Journal of Safety Research 42 (2011) 117–129

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

# Case–control and cross-sectional methods for estimating crash modification factors: Comparisons from roadway lighting and lane and shoulder width safety effect studies

Frank Gross [a,*], Eric T. Donnell [b]

[a] Vanasse Hangen Brustlin, (VHB), Inc. 333 Fayetteville St, Suite 1450, Raleigh, NC 27601, USA
[b] The Pennsylvania State University, Department of Civil and Environmental Engineering, 212 Sackett Building, University Park, PA 16802, USA

ARTICLE INFO

Available online 22 March 2011

*Keywords:*
Road safety
Case–control
Cross-sectional
Crash modification factor
Countermeasure evaluation

ABSTRACT

*Problem:* While observational before–after studies are considered the industry standard for developing crash modification factors (CMFs), there are practical limitations that may preclude their use in highway safety analysis. There is a need to explore alternative methods for estimating CMFs. *Method:* This paper employs case–control and cross-sectional analyses to estimate CMFs for fixed roadway lighting and the allocation of lane and shoulder widths. *Results:* Based on the case–control method, the CMF for intersection lighting is 0.886, while the cross-sectional study indicates a CMF of 0.881. The CMFs developed for lane and shoulder widths are also similar when comparing the two methods. *Conclusions:* This paper suggests that case–control and cross-sectional studies produce consistent results if care is taken in the study design and model development. *Impact on industry:* Case–control and cross-sectional studies may provide a viable alternative to estimate CMFs when a before–after study is impractical due to data restrictions.

## 1. Introduction

In an observational before–after study, a safety countermeasure is implemented at a location and the safety effect is estimated by comparing the observed crash frequency after implementation to an estimate of the expected number of crashes that would have occurred had the countermeasure not been implemented. If the only change to the site being evaluated is the implementation of a single safety countermeasure, then it is reasonable to conclude that the countermeasure caused the observed change in crashes at the site. Harwood, Council, Hauer, Hughes, and Vogt (2000) suggest that well-designed observational before–after studies offer several advantages over other safety countermeasure evaluation methods; however, there are several practical limitations, including:

1. *Confounding factors:* several improvements may be implemented simultaneously, making it difficult to isolate the effect of a single countermeasure using a before–after study. Similarly, changes in traffic volume, driver population, vehicle mix, and other factors may occur over the analysis time period in a before–after study.
2. *Sample size:* it is sometimes difficult to find an adequate sample of sites where the treatment of interest has actually been implemented. Results from a limited sample will have a high level of statistical uncertainty. When there are few or no sites being treated with the countermeasure of interest, a before–after study is difficult to employ.
3. *Study period:* a before–after study requires a time sequence, where it is necessary to implement a countermeasure and wait for sufficient data in the after period. While data collection can be time consuming for any safety evaluation, waiting several years after implementation is a practical concern in before–after studies.

Given the limitations associated with observational before–after studies, alternative evaluation methods are sometimes needed to provide estimates of countermeasure safety effectiveness. *A Guide to Developing Quality Crash Modification Factors* introduces various methods for estimating the safety effects of countermeasures (Gross, Persaud, & Lyon, 2010). Several variations of observational before–after studies are presented along with alternative methods. Two of the alternative methods are explored in this paper.

One alternative method to estimate the safety effectiveness of a countermeasure is a cross-sectional study. Hauer (2010) argues that cross-sectional studies have not proven successful to identify cause and effect in road safety because multivariable regression typically does not produce consistent results between studies. He suggests that an observational epidemiological approach may, however, be a viable method to control for the many sources of variation present in cross-sectional data. A case–control study is one example of an observational epidemiology evaluation method that may be used to develop crash modification factors (CMFs) for a countermeasure. There is a need to compare CMFs developed from observational case–control,

---

* Corresponding author. Tel.: +1 919 834 3972; fax: +1 919 834 3970.
*E-mail addresses:* fgross@vhb.com (F. Gross), edonnell@engr.psu.edu (E.T. Donnell).

0022-4375/$ – see front matter © 2011 National Safety Council and Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsr.2011.03.003

before–after, and cross-sectional studies, given the same data limitations, to investigate their potential as alternative methods for safety evaluations.

The objective of this study is to compare case–control and cross-sectional methods to estimate measures of safety effectiveness using two independent datasets. The safety effects of fixed, at-grade intersection lighting in Minnesota were estimated using both evaluation methods. Similarly, the safety effects of lane and shoulder width dimensions were evaluated using two-lane, rural highway data from Pennsylvania. An observational before–after evaluation was not considered in the present study because roadway lighting is seldom the only countermeasure applied to a site, making it difficult to isolate the safety effects of roadway lighting using this method. Similarly, the lane and shoulder width evaluation did not involve a treatment; rather, various lane and shoulder width combinations were compared to a baseline condition.

## 2. Background

### 2.1. Cross-sectional studies

Cross-sectional studies are commonly used in transportation safety research to estimate the expected number of crashes on a roadway segment, interchange, or intersection. CMFs derived from cross-sectional data are based on a prescribed time period under the assumption that the ratio of average crash frequencies for sites with and without a feature is an estimate of the CMF for implementing that feature. The strength of a cross-sectional study is that a large number of sites with and without a specific countermeasure can often be identified. A weakness of a cross-sectional study is that it is difficult to determine the reason that certain safety countermeasures exist at one location and not at other similar locations. As such, the observed difference in crash experience can be due to known or unknown factors other than the feature of interest. Known factors, such as traffic volume or geometric characteristics, can be controlled for in principle by estimating a multivariate regression model. However, the issue is not completely resolved since it is difficult to properly account for unknown, or known but unmeasured, factors. Several examples of developing CMFs from a cross-sectional study are contained in the literature. Lord and Bonneson (2007) developed CMFs for lane width, shoulder width, and edge-line marking presence for frontage roads in Texas. Bonneson and Pratt (2008) recently proposed a procedure to develop CMFs for curve radius along two-lane rural highways. Additionally, Fitzpatrick, Lord, and Park (2008) developed CMFs for median width on freeways and rural multi-lane highways in Texas.

### 2.2. Case–control studies

Case–control designs are well established in epidemiology where they are used to relate risk factors within a population to a particular outcome or disease. In the highway safety context, their use has often been limited to studies of the road-user and vehicle (Tsai, Wang, & Huang, 1995; Stevenson, Jamrozik, & Spittle, 1995; Jovanis, Park, Chen, & Gross, 2005). More recently, the case–control method has been applied to estimate the safety effectiveness of lane and shoulder width (Gross & Jovanis, 2007; Gross & Jovanis, 2008); the results of this research showed a striking resemblance to CMFs proposed in the Highway Safety Manual for two-lane, rural highways (American Association of State Highway Transportation Officials [AASHTO], 2010). This study concludes that case–control methods may be a viable alternative for estimating CMFs when before–after studies are not practical or feasible.

Case–control studies assess whether exposure to a potential risk factor is disproportionately distributed between cases and controls, thereby indicating the likelihood of an outcome given the presence of the risk factor. Case–control studies produce an estimate of the odds ratio, which can be used as a direct estimate of safety effectiveness. The odds ratio is a measure of the percent change in the chance of an outcome given the presence of a risk factor compared to the baseline level of the risk factor. This lends itself well to the approximation of CMFs because the purpose is to provide an estimate of the incremental safety effect of a particular feature in relation to a certain baseline level.

The case–control method, in general, is associated with several advantages over alternative safety evaluation methods, and the matched case–control design has additional distinct advantages as follows:

- *Studying rare events*: the case–control design is ideal for studying rare events, such as crashes, because the sample may be selected so that a pre-specified number of cases are enrolled in the study, ensuring an adequate sample for analysis.
- *Evaluating multiple risk factors from a single sample:* the sample is selected based on outcome status and investigated to determine potential risk factors. Any variables not included in the case definition or matching scheme may be assessed, simultaneously, as individual risk factors.
- *Controlling for confounding variables:* confounders include variables that completely or partially account for the apparent association between an outcome and risk factor. Specifically, a confounder is a variable that is a risk factor for the outcome under study, and is associated with, but not a consequence of, the risk factor in question (Collett, 2003). In highway safety, an example of a confounder is average daily traffic (ADT). ADT has been shown to be associated with crash risk and is also associated with, but not a consequence of, several geometric features (e.g., lane width, shoulder width, and horizontal curvature).
- *Matching:* the primary reason for a matched design is to directly control for confounding variables. Control sites are matched to each case through random sampling based on similar values of potential confounding variables.
    - o Matching provides a balanced design and adjusts for the effects of variables included in the matching scheme.
    - o Matching ensures that adjustment is possible when the confounding variable is distributed differently within the case and control populations. In rare cases, the distribution of a confounding variable may not overlap for a random sample of cases and controls. In this case, there would be no way to adjust the results during the analysis phase.
    - o Matching improves the efficiency of the design, requiring smaller sample sizes or resulting in estimates with a narrower confidence interval. However, this only holds when the matching is based on true confounders (Woodward, 2005).

Case–control designs are appealing due to their ability to estimate risk while properly controlling for confounding variables; however, there are disadvantages that must be recognized and addressed. Disadvantages of the case–control method are as follows:

- Case–control studies cannot be used to measure the probability of an event (e.g., crash, severe injury) in terms of expected frequency. They are more often used to show the relative effects of risk factors.
- Case–control studies often rely on collecting retrospective data for risk factors and outcome status, relying on the availability of historical documentation to provide information regarding risk factors and outcomes.
- Case–control studies are based on cross-sectional data; however, they should not be confused with cross-sectional studies in general. Case–control studies select subjects based on outcome status where cross-sectional studies generally sample based on risk factor status. Whether used for a case–control design or cross-sectional design, cross-sectional data do not involve a time sequence of data collection. Hence, they can only demonstrate associations, not causality.
- Although case–control studies may be used to explore multiple risk factors, they can only investigate one outcome per sample.