CONNECTICUT TRANSPORTATION INSTITUTE

ESTIMATING BENEFITS FROM SPECIFIC HIGHWAY SAFETY IMPROVEMENTS: PHASE I – FEASIBILITY STUDY

Final Report

MAY 1999

JHR 99-268 Project 97-1

SCHOOL OF ENGINEERING
UNIVERSITY OF CONNECTICUT
STORRS, CONNECTICUT

ESTIMATING BENEFITS FROM SPECIFIC HIGHWAY SAFETY IMPROVEMENTS: PHASE I – FEASIBILITY STUDY

Final Report

MAY 1999

JHR 99-268 Project 97-1

Principal Investigators:

Fei Yuan, Graduate Assistant
John Ivan, Assistant Professor
Norman Garrick, Associate Professor
Christian Davis, Professor

This research was sponsored by the Joint Highway Research Advisory Council (JHRAC) of the University of Connecticut and the Connecticut Department of Transportation and was carried out at the Connecticut Transportation Institute of the University of Connecticut.

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

Technical Report Documentation Page

1. Report No.	2.	3. Recipient's Catalog No.
JHR 99-268	PB99-147647	N/A

4. Title and Subtitle
Estimating Benefits from Specific Safety Improvements: Phase I: Feasibility Study

5. Report Date
May 1999

6. Performing Organization Code
N/A

7. Author(s)
Fei, Yuan, John Ivan, Christian Davis, and Norman Garrick

8. Performing Organization Report No.
JHR 99-268

9. Performing Organization Name and Address
University of Connecticut
Department of Civil and Environmental Engineering
191 Auditorium Road, Box U-37
Storrs, CT 06269

10 Work Unit No. (TRAIS)
N/A

11. Contract or Grant No.
N/A

12. Sponsoring Agency Name and Address
Connecticut Department of Transportation
280 West Street
Rocky Hill, CT 06067-0207

13. Type of Report and Period Covered
Final

14. Sponsoring Agency Code
N/A

15 Supplementary Notes
N/A

16. Abstract

One task of a traffic safety engineer is to identify high accident locations and to design treatments for reducing the number of crashes. This process relies on the availability of accurate information on the crash reduction efficacy of the various treatments. Currently, most agencies rely for their prediction on crash reduction factors that were developed in the 1960's. Given the revolution in vehicles and highway facilities since that time we feel that a re-assessment of these factors is long overdue.

Currently we are conducting a before/after type study to develop crash reduction factors for modern conditions on two-lane rural highways. This paper reports on the results of the first phase of this study, which was aimed at determining the feasibility of collecting the required information and developing a procedure for evaluating improvements. The issues addressed at this phase were an assessment of the availability and reliability of data and development of a comprehensive procedure for the overall study, including suitable analytical methodologies. In this paper we report on the site selection process, the data sources (including use of the Connecticut Department of Transportation Photolog) and an analysis of before and after data for several sites treated. Two different methods of analysis were investigated - confidence intervals and likelihood functions. We feel that both methods have merit for quantifying the magnitude of the accident reduction factors.

17. Key Words	18. Distribution Statement
highway safety, before/after study, roadway realignment, two-lane highways, and photolog	No restrictions. This document is available to the public through the National Techical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	35	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimetres	mm
ft	feet	0.305	metres	m
yd	yards	0.914	metres	m
mi	miles	1.61	kilometres	km
		AREA		
in^2	square inches	645.2	millimetres squared	mm^2
ft^2	square feet	0.093	metres squared	m^2
yd^2	square yards	0.836	metres squared	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	kilometres squared	km^2
		VOLUME		
fl oz	fluid ounces	29.57	mililitres	mL
gal	gallons	3.785	Litres	L
ft^3	cubic feet	0.028	metres cubed	m^3
yd^3	cubic yards	0.765	metres cubed	m^3

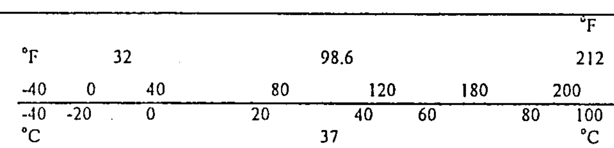
NOTE: Volumes greater than 1000 L shall be shown in m^3

Symbol	When You Know	Multiply By	To Find	Symbol
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9	Celcius temperature	°C

* SI is the symbol for the International System of Measurement

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimetres	0.039	inches	in
m	metres	3.28	feet	ft
m	metres	1.09	yards	yd
km	kilometres	0.621	miles	mi
		AREA		
mm^2	millimetres squared	0.0016	square inches	in^2
m^2	metres squared	10.764	square feet	ft^2
ha	hectares	2.47	acres	ac
km^2	kilometres squared	0.386	square miles	mi^2
		VOLUME		
mL	millilitres	0.034	fluid ounces	fl oz
L	litres	0.264	gallons	gal
m^3	metres cubed	35.315	cubic feet	ft^3
m^3	metres cubed	1.308	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams	1.102	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C+32	Fahrenheit temperature	°F



ACKNOWLEDGMENTS

This research was supported by the Joint Highway Research Advisory Council of the University of Connecticut and the Connecticut Department of Transportation and was conducted at the Connecticut Transportation Institute.

The authors appreciate assistance provided by Sylvia Zajac in the data collection process, and thank Nelson Bernardo and Chunyan Wang for additional contributions.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Tables	vi
List of Figures	vii
1. Introduction	1
2. Literature Review	2
3. Procedure for Before-and-After Study	5
4. Sample Application	7
5. Results	12
6. Conclusions	24
7. References	26

LIST OF TABLES

Table 1	Study Intersections	8
Table 2	Number of Crashes	9
Table 3	Crash Reduction Factors for Approach Angle Realignment	18
Table 4	Crash Reduction Factors for Horizontal Curve Realignment	18
Table 5	Time Spent on Phase 1 Tasks	25

LIST OF FIGURES

Figure 1	Confidence Interval Procedure	6
Figure 2	Likelihood Function Procedure	7
Figure 3	Site 2 Crash by Type	12
Figure 4	Site 2 Crash Severity	13
Figure 5	Site 4 Crash by Type	13
Figure 6	Site 4 Crash Severity	14
Figure 7	Site 3 Crash by Type	14
Figure 8	Site 3 Crash by Location	15
Figure 9	Site 3 Crash Severity	15
Figure 10	Site 5 Crash by Type	16
Figure 11	Site 5 Crash by Location	16
Figure 12	Site 5 Crash Severity	17
Figure 13	Total Crash Reduction Likelihood Functions for Approach Angle Realignment	20
Figure 14	Target Crash Reduction Likelihood Functions for Approach Angle Realignment	20
Figure 15	Injury Reduction Likelihood Functions for Approach Angle Realignment	21
Figure 16	Property Damage Reduction Likelihood Functions for Approach Angle Realignment	21
Figure 17	Total Crash Reduction Likelihood Functions for Horizontal Curve	22

Realignment

Figure 18 Target Crash Reduction Likelihood Functions for Horizontal Curve Realignment 22

Figure 19 Injury Reduction Likelihood Functions for Horizontal Curve Realignment 23

Figure 20 Property Damage Reduction Likelihood Functions for Horizontal Curve Realignment 23

1. INTRODUCTION

In the United States, over 40,000 people are killed and more than 3.5 million are injured each year as a result of motor vehicle crashes on the nation's highways [1]. In fact, more people have been killed in traffic crashes than in all the wars in which this nation has been involved [2]. Overall, crash costs represent about fourteen percent of the total costs of highway travel, creating a loss ratio that most business and industrial activities could not support. Additionally, the social and economic costs are high since sixty percent of those who died and nearly seventy percent of those injured in traffic crashes are in the highly productive fifteen to forty-five-year old age group [3].

A routine exercise in traffic safety engineering is to identify locations that experience high crash rates and apply treatments such as increasing lane and shoulder widths to reduce the number of crashes. The most widely known procedure for identifying and eliminating hazardous locations was proposed by Laughland *et al.* [4]. The system consists of the following six steps:

1. Hazardous location evaluation
2. Alternative improvements selection
3. Alternative improvements evaluation
4. Improvements programming and implementation
5. Implemented improvements evaluation
6. Highway safety program evaluation

Within a given budget, the traffic engineer must identify the hazardous locations and select the improvement that will yield the greatest benefit-cost ratio. Each proposed improvement must first be identified and evaluated to ensure that it has a sufficiently high likelihood of reducing the number and the severity of crashes so as to result in an overall economic benefit. The key steps in the procedure are to identify crash causation, and predict the reduction in crashes through highway improvements. Much of the predictive process is based on a study using data (much of questionable validity) dating back to 1947 [5]. Clearly, there is a need to update the predictions of crash reduction rates using new data.

This report documents the work completed in Phase 1 of a Joint Highway Research Advisory Council project titled "Estimating Benefits from Specific Highway Safety Improvements." The overall objective of this project is to update the prediction procedure. The Phase 1 feasibility study focused on determining whether or not the required data were available, where to get them, and how much processing was required to prepare them for analysis. In this phase an appropriate methodology and procedure for conducting a study of the effectiveness of crash reduction methods was developed.

In the first section of this report, a literature review outlines several proven methods for calculating and evaluating crash reduction factors. Next, the procedure for conducting the before-and-after analysis is presented. Then this analysis procedure is applied to a small data sample. Finally, conclusions about findings and experience with the procedure are discussed, along with proposed plans for Phase 2.

2. LITERATURE REVIEW

A crash reduction factor quantifies the effectiveness of an improvement designed to reduce either the frequency or severity of crashes at a location. (Note that Federal highway safety agencies prefer the term "crash," because "accident" implies an unavoidable event.) Crash reduction factors are widely used because the concept is simple. Before-and-after study and cross-sectional analysis are the two methods often used to develop crash reduction factors. In before-and-after studies, the safety effect of an improvement is determined by comparing the number of crashes occurring before the improvement with the number of crashes occurring after the improvement. This project is designed to be a before-and-after study.

The simple before-and-after method is based on the observed number of accidents. The most frequently discussed problem associated with this kind of study is the regression to mean effect. The first part of this section will focus on different methods of reducing the bias due to this effect. Another problem associated with this type of study is determining how to conduct a conclusive statistical experiment for the analysis. In the second part of this section, methods of statistical inference will be discussed. Then, a procedure is outlined for conducting the study.

A. Methods of Reducing the Regression to Mean Effect

The regression-to-mean effect, or bias-by-selection, is a phenomenon whose existence has been known about for many years. The observed accident frequencies before a treatment are not good estimators of the long-term mean accident rate at a site: they are usually biased upwards [6]. This is because normally a location is selected to implement improvements because it has a higher crash frequency than other similar sites. These locations with high crash frequency for the before periods thus tend to have higher reductions in crash frequency after improvements, since even without the treatments, the crash frequency would probably reduce due to natural fluctuations. Hence, the observed change in accidents will tend to overestimate the true benefits of a treatment. Danielsson has mathematically proven that the expected regression effects are the same for all types of crash [7].

There are different methods for mitigating the regression to mean effect. The idea is to estimate the number of crashes expected to be recorded during the after period had the treatment not been implemented. This expected number of crashes for the after period without the improvement is compared with the observed number of crashes with the improvement for the same period. The reduction factor is then calculated from this difference. The problem then becomes that of obtaining an estimate of the expected number of crashes for no improvement.

One approach is to use matched-control-group methods that involve a classical experimental design [8, 9]. The changes in crash rates at the treated sites are compared with those for a carefully matched control group. The distribution of observed crash frequencies between sites in the control group should be similar to that for the treatment group, and the physical characteristics of the sites in the two groups should also be similar. Crash data for both before and after periods of the control group are required. In principle, this type of method avoids the regression-to-mean effect completely and the problem of bias does not arise. However, this method has some practical problems because it is often difficult to find a sufficient number of similar sites left without treatment.

Many researchers have suggested empirical estimation methods, which assume a particular form of distribution of mean crash rates between sites [8-13]. This kind of analysis assumes that traffic crashes at any particular location in the absence of any highway improvement fit the Poisson distribution. The expected number of crashes is a random variable with a gamma probability distribution over the population of a number of sites, and the expected crash rate is a random variable with a gamma probability distribution.

This approach is often called the Empirical Bayesian method. It is distinguished from other statistical methods by the fact that any parameter in a problem (such as the true crash rate at a location) is regarded as the value of a random variable having a probability distribution. Hauer presented some discussion and the derivation of formulas to utilize this method for estimating the expected number of crashes in the after period in a before and after study [8-10]. This method does not require crash data in the after periods for the control cases.

In Hauer's study, '*m*' is defined as the expected number of crashes at a location, and the actual count of crashes which is subject to random variation is denoted by '*x*'. The actual crash count should be treated like one observation from a random variable because of natural fluctuations. The distribution of *m*'s in a group of sites can be described by a gamma probability distribution function. With this in mind, one can estimate the expected number of crashes for a treated site, and compare this estimator with the observed after count to get the crash reduction factor, thus mitigating the regression-to-mean effect.

The following formula is used to estimate m for a site at which the observed crash count is x [8, 9]:

$$\varepsilon = x + \left[E\{m\}/(VAR\{m\} + E\{m\})\right] \times \left[E\{m\} - x\right] \qquad (1)$$

$$= \alpha E\{m\} + (1 - \alpha)x \qquad (2)$$

Where
ε is the estimator of m for an intersection that recorded x crashes,
x is the crash count,
$E\{m\}$ is the expected value of m,
$VAR\{m\}$ is the variance of m, and
α is defined by the following expression

$$\alpha = (1 + VAR\{m\}/E\{m\})^{-1} \qquad (3)$$

By its measured causal factors, the site belongs to a population of sites in which the *m*'s have a mean *E{m}* and variance *VAR{m}*. The equation shows that for a given site, the estimated true expected crash count is equal to a weighted sum of the observed crash count (x) and the mean crash count for the group to which the intersection belongs (*E{m}*). The weight (α) of *E{m}* is always a number between 0 and 1. When the *VAR{m}* is much bigger than *E{m}*, α will be very small and $\varepsilon \approx X$. That is, little can be learned from the fact that a site belongs to the indicated group because sites in that group have widely differing crash counts. Conversely, when *VAR{m}* is much smaller than *E{m}* the weight $1 - \alpha$ will be very small and $\varepsilon \approx E\{m\}$. In this case, little weight attaches to x, which is given to random fluctuations, and one should rely mainly on the fact that sites in this group all have very similar *m*'s.

Methods for estimating *E{m}* and *VAR{m}* are provided by Hauer for populations having a gamma distribution [10]. The expected value of the crash counts in the reference population is the same as the expected value of the expected number of crashes during a specified time period in the reference population. The variance of the crash counts in the reference population is the sum of the expected value of the crash counts during a specified time period in the reference population and their variance. The following equations are used to calculate *E{m}* and *VAR{m}*. The derivation can be found in [10].

$$\hat{E}\{m\} = \overline{X} \tag{4}$$

$$VA\hat{R}\{m\} = s^2 - \overline{X} \tag{5}$$

B. Methods of Statistical Inference

Point estimation is one of the most common methods used to make inferences about crash reduction factors. In these types of studies, statistical hypothesis tests are used to make inferences about the safety effect of the countermeasures studied. In real-life, studies are conducted with relatively small samples and deal with countermeasures of which the effect is typically small. The built-in conservatism in the hypothesis test usually tends to return the answer that the hypothesis "no effect" cannot be rejected. In addition, whether the null hypothesis is rejected or not depends largely on the chosen level of significance, which is typically chosen by convention.

One alternative to hypothesis testing is to use point estimates with confidence intervals [14]. The confidence interval is a useful index of the degree of uncertainty in the estimator. The distribution of the expected crash reduction factors at different locations is assumed to be normally distributed and the probability that the expected crash reduction factor falls in a certain confidence interval is given. As the body of experimental sites grows, this confidence interval becomes narrower.

Another method preferred by Hauer is the likelihood method [15]. The likelihood function identifies the most likely value of crash reduction and presents the uncertainty surrounding it in an intuitively clear fashion. It preserves in condensed form all that can be extracted from a data set. It also represents a structured process for the accumulation of information and for learning from experience. When new data become available, the corresponding likelihood function is used to update reduction factors and to create a new state of knowledge. So, the likelihood function can facilitate the use of formal decision analysis and can be an essential ingredient for making coherent decisions.

3. PROCEDURE FOR BEFORE AND AFTER STUDY

The procedure for conducting a before-and-after study is summarized as follows:
Step 1 Site Selection
Step 2 Data Collection and Preparation
Step 3 Crash Frequency Estimation
Step 4 Comparison and Statistical Inference

Step 1 - Site Selection

Study sites should be selected such that there is a sufficient number of years (*i.e.*, two) of crash data available both before and after the construction period. In addition, the site selection process should be consistent with statistical sampling theory.

Step 2 - Data Collection and Preparation

Here, geometric features as well as traffic volume and crash history data at the study sites are collected. Sites with the same kind of highway improvement and similar geometric characteristics may be classified into the same group. For each site, the crash data are reviewed for aberrations such as the presence of a time trend in the data; such trends must be removed before analysis. The crashes are then aggregated into groups according to the type, location or cause of the crash. This permits the analysis to consider crash reduction effects for different crash groups separately, as well as for all crashes on an aggregate basis. Then we can predict the effectiveness of a specific highway improvement for reducing specific crash types. The output of the second step will be the data set that can be directly used for the analysis in Step 3.

Step 3 - Crash Frequency Estimation

Crash frequencies vary by factors other than the improvement that was made, so we must control for these differences, and we must analyze intersections in groups with similar factors. Also, we are trying to get a good estimator of the crash rate without the improvement to see if the after rate is within the range. To estimate the expected crash frequencies, we need a sufficient number of control cases. Control sites can be used to estimate the mean crash counts and their variance for the study sites, and obtain a better estimator of the expected number of crashes for the before period, thus mitigating the regression-to-mean effect. To select appropriate control cases, we must find the important site characteristics that affect the crash frequencies, and use them to define the study and the control groups. Previous studies suggest that factors affecting the number of crashes occurring at an intersection include time trends and traffic demand, urban or rural environment, type of traffic control, frequency of access points, speed limit, shoulder width, median type and width, lighting level, availability of left turn bays, number of legs, and number of traffic lanes [16,17,18]. It was determined that environment (urban versus rural), traffic control (signal, sign), and geometry (number of legs) were among the most important factors to be controlled [16], so these features define the analysis groups and the control groups.

Crash history data for the control cases are required to estimate the crash frequency for a certain type of site. Different methods of estimation have different data requirements. The output of this step is the expected crash frequencies for the treated sites if the improvements had not been done.

Step 4 - Comparison and Statistical Inference

In step 4, we compare the predicted accident frequency from the previous step with the observed after count, and compute the accident reduction factor. Either point estimation with confidence intervals or likelihood functions can be used to make the statistical inferences. In the first method, we assume the mean reduction factors for a specific improvement are normally distributed, then obtain the point estimator and compute its confidence interval on a certain level. The other method requires inputting the likelihood function and finding the most likely value of the crash reduction and its probability distribution. Figures 1 and 2 show the procedures that based on the above discussion for point estimator with confidence interval and the likelihood functions, which are outlined in greater detail in subsequent sections, respectively.

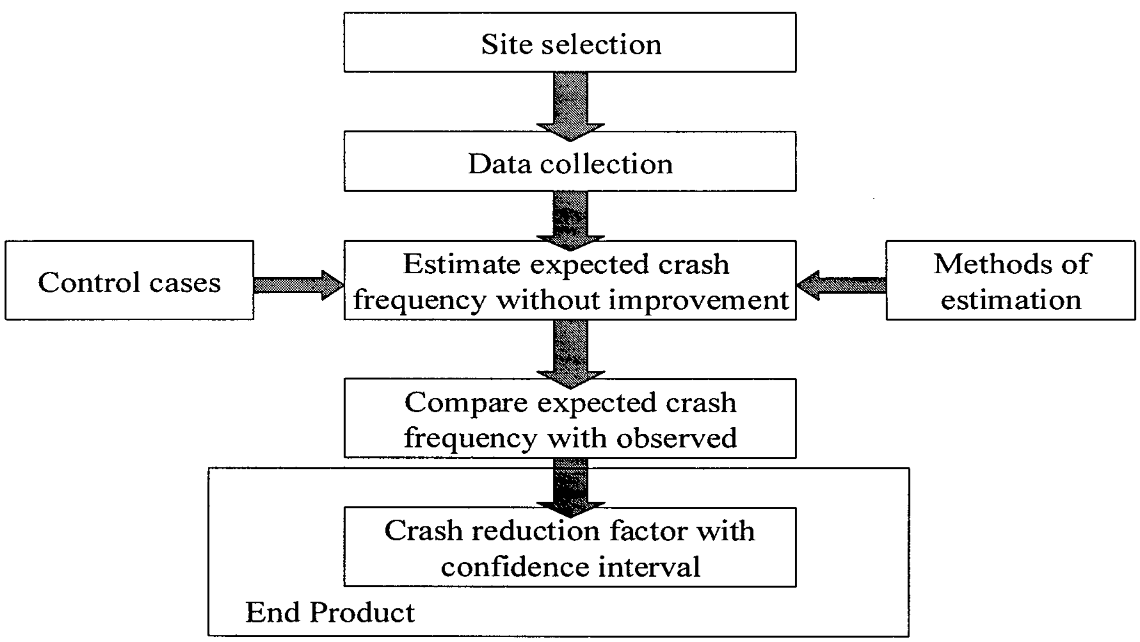


Figure 1. Confidence Interval Procedure

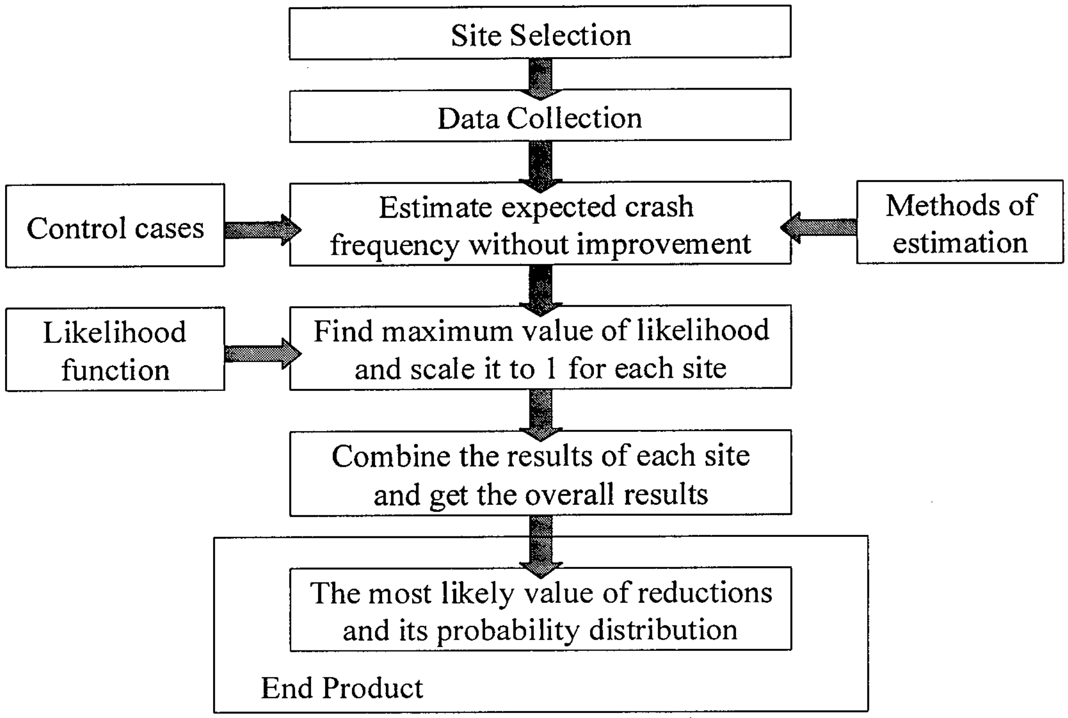


Figure 2. Likelihood Function Procedure

4. SAMPLE APPLICATION

This section describes a sample application of this procedure based on a small sample of intersection improvement sites.

A. Study Design

Focus of study. We studied a small number of highway locations at which specific crash reduction treatments had been applied. To focus the analysis, locations were restricted to rural, two-lane highways, which had been the subject of roadway realignment projects. These realignments fell into one of two groups: (1) realignment of skewed intersection approaches or (2) realignment of horizontal curves at intersections.

Site selection. The study sites were selected from the ConnDOT Pre-Construction Management System (PCMS) list of projects that have been implemented in recent years. The date that construction was completed and a short description of the improvement are also given in this list. The main standard for selection of the sites was the availability of a sufficient number of years of crash data before and after construction. The nine sites selected are listed in Table 1.

Table 1. Study Intersections

Site	Town	Location	Type of Improvement
1	Glastonbury	Rt. 17 & Main St.	realigned channelization, added left turn lane, and added traffic signal
2	Madison	Rt. 79 & SR 450	realigned approach angle and added left turn lane
3	Montville	Rt. 163 & Maple Ave.	realigned horizontal curve
4	New Canaan	Rt. 106 & Weed St.	realigned approach angle
5	New Milford	Rt. 7 & Candlewood Lake Rd.	realigned horizontal curve
6	Windsor	Rt. 159, Deerfield Rd. & Rood Ave.	realigned offset intersection and added left turn lane
7	Coventry	Rt. 44, Rt. 31 & Grant Hill Rd.	realigned approach angle and added left turn lane
8	Vernon	Rt. 31 & Rt. 30	realigned offset intersection and added left turn lane
9	Vernon	Rt. 83 & Rt. 74	realigned approach angle and added traffic signal

Define study area. The functional area of an intersection is larger than the actual physical area because crashes attributable to the intersection do not necessarily occur within the boundaries of the intersection. Logic suggests that the functional area should comprise the distance traveled during perception and reaction time, the distance required to come to a stop and any required storage length. Stover gives maneuver distances and total distances (maneuver plus perception and reaction distance but excluding storage) for different driving speeds [19]. These distances represent the minimum functional length of an approach to an intersection. In this project, all of the selected sites have posted speed limits between 30 and 45 mph, and the total distance required to finish the approaching intersection maneuver for a running speed of 45 mph is between 465 and 630 ft. Based on this, we used 0.1 mile (528 ft.) from each approach as the boundary of the area that would be affected by the intersection. All the geometric and crash data were collected for the area within this boundary.

Data collection. Crash data from 1989 to 1996 were obtained from the Connecticut Department of Transportation Accident Experience database. These records are extracted from original police reports and include all reported crashes that occurred on state maintained roads. They include information about crash participants, how the crash occurred, and the prevailing environmental and road conditions for each crash. We calculated the crash rate as crashes per million vehicles entering the intersection.

Most of the projects were performed in 1993 and 1994, so there were about five years of crash data for the before period and about two years of crash data for the after period. Crashes occurring during the construction period were excluded from the analysis. The number of crashes and the construction period for each study site are presented in Table 2. Crash variables include location of the crash, time of occurrence, collision type, severity, number of different types of participants, causal factors, weather, pavement conditions, and other relevant features of each crash.

Table 2. Number of crashes

Site	Number of Crashes		Exposure		Construction period	
	Before	After	Before (MV*/Day)	After (MV*/Day)	Beginning	Completion
2	40	14	12249	14200	10/11/93	6/23/94
3	19	3	4853	4900	4/1/93	7/5/94
4	53	2	12237	10800	4/14/94	3/17/95
5	40	8	11400	10801	5/3/93	10/7/94

* Million vehicles

Most of the geometric data were collected using the ConnDOT photolog archives. The photolog archives are a videodisc system containing images of the entire 6300 centerline kilometers (3900 miles) of the state-maintained highway network [20]. Important site characteristics that may affect crash rate were collected for a distance of 0.1 mile along each approach to each study intersection. These characteristics included lane and shoulder widths, number of roadside objects (buildings, utility poles, mailboxes, trees) within a certain distance from the traveled way, number of driveways by type, number of minor intersections, roadway lighting conditions, traffic control devices, number of approach legs, sight distances from the intersection, type and visibility of warning devices, and presence of guide-rail and their type and end treatment. The 1988 videodiscs were used to collect the before characteristics of intersections and the 1996 videodiscs were used to collect the after period data. For this feasibility study, the site characteristics collected from the photolog were assumed to be consistent over the five years of the before period. For the next phase of the project, the site conditions will be checked for each year to make sure no changes were made to the site during the relevant periods.

Some information such as intersecting angles, turning radius, and grades were not visible on the photolog. These were measured directly from the as-built plans. Site visits were conducted to collect information for intersecting minor roads not covered by the photolog system. In such cases, the before data are also not available from any source.

B. Analysis:
Two statistical models were considered for use in this project: (1) a before and after model using a control group and (2) a before and after model not using a control group. While a control group adds statistical robustness to the before and after statistical model, it was rejected for use in the feasibility phase because of the difficulty in matching the construction sites with a sufficient number of control locations of similar size and physical and traffic characteristics. Also there was a sufficient number of years of accident data for the before period, which can partly remove the regression to mean effect. This is because the longer the time period, the closer the crash count will be the true mean for the site in the long run.

Of the nine original sites selected, only four were considered for full-scale analysis. Sites 1 and 8 in Table 1 were discarded because each had conditions which were totally unique with respect to the others, so they could not be matched into groups. Site 6 was excluded because it was in a suburban rather than rural area. Site 9 was excluded due to the lack of sufficient number of years of before crash data. Site 7 was discarded because it was the only signal controlled site left. The remaining four intersections were classified into two groups according to their specific improvements: approach angle realignment (Site 2 and 4), and horizontal curve realignment (Site 3 and 5).

Crashes were classified in three different ways: by type, location, and severity. In the horizontal curve realignment group, crashes that occurred on the realigned curve were considered to be the target crash for the improvement. In the approach angle realignment group, multi-vehicle intersection crashes were the target crashes. The multi-vehicle intersection crashes were also divided into three groups (head-on, rear-end, and other) in order to evaluate the effect of the improvement on the different types of crash.

The crash reduction factors Φ_i were calculated using the formula:

$$\Phi_i = \frac{(\mathrm{N}_{\mathrm{B_i}}/\mathrm{V}_{\mathrm{B_i}}) - (\mathrm{N}_{\mathrm{A_i}}/\mathrm{V}_{\mathrm{A_i}})}{(\mathrm{N}_{\mathrm{B_i}}/\mathrm{V}_{\mathrm{B_i}})} \tag{6}$$

N_{B_i} and N_{A_i} are the number of crashes of a certain type (total, injury, etc.) before and after the improvement for site *i* respectively. V_{B_i} and V_{A_i} are defined as the traffic exposure (million-vehicles entering the intersection) of site *i* for the before and after period, respectively.

The likelihood functions for the reduction factors were also determined, and the results were compared with the point estimators. The likelihood function is of the following form [15]:

$$L(\theta) = \prod_{i=1}^{n} \theta^{N_{A_i}} \left[B_i + \alpha_i + (V_A/V_B)_i A_i \theta\right]^{-(N_{B_i}+\beta_i+N_{A_i})} \tag{7}$$

$$\theta = \frac{N_{Ai}}{N_{B_i}} \tag{8}$$

The parameters α_i and β_i are estimates given by

$$\alpha_i = \overline{N}_{B_i} \Big/ \left| S^2_{B_i} - \overline{N}_{B_i} \right| \tag{9}$$

$$\beta_i = \overline{N}^2_{B_i} \Big/ \left| S^2_{B_i} - \overline{N}_{B_i} \right| \tag{10}$$

Where

$\overline{N}_{B_i}$ is the sample mean of the number of before crashes for the group to which site i belongs,

$S^2_{B_i}$ is the sample variance of the number of before crashes for the group to which site i belongs,

$(V_A/V_B)_i$ is the ratio of exposure in the after to the before period for site i,

N_{B_i} is the observed number of crashes before for site i,

N_{A_i} is the observed number of crashes after for site i,

B_i is the number of years in the before period for site i,

A_i is the number of years in the after period for site i.

The variable θ serves here as the index of the safety effect. If a measure reduces the expected number of accidents to 90 percent of its previous value, then $\theta = 0.90$. If it causes an increase of 5 percent, then $\theta = 1.05$. In other words, the reduction factor is $1-\theta$. The $L(\theta)$ value is scaled between 0 and 1: the larger the value of $L(\theta)$, the more likely is the value of $1 - \theta$ to be the true reduction factor.

5. RESULTS

Figures 3 through 12 present comparisons of the before and after crash rates for each site by crash classification. Note that horizontal curve realignment reduces overall crash rates at both sites. The on-curve (target) crashes are reduced most in both cases. On the other hand, the off-curve or non-target crashes actually increased at the New Milford site. The photolog reveals that this site had driveways off the realigned curve and no driveways on the curve. Three crashes occurred at driveways off the realigned curve during the after period whereas there were no such crashes before. Thus the increase in off-curve crash rate is possibly explained by accident migration: straightening the curve increases the vehicle speed in the vicinity of the intersection and these driveways, so that the driveways became more dangerous. Non-visible injuries also increased at the New Milford site, but the overall injury severity was reduced.

For the approach angle realignment group, the total and multi-vehicle intersection (target) crash rates at both sites were reduced. In addition, the target crashes at both sites were reduced more than the total crashes. A greater reduction is observed at the New Canaan site. This may be because the approach angle at this site was more skewed in the before period.

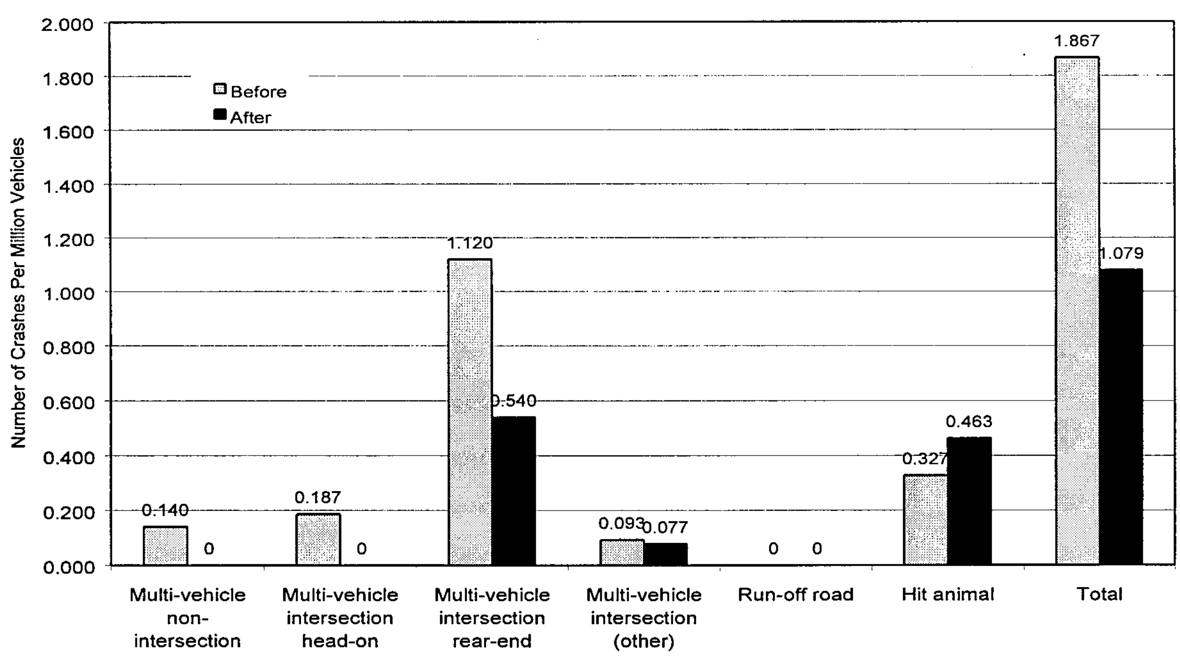


Figure 3. Site 2 (Approach Angle Realignment, Madison) Crashes by Type

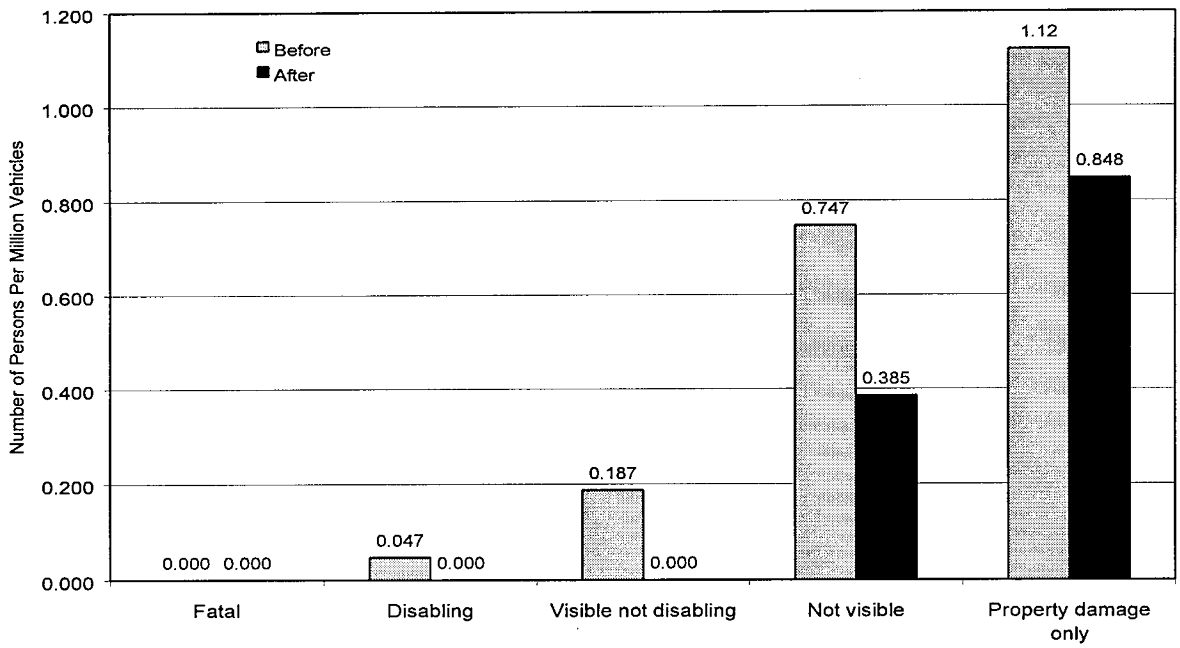


Figure 4. Site 2 (Approach Angle Realignment, Madison) Crash Severity

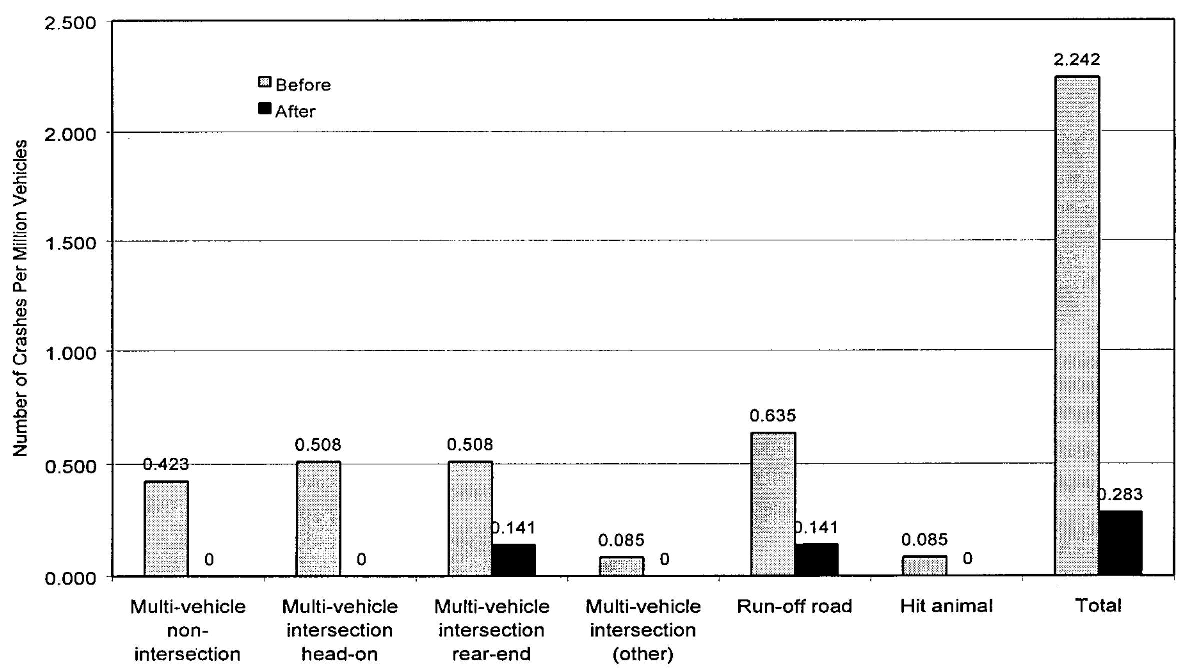


Figure 5. Site 4 (Approach Angle Realignment, New Canaan) Crashes by Type

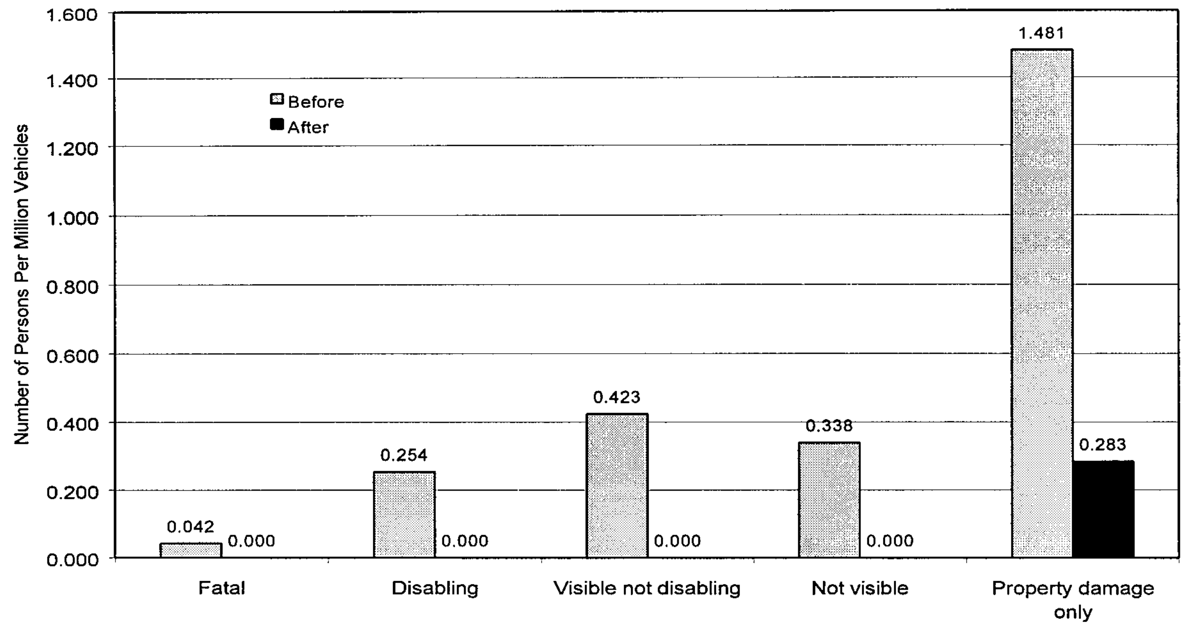


Figure 6. Site 4 (Approach Angle Realignment, New Canaan) Crash Severity

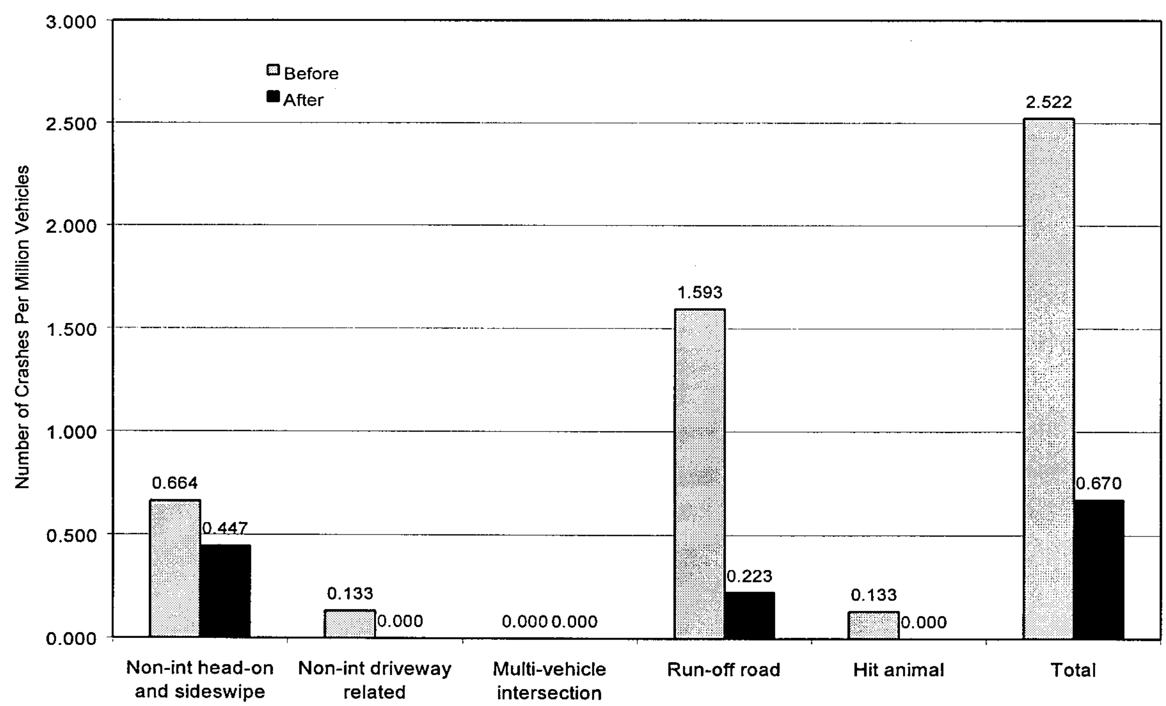


Figure 7. Site 3 (Horizontal Curve Realignment, Montville) Crashes by Type

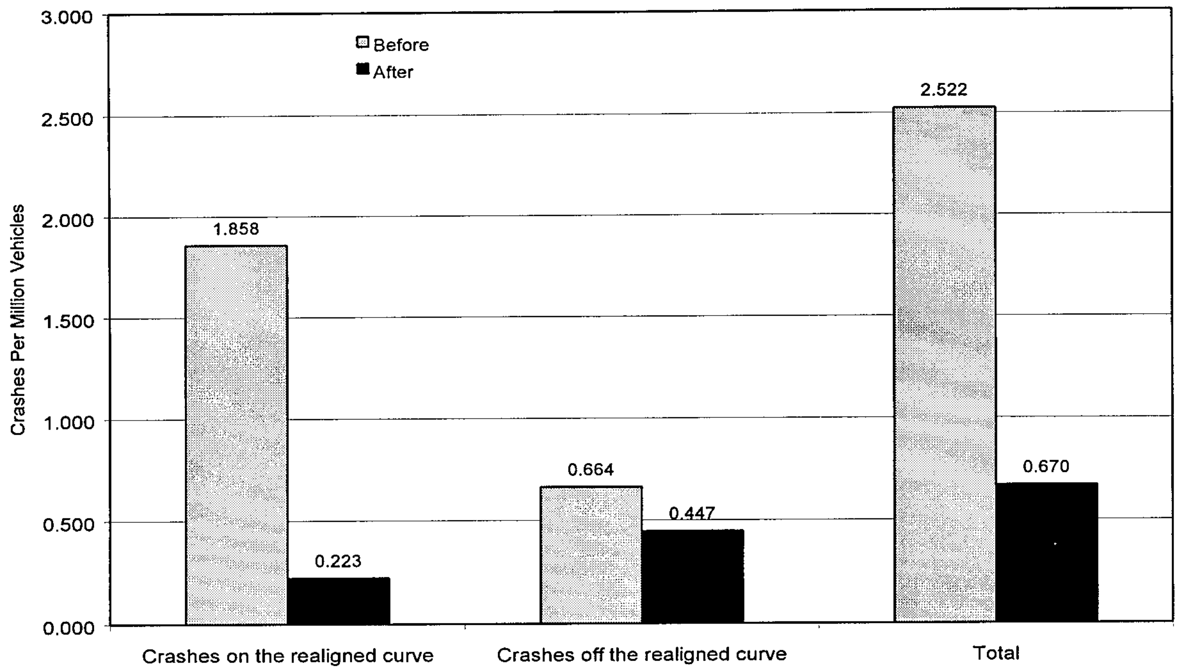


Figure 8. Site 3 (Horizontal Curve Realignment, Montville) Crashes by Location

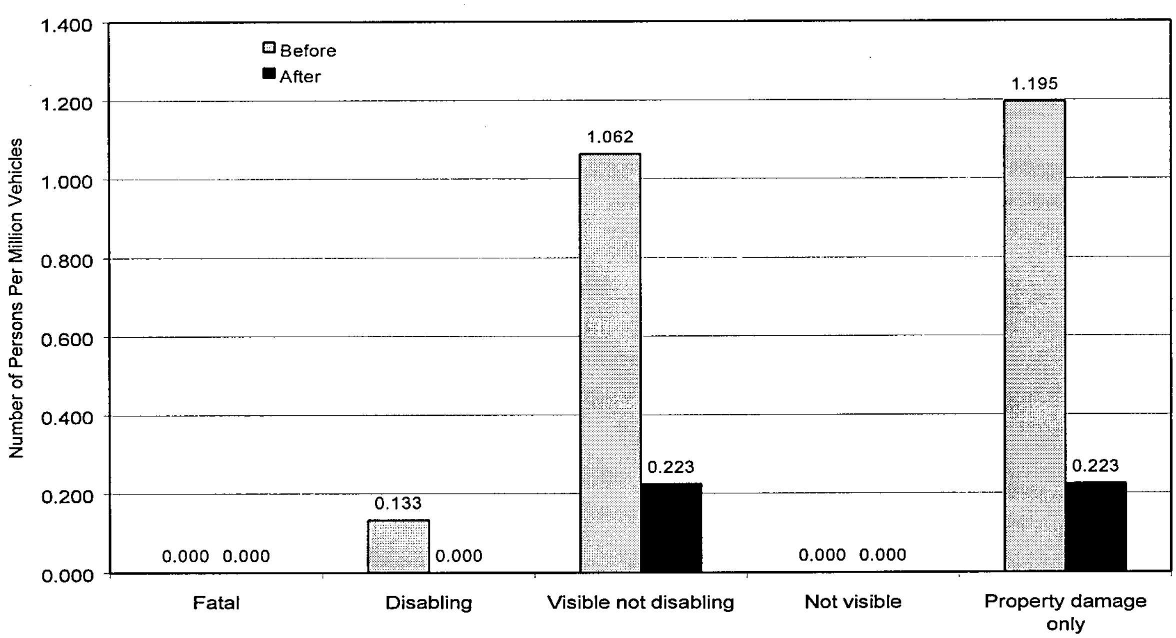


Figure 9. Site 3 (Horizontal Curve Realignment, Montville) Crash Severity

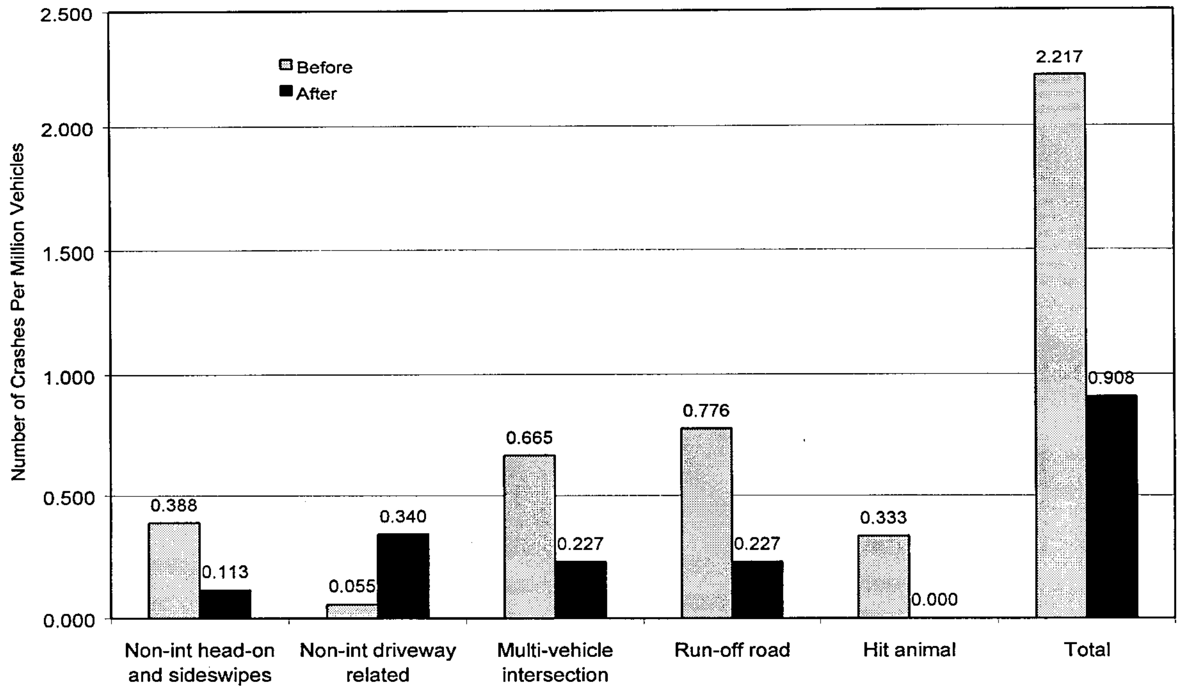


Figure 10. Site 5 (Horizontal Curve Realignment, New Milford) Crashes by Type

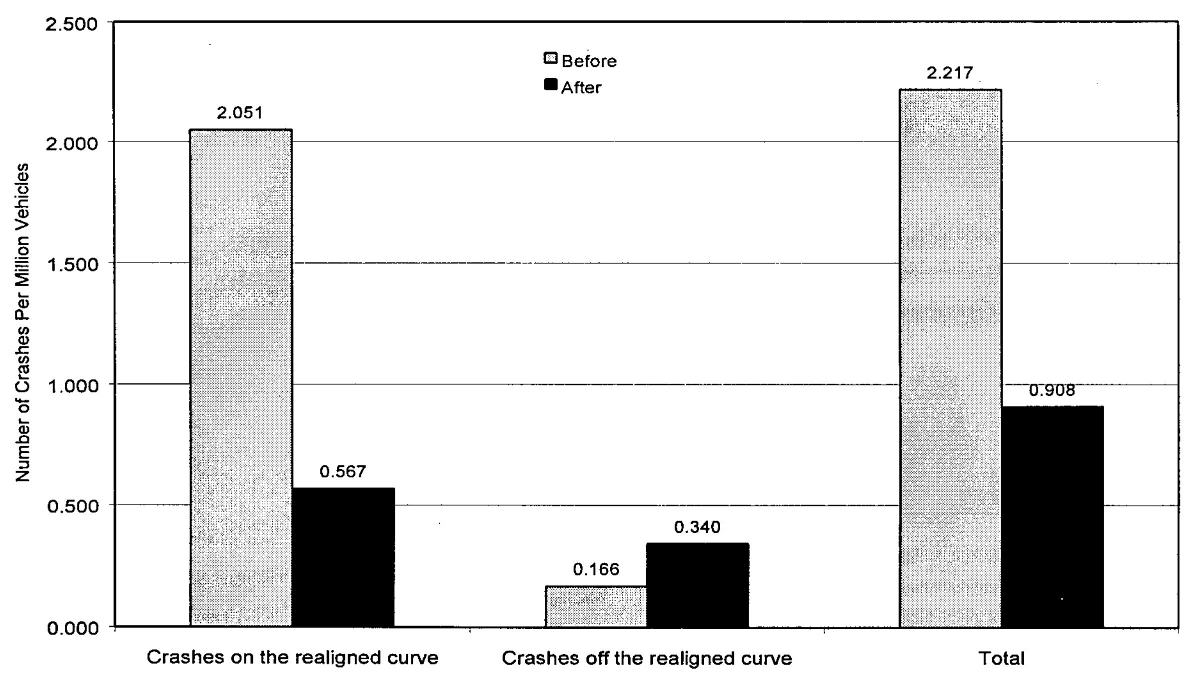


Figure 11. Site 5 (Horizontal Curve Realignment, New Milford) Crashes by Location

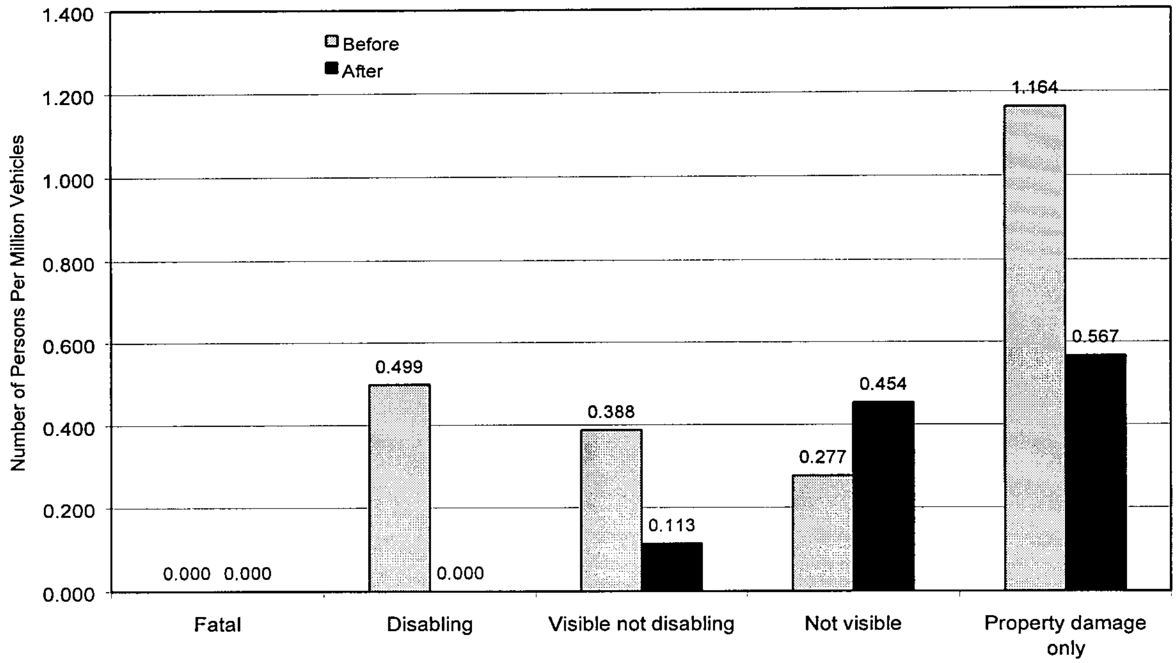


Figure 12. Site 5 (Horizontal Curve Realignment, New Milford) Crash Severity

The crash reduction factors computed for approach angle realignment and horizontal curve realignment are presented in Tables 3 and 4 respectively. The information includes the reduction factor for each site and each type of accident. The expected reduction factors for each type of improvement were calculated, along with a 90% confidence interval. The values are presented as percentages. Negative numbers denote an increase in accidents, and "NA" denotes division by zero (no accidents during the before, or total study period).

Table 3. Crash Reduction Factors for Approach Angle Realignment

Crash Category			Site 2	Site 4	Mean	Std. Dev.	90% Confidence Interval	
							Lower	Upper
Type	Multi-vehicle non-intersection		100%	100%	100%	0%	100%	100%
	Multi-vehicle intersection	Head-on	100%	100%	100%	0%	100%	100%
		Rear-end	52%	72%	62%	14%	39%	85%
		Other	17%	100%	59%	59%	-38%	155%
		Total	56%	87%	72%	22%	35%	108%
	Run-off road		NA	78%	78%	NA	NA	NA
	Hit animal		-42%	100%	29%	100%	-135%	194%
Severity	Killed		NA	100%	100%	NA	NA	NA
	Injury	Disabling	100%	100%	100%	0%	100%	100%
		Visible not disabling	100%	100%	100%	0%	100%	100%
		Not visible	48%	100%	74%	36%	14%	134%
		Total injury	61%	100%	80%	28%	35%	126%
	Property		24%	81%	53%	40%	-13%	118%
Total			42%	87%	65%	32%	12%	117%

Table 4. Crash Reduction Factors for Horizontal Curve Realignment

Crash Category			Site 3	Site 5	Mean	Std. Dev.	90% Confidence Interval	
							Lower	Upper
Type	Non-intersection head-on & sideswipe		34%	71%	52%	26%	9%	96%
	Non-intersection driveway related		100%	-518%	-209%	437%	-928%	510%
	Multi-vehicle intersection		NA	66%	66%	NA	NA	NA
	Run-off road		86%	71%	78%	11%	61%	96%
	Hit animal		100%	100%	100%	0%	100%	100%
Location	On curve		88%	72%	80%	11%	62%	98%
	Off curve		33%	-105%	-36%	97%	-196%	124%
Severity	Killed		NA	NA	NA	NA	NA	NA
	Injury	Disabling	100%	100%	100%	0%	100%	100%
		Visible not disabling	79%	71%	75%	6%	65%	84%
		Not visible	NA	-64%	-64%	NA	NA	NA
		Total injury	81%	51%	66%	21%	31%	101%
	Property		69%	61%	65%	6%	55%	75%
Total			73%	59%	66%	10%	50%	83%

From Tables 3 and 4, we can see that in the approach angle realignment group, the mean reduction was 65 percent for the total crash rate, 72 percent for the target crash (multi-vehicle intersection) rates, and the number of persons injured was reduced by 80 percent. The difference for the two study sites in terms of total crash reduction was greater than that the reduction for the target crashes. This may partly due to the large portion of animal crashes at site 2 (Madison) and other unobserved features of each site.

In the horizontal curve realignment group, the mean reduction was 66 percent for the total crash rate, 80 percent for the target crash (multi-vehicle intersection) rate, and the number of persons injured was reduced by 66 percent. The variance in the crash reduction factors for the two study sites in this group was relatively small compared with that in the approach angle realignment group.

Likelihood functions are plotted in figures 13 through 16 for total, target, and injury, respectively, for the approach angle realignment. Figures 17 through 20 give the same information for the horizontal curve realignment. The most probable value of total crash reduction for the approach angle realignment was 0.611 (the maximum value of $L(\theta)$ corresponds to a θ of 0.389, so the most likely value of the crash reduction factor should be (1- 0.389) or 0.611). Following the same procedure, we have 0.672 for the target reduction, and 0.759 for the injury reduction. The most probable value of total crash reduction for the horizontal curve realignment was 0.653, and 0.777 for the target, 0.620 for the injury reduction.

The most likely values for the crash reduction factors calculated from likelihood functions are very close to the point estimators. For example, the total crash reduction factor for curve realignment was 66 percent by the point estimator, and the most likely value obtained from the likelihood functions was 0.653. The point estimators seem to have higher values in general. Since the likelihood functions reflect the probability distribution of the reduction factor, they provide a clear picture of the uncertainty of the results.

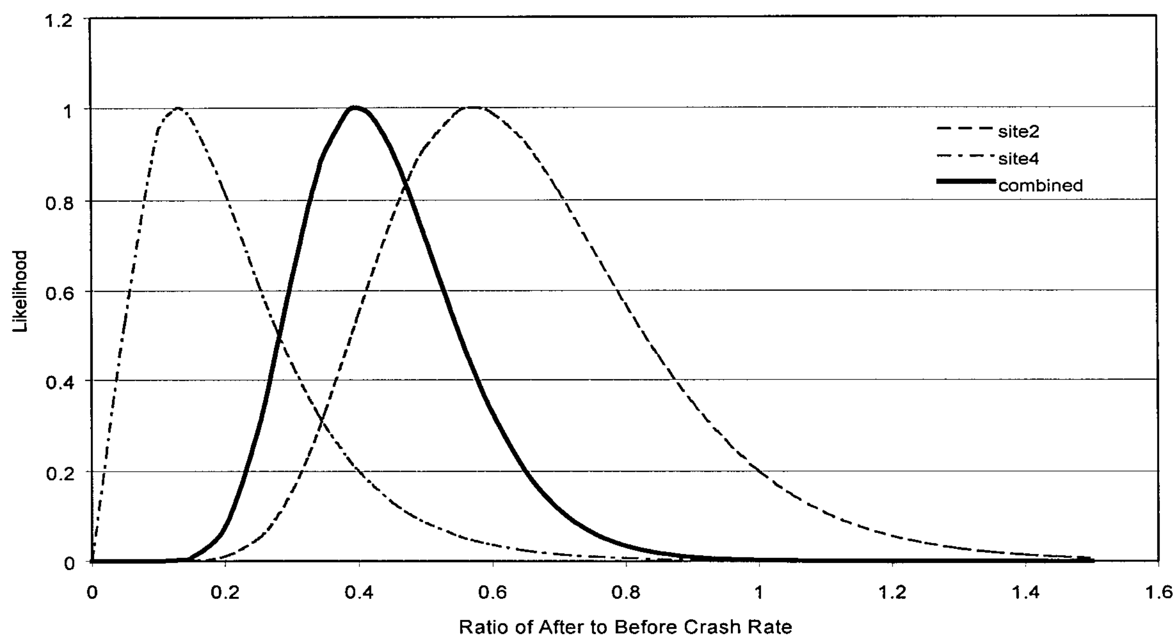


Figure 13. Total Crash Reduction Likelihood Functions for Approach Angle Realignment

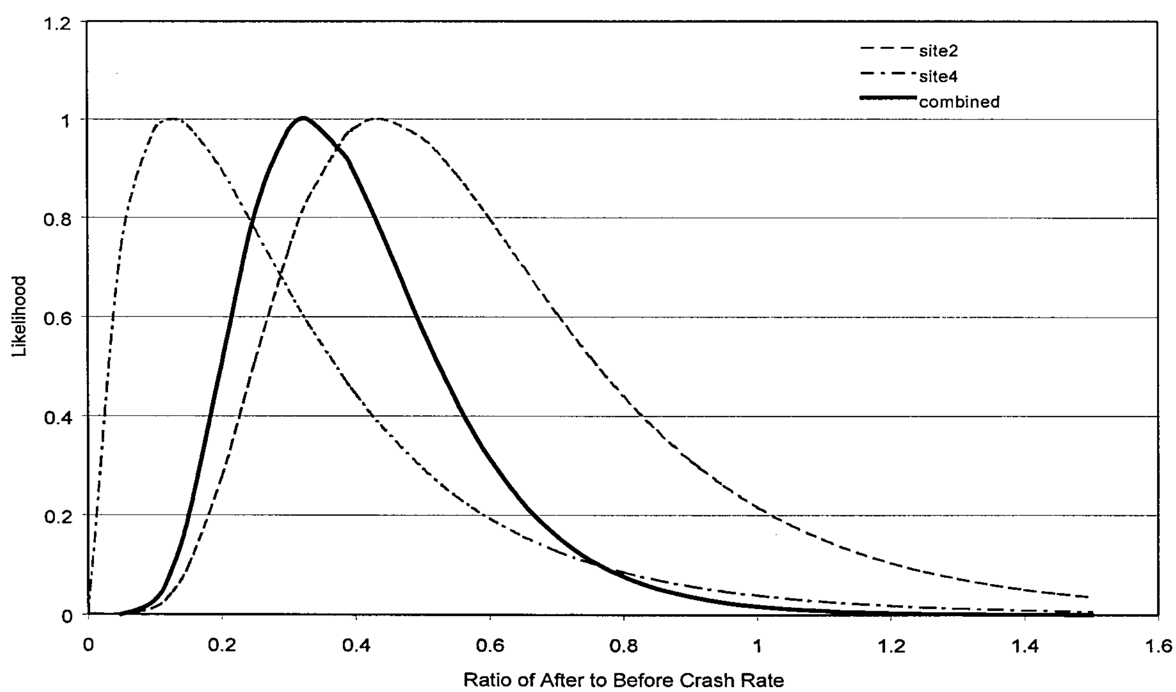


Figure 14. Target Crash (Intersection Multi-vehicle) Reduction Likelihood Functions for Approach Angle Realignment

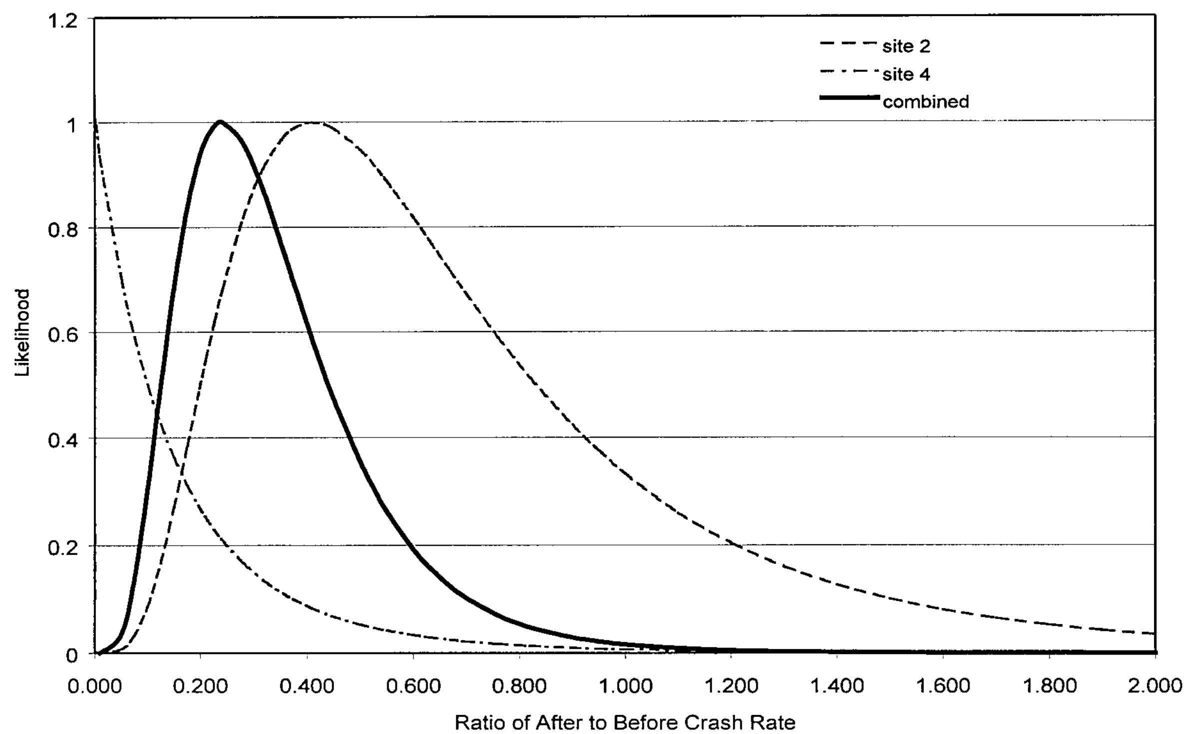


Figure 15. Injury Reduction Likelihood Functions for Approach Angle Realignment

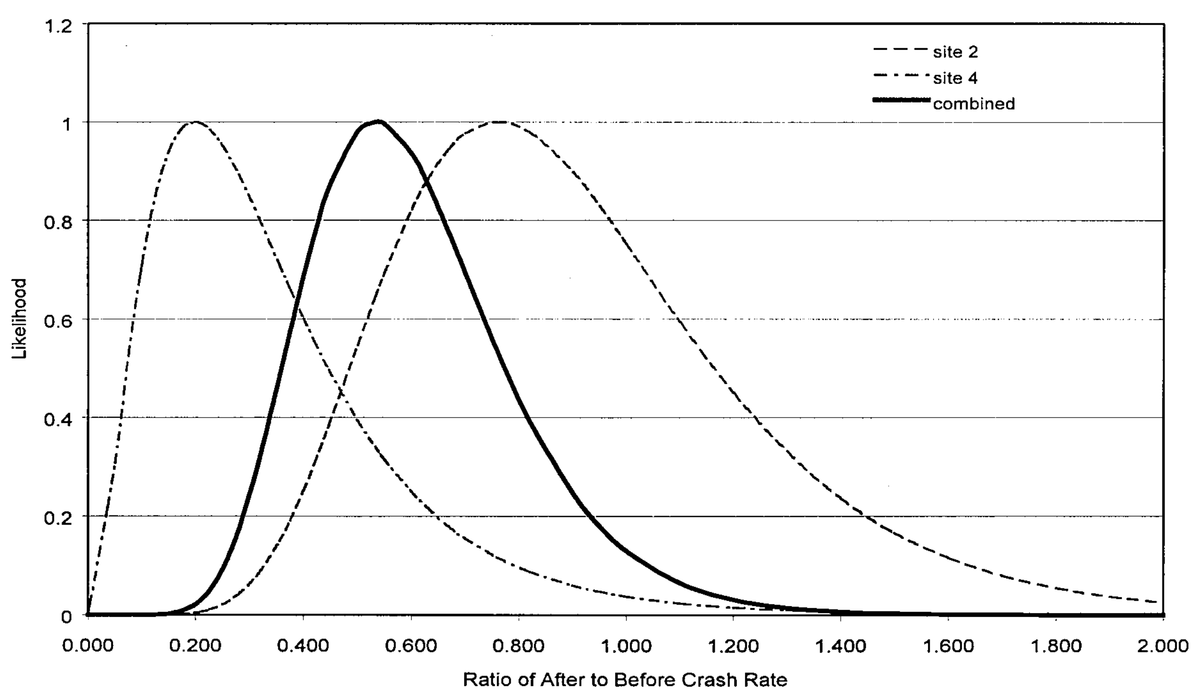


Figure 16. Property Damage Reduction Likelihood Functions for Approach Angle Realignment

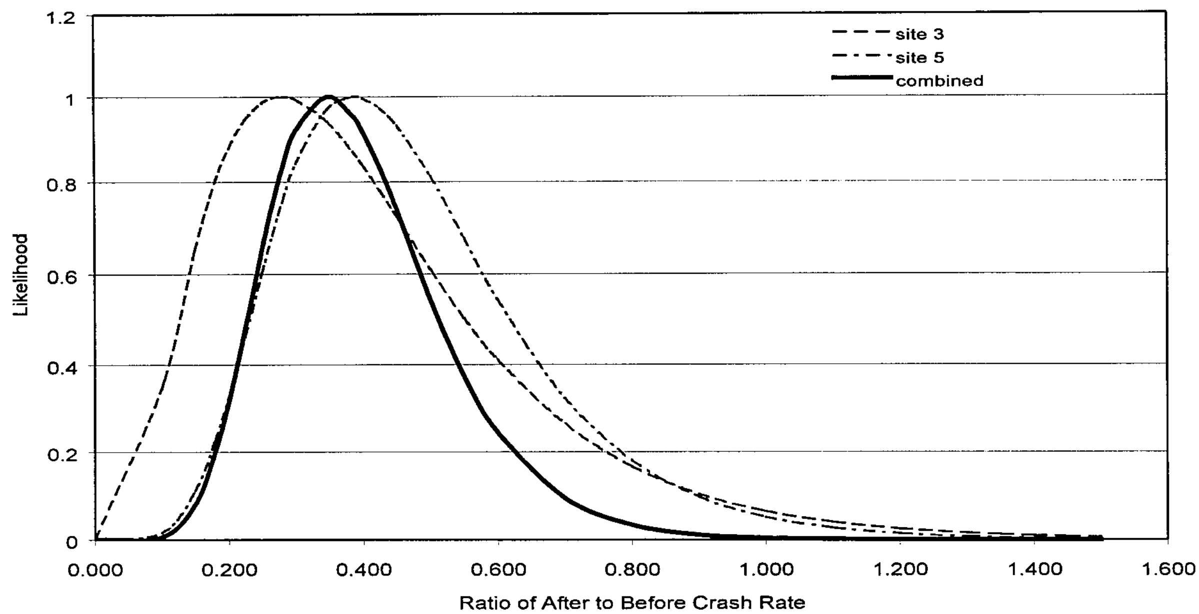


Figure 17. Total Crash Reduction Likelihood Functions for Horizontal Curve Realignment

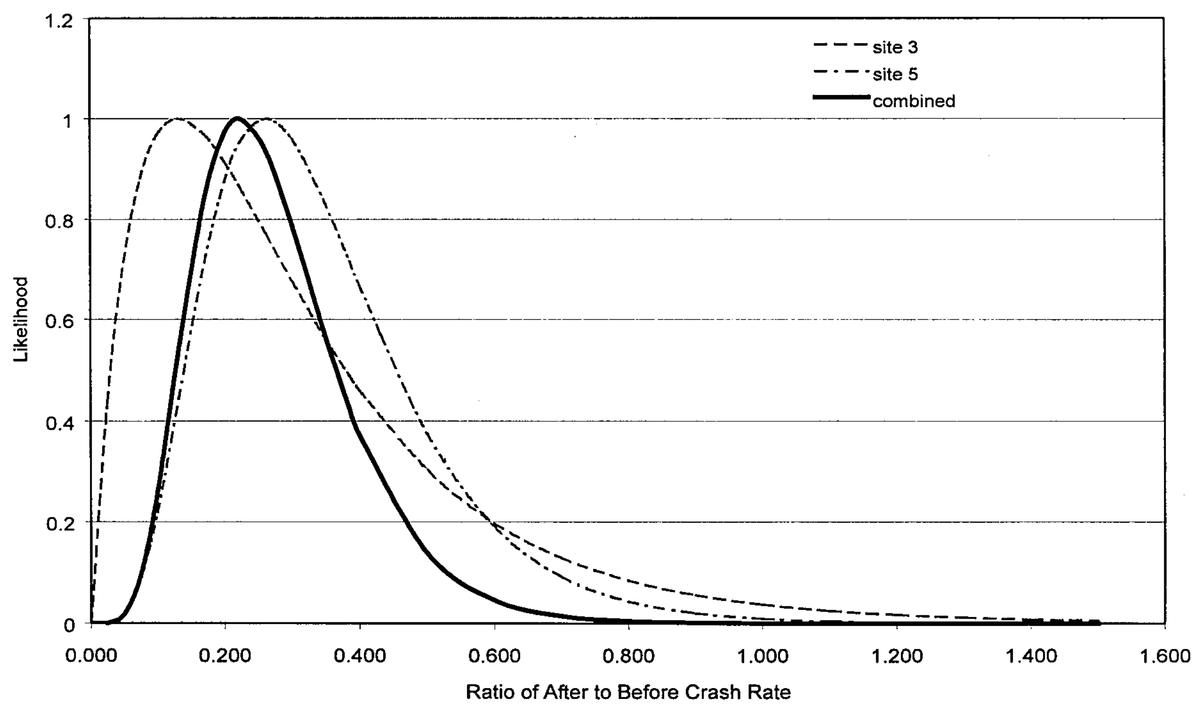


Figure 18. Target (On-Curve) Crash Reduction Likelihood Functions for Horizontal Curve Realignment

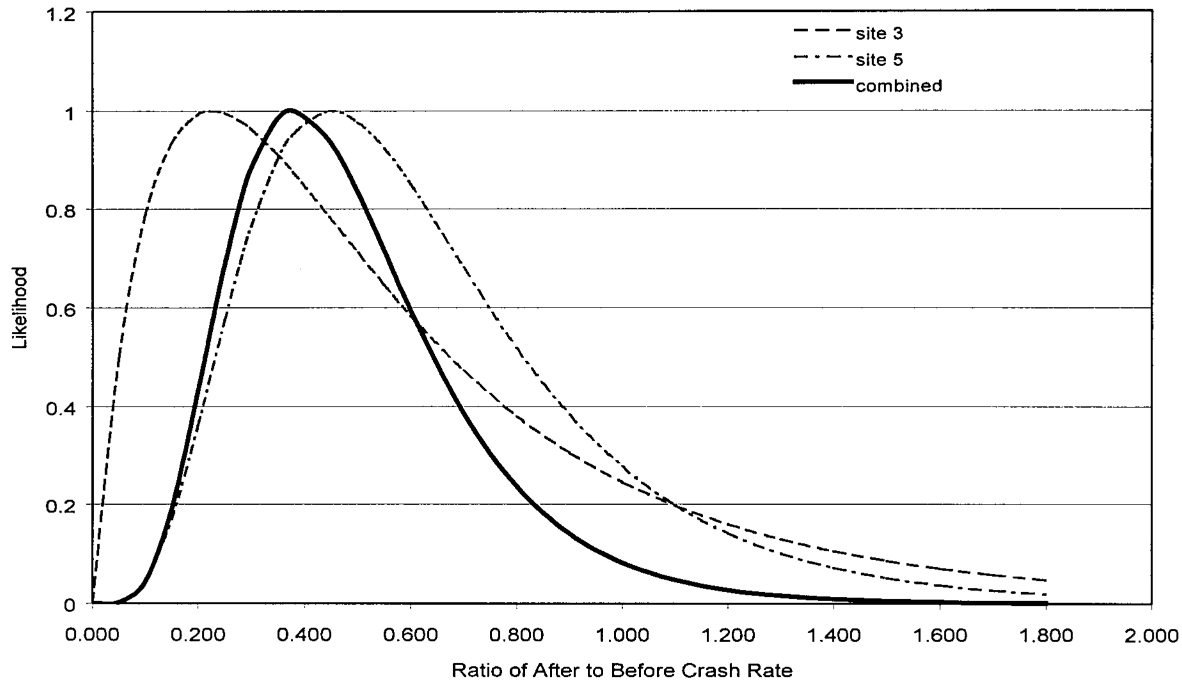


Figure 19. Injury Reduction Likelihood Functions for Horizontal Curve Realignment

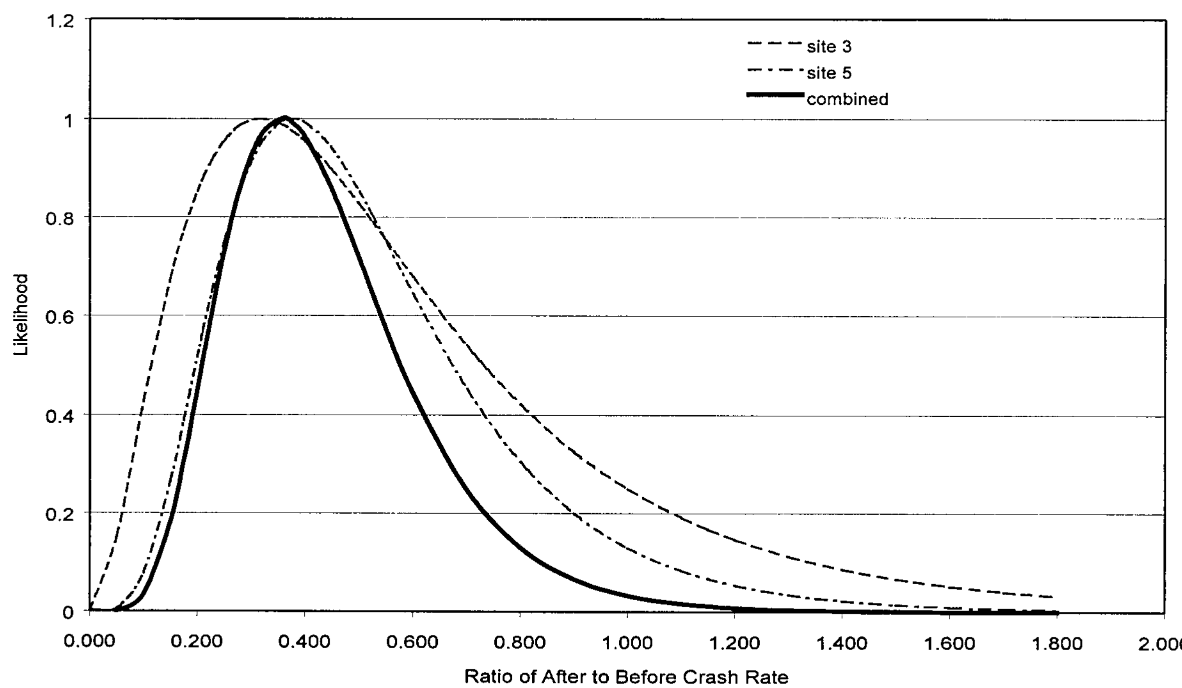


Figure 20. Property Damage Reduction Likelihood Functions for Horizontal Curve Realignment

6. CONCLUSIONS

A. Summary

Identifying crash causation and predicting the reduction in crashes through highway improvements is a very important aspect of traffic safety engineering. In Connecticut, much of the predictive process is based on an earlier study using data (much of questionable validity) dating back to 1947 [2]. There is a need to update the prediction of crash reduction factors based on current data.

The ultimate objective of this project is to update the prediction procedure. The study for Phase 1 focused on the availability of data, ease of data collection, processing requirements, methodology and procedure for conducting the crash reduction study. Two methods of before-and-after analysis for calculating crash reduction factors were demonstrated in this phase. Due to scope of work for this first phase of the project, the extra effort required to collect data for control cases was not practical. Including such cases would improve the statistical reliability of the results, so the next phase of analysis will do so.

In the example application, study sites were selected from the PCMS list, which gives information on projects that have been implemented in recent years by ConnDOT. Accident data were available from 1989 to 1996 in the Connecticut Department of Transportation Accident Experience database. Traffic exposures were calculated from automatic traffic recorder data. Geometric data were collected through the photolog, as-built plans, and site visits. The only information not available was the geometric characteristics on some minor roads before the improvement. However, the traffic volumes on these minor roads were very low compared with that on the main roads, so the absence of these data is likely to have a negligible effect on the overall results.

Approximately four months were spent collecting the data. This included one to two months collecting geometric features from the photolog, construction plans, site visits, approximately one-month extracting crash data from ConnDOT's database, one week calculating the exposure, and probably one-month data entry. These estimates are based on one person working forty hours a week. Table 5 presents the information of the approximate time that spent on different tasks in Phase 1.

Table 5. Time Spent on Phase 1 Tasks

Tasks		Time spent on tasks
Geometric data collection	Photolog	five weeks
	Construction plan	one week
	Site visit	one week
	Data entry	three weeks
Crash data collection	Data extraction	two weeks
	Data entry	four weeks
Calculation of exposure and data entry		one week

The procedure for estimating safety benefits from specific highway improvement is developed now. The application of the procedure has been demonstrated in this study. It has been shown that this procedure is practical and feasible.

B. Plans for phase 2

In Phase 1, we developed the methodology and procedure, and conducted a sample application to test the applicability of this procedure. In Phase 2, we need to conduct the analysis based on a larger sample size. In addition more treatments such as adding left turn lanes and installation of signals will be studied.

For the treatments already examined, we only have two sites in each category at present. Therefore, the results are not transferable and cannot be used in decision making. In the next phase, more sites will be added in each category. Control cases will also be used to reduce the regression-to-mean effect. The results will be compared to those from the first phase computed without control cases to see how the regression to mean effect can affect the results. The use of control cases will also give better estimations of mean and variance, thus giving a better distribution shape for the likelihood function. In addition, accident data for 1997 are now available, so one more year of crash data for the after period can be used in the analysis. All of these factors will help us to obtain a more robust analysis. This phase did not deliberately apply methods of statistical sampling or data collection costs, but a more rigorous study design will be conducted in the next Conflict analysis methods may also be introduced into the next phase. Because it can take a long time to accumulate sufficient quantities of crash data for analysis, some investigators have suggested that the "traffic conflict" technique can be used as a surrogate for crash records. The next phase will first examine the reliability of the techniques and then, if appropriate and feasible within the scope of the project, it will be used to extend the "after" database.

7. REFERENCES

1. R. Clark Bennett. Highway Safety Management System. *Transportation Engineers 63rd Annual Meeting, Compendium of Technical Papers,* Sep. 1993, pp. 388-390.
2. W. McShane, and R. Roess. *Traffic Engineering*. Prentice Hall, New Jersey, 1990.
3. C.C. Robinson. Safety - An important Responsibility. *ITE Journal*, Vol. 61, No. 7, Jul. 1991, pp. 21-24.
4. John C. Laughland, et al. *Methods for Evaluating Highway Safety Improvements. NCHRP Report 162*, Transportation Research Board, National Research Council, 1975.
5. Roy Jorgensen. *Evaluation of Criteria for Safety Improvements on the Highway.* Roy Jorgensen and Associates, Westat Research Analysts, Inc. 1966, p. 227.
6. C. Abbess, D. Jarret, and C.C. Wright. Accidents at Blackspots: Estimating the Effect of Remedial Treatment with Special Reference to the Regression-to-Mean Effect, *Traffic Engineering and Control*, Vol. 22, No.10, Oct. 1981, pp. 535-542.
7. Stig Danielsson. Estimation of the effects of Countermeasures on Different Types of Accidents in the Presence of Regression. *Accident Analysis and Prevention*, Vol. 20, No. 4, Aug. 1988, pp. 289-298.
8. E. Hauer. On the Estimation of the Expected Number of Accidents, *Accident Analysis and Prevention*, Vol. 18, No. 1, Feb. 1986, pp. 1-12.
9. E. Hauer, and B. N. Persaud. How to Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effects of Warning Devices. *Transportation Research Record 1114,* TRB, National Research Council, Washington D.C. 1988, pp. 131-140.
10. E. Hauer. *Observational Before – After Studies in Road Safety*. Pergamon, 1997.
11. C. C. Wright, C. R. Abess, and D. F. Jerrett. Estimating the Regression-to-Mean Effect Associated with Road Accident Black Spot Treatment: Towards a More Realistic Approach. *Accident Analysis and Prevention*, Vol. 20, No. 3, Jun. 1988, pp. 199-214.
12. Hashem R. Al-Masaeid, Kumares C. Sinha, and Thomas Kuczek. Evaluating of Safety Impact of Highway Projects. *Transportation Research Record 1401,* TRB, National Research Council, Washington D.C. 1993, pp. 9-16.
13. Michael Yiu-Kuen Lau, Adolf D. May. *Accident Prediction Model Development for Unsignalized Intersections: Final Report.* Institute of Transportation Studies, University of California at Berkeley, 1989.
14. E. Hauer. Reflections on Methods of Statistical Inference in Research on the Effect of Safety Countermeasures. *Accident Analysis and Prevention.* Vol. 15, No.4, Aug. 1983, pp. 275-285.
15. Ezra Hauer, and Jane Lovell. New Directions for Leraning about the Safety Effect of Measures. *Transportation Research Record 1068,* TRB, National Research Council, Washington D.C. 1986, pp. 96-102.
16. James A. Bonneson, and Patrick T. McCoy. Estimation of Safety at Two-way Stop-controlled Intersections on Rural Highways. *Transportation Research Record 1401*, TRB, National Research Council, Washington D.C. 1993, pp. 83-89.
17. Michael Yiu-Kuen Lau, and Adolf D. May. Injury Accident Prediction Models for Signalized Intersections *Transportation Research Record 1172*, TRB, Washington D.C. 1988, pp. 58-67.

18. John N. Ivan, and, Patrick J. O'Mara. Prediction of Traffic Accident Rates Using Poisson Regression. Paper Presented at the Transportation Research Board 76th Annual Meeting, January 12-16, 1997, Washington, D.C., Paper No. 970861
19. Vergil G. Stover. Access Control Issues Related to Urban Arterial Intersection Design. *Transportation Research Record* 1385, TRB, Washington D.C. 1993, pp. 22-31.
20. Research Pays Off: Laser Videodisc Technology Meets Changing Operational Demands. *TR News 176*, 1995, pp. 24-25.
21. Poisson Regression. Paper Presented at the Transportation Research Board 76th Annual Meeting, January 12 - 16, 1997, Washington, D.C., Paper No. 970861
22. Vergil G. Stover. Access Control Issues Related to Urban Arterial Intersection Design. *Transportation Research Record 1385*, TRB, Washington D.C. 1993, pp. 22-31.
23. Research Pays Off: Laser Videodisc Technology Meets Changing Operational Demands. *TR News 176*, 1995, pp. 24-25.