of different changes in the textbox; the number of chains is equal to the number of dependent variables that are being modeled. For example, a univariate study looking at all accidents will have 1 chain, while a multivariate study looking at accidents of 3 levels of severity will have 3 chains.

Under the headings “Data 1”, highlight the word “list” and press the “load data” button. In the “Data 2” section, highlight all the variables (excluding the actual data), and press the “load data” button.
After loading both data, select the “Compile” button to compile the model. At this point, the model has now been compiled; however, before the simulation can run, initial values must be given. A similar procedure applies for loading initial values. Under the “init” heading, select the “list” for the first set of initial values, and press the “Load Inits” button. Repeat this for the second set of initial values, and press the “Gen Inits” button to generate the initial values. The model is now for simulation.

Before running the simulation, the next step is to first specify the outputs of the analysis. This includes extracting the odds ratio, standard errors, regression coefficients, etc. Under the menu “Inference”, select the option “Samples.”

In the field “node” enter all the relevant variables to extract from the simulation. This includes the overall odds ratio “OR”, the individual odds ratio for each of the different locations “ORi”, and applicable regression coefficients, such as b0, b1, b2, etc.

In the “Model” menu, select the options “Update” to pull up the Update Tool windows. This is the step that will give instructions for the model to begin running simulations.
Without going into too much detail, the “Updates” field indicate the desired number of iterations; the “Refresh” field indicates how often the iteration indicator should update. Typical studies using WinBUGS use a minimum of 10 000 iterations, with half of them used for a burn-in period. However, with current advancements in computational technology, 100 000 burn-in iterations and 100 000 simulation iterations should take less than an hour. For this study, all models run through 100 000 burn-in and 100 000 simulation iterations. The two sets of iterations are run separately, so that a particular statistic, known as Deviance Information Criterion (DIC), can be reset after the burn-in period.

After running the model for the burn-in period, the DIC Tool can be found under the “Inference” menu. Selecting the “set” button will instruct the program to start calculating the DIC for the subsequent iterations.

After the DIC Tool has been set, the 100 000 simulation iterations can be carried out. After the update, it is then time to extract the results from the simulation. The variable data is found in the Sample Monitor Tool used earlier. On the right hand field, the various percentiles of each variable can be indicated for display. Entering an asterisk (*) into the nodes field and pressing the statistics button will instruct the program to generate a results table for all the variables. A sample output is shown in the figure below.
Before discussing the implications of each of the variables, the DIC statistics should also be extracted first. This can be done by pulling up the DIC Tool again, and selecting “Stats.” This will create a second window with the DIC outputs.

For convenience, these outputs can be copied and pasted from this window into the samples output and saved as a single file. The data is now ready for interpretation. Instead of giving a single value for each variable, statistical methods often give a number of different attributes which describe the overall distribution of the variables. These attributes usually include:

- **Mean:** This indicates the average value of the particular variable. This is usually the attribute of interest, but should be taken with consideration to other attributes which describe the “spread” of the data.
- **Sd:** This stands for the standard deviation of the data.
- **valXXpc:** This value represents the specific value of the variable at the particular percentile XX. For example, for a given variable, the attribute val2.5pc would indicate the value of that variable at the 2.5th percentile of the distribution. Percentile values are useful for testing the significance of a variable.

Statistical inference will always have an allowed margin of error, usually represented by the confidence level. The aforementioned attributes are combined to interpret whether the particular findings are significant or not. For example, imagine a variable which has a mean value of 0.9. If the chosen confidence level is 90%, then the percentile values of interest are the 5th and 95th. If the range of these two values pass through 1, then the variable is considered insignificant at that confidence level. A result of 0.9 (0.8 – 0.95) would indicate significance, while a result of 0.9 (0.8 – 1.05) would not. It should be noted that these values are all in reference to the *chosen* confidence level. Choosing a higher confidence level will naturally lead to higher confidence in results, at the risk of excluding variables due to insignificance. In practice,
typically a minimum confidence level of 90% is selected; when representing data from statistical inference, it is always necessary to indicate not only the mean, but also the range (the percentile values), and also the particular confidence level. A discussion of the different variables is as follows.

- **OR**: This stands for the odds ratio; a simpler way of explaining odds ratio is that \((1 - \text{odds ratio})\) represents the actual percentage reduction of accidents. For example, an odds ratio of 0.9 would mean that the actual reduction is \((1 - 0.9) = 0.1\), which represents a 10% reduction in accidents. For a multivariate study, a bracketed number will indicate the particular variable. For example, for a study that looks at PDO and SEV accidents, then \(\text{OR}[1]\) and \(\text{OR}[2]\) will express the odds ratio for PDO and SEV accidents, respectively.

- **ORi**: The differentiation between \(\text{OR}\) and \(\text{ORi}\) is that while \(\text{OR}\) represents the average \(\text{OR}\) across all locations, \(\text{ORi}\) indicates the odds ratio for each location \(i\). \(\text{ORi}[1,2]\) would represent the odds ratio for treatment site 1 and SEV accident, while \(\text{ORi}[3,1]\) would represent the odds ratio for treatment site 3 and PDO accidents.

- **\(b_0, b_1, \text{etc}\)**: The remaining variables are the regression coefficients which are generated from the simulation. Recall that Bayesian inference techniques depend on safety performance functions to estimate the expected accident rates at locations. Empirical Bayes techniques require that SPF’s are already formulated, while the Full Bayes methodology integrates it into one step. Extracting these regression coefficients and matching with the specified model form for the analysis will allow one to construct the Safety Performance Function that was generated from the simulation.

Thus, the critical step for the evaluation is to extract the averages for the odds ratio (overall and for each location), along with their percentiles in order to test for significance. Interpretation of these results are used to conclude whether the traffic safety intervention was successful or not, and to what degree.
APPENDIX C – DEVIANCE INFORMATION CRITEREON

Deviance information criterion (DIC) is a hierarchical modeling generalization of the Akaike information criterion (AIC). It is fairly useful in Bayesian model analysis, and has a tendency to penalize for model complexity. It represents a goodness-of-fit of the proposed model to the dataset.

\[ D(\theta) = -2 \log(p(y|\theta)) + C \]

Where:

- \( D(\theta) \) = Deviance
- \( Y \) = Data
- \( \Theta \) = Unknown parameters of the model
- \( p(Y|\theta) \) = Likelihood function
- \( C \) = Constant

In computation of its value, a lower DIC is more favorable. For comparison purposes of two different models, a difference in DIC of more than 5 suggests significant difference in the goodness of fit, while a difference of more than 10 is enough to rule out the model with the higher DIC.
## APPENDIX D – DATA FILES

### Volume and Dataset Files

<table>
<thead>
<tr>
<th>File Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume Data Consolidation.xlsx</td>
<td>Consolidation of major and minor AADT traffic for locations, including forecasted estimates using population and total registered vehicles.</td>
</tr>
<tr>
<td>Volume Adjustment.xlsx</td>
<td>Forecasted estimates for AADT using Traffic Growth Factors estimated from nearby sites.</td>
</tr>
<tr>
<td>Excel Template_Yearly_SevPdo.xlsx</td>
<td>Data template for yearly analysis, ready to be imported into WinBUGS</td>
</tr>
</tbody>
</table>

### Model and Output Files

<table>
<thead>
<tr>
<th>File Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mvplni_yearly.odc</td>
<td>WinBUGS compatible file, contains model and data of MultiVariate Poisson-Lognormal Intervention Model</td>
</tr>
<tr>
<td>mvplni_yearly_out.odc</td>
<td>Contains the outputs of the said model</td>
</tr>
<tr>
<td>mvplniJ_yearly.odc</td>
<td>WinBUGS compatible file, contains model and data of MultiVariate Poisson-Lognormal Intervention w/ Jump Model</td>
</tr>
<tr>
<td>mvplniJ_yearly_out.odc</td>
<td>Contains the outputs of the said model</td>
</tr>
<tr>
<td>rtrn_uni_pln_J-RP_yearly.odc</td>
<td>WinBUGS compatible file, contains model and data of UniVariate Poisson-Lognormal Intervention Model for Right-Turn collisions</td>
</tr>
<tr>
<td>rtrn_uni_pln_J-RP_yearly_out.odc</td>
<td>Contains the outputs of the said model</td>
</tr>
<tr>
<td>Output.xlsx</td>
<td>Consolidation of Output Files</td>
</tr>
</tbody>
</table>