## Title

Machine Learning Can Reveal Effectiveness of Traffic Safety Countermeasures



**Authors**
Li, Jia, PhD
Qi, Yanlin
Zhang, Michael, PhD

**Publication Date**
2025-08-01

**DOI**
10.7922/G2VD6WVF

POLICY BRIEF

Institute of Transportation Studies | UNIVERSITY OF CALIFORNIA

# Machine Learning Can Reveal Effectiveness of Traffic Safety Countermeasures

**Jia Li, Ph.D., Department of Civil and Environmental Engineering, Washington State University**
**Yanlin Qi, Institute of Transportation Studies***
**Michael Zhang, Ph.D., Department of Civil and Environmental Engineering***
***University of California, Davis**

**August 2025**

## Issue

Emerging machine learning capabilities can be leveraged to make transportation infrastructure safer and reduce fatalities by informing decisions about which countermeasures to apply at crash-prone locations. At this time, project prioritization typically involves assessing effectiveness, cost-benefit ratios, and available funding. Crash Modification Factors (CMFs) play an essential role in project assessment by predicting the effectiveness of safety countermeasures. Their applicability has limitations, however. Some of these may be overcome with innovative approaches such as knowledge-mining.

The US Department of Transportation's (DOT) CMF Clearinghouse provides practitioners with a list of reliable CMFs developed from individual studies. However, available CMFs do not cover all potential scenarios-of-interest to State DOTs because unique projects may feature novel infrastructure types or countermeasures. Experimental or observational studies are the dominant tools for estimating CMFs. However, these approaches may require years of effort to collect adequate crash data.

To address these challenges, we developed a machine learning framework that mines CMF Clearinghouse data to uncover previously unidentified relationships. This provides a cost-effective and time-efficient solution to assessing CMFs not covered by the CMF Clearinghouse. Our proposed framework fully explores existing CMF data.

We extensively trained and tested the proposed approach on CMF Clearinghouse data with experiments and showed that the framework can predict CMFs with reasonable accuracy. The framework flexibly incorporates heterogeneous data from the CMF Clearinghouse, captures the semantic contexts of countermeasures, and maintains data compatibility.

## Key Research Findings

The two dominant approaches for estimating CMFs are experimental or observational studies based on before/after or cross-sectional data. Experimental studies are rigorous, requiring years of effort to collect data. This long study cycle hinders wide applications. Observational studies, in contrast, bring fewer ethical concerns than experimentation and can leverage data from sites with already-implemented treatment(s). Before-after studies typically require multi-year crash data collected before and after a treatment. Most cross-sectional studies require crash inventory data collected from multiple sites with and without a treatment in a single period. Both types of studies require substantial quantities of crash data.

**A machine-learning framework can mine CMF Clearinghouse data.** As a one-stop repository, the CMF Clearinghouse has long been used as a search database to support practitioners responsible for infrastructure safety. With thousands of existing CMF records and continuing entries, the CMF Clearinghouse has tremendous potential as a knowledge base for deriving additional CMFs. Our knowledge-mining approach for deriving additional CMFs uses information already available, thereby reducing the data-collection and study duration burden of estimating CMFs not covered by the CMF Clearinghouse.

**Machine learning for estimating CMF effectiveness is a work-in-progress.** Our approach is meant to provide a preliminary estimate of CMFs and is not intended to replace traditional methods for developing CMFs for specific scenarios. The accuracy of machine learning output can be improved with additional development, for example, if structured or domain knowledge is leveraged. We developed a specialized model handling structured countermeasure descriptions and this model indeed obtained a better accuracy. In the current framework, the values of CMFs are predicted, however, the confidence levels of these predictions are not provided. This information would be helpful in determining the fidelity of CMF predictions.

**Extensions to the framework can make machine learning more effective in estimating CMFs.** Promising extensions include: (1) combining the proposed approach with traditional approaches to improve CMF estimation accuracy and reduce cost, (2) using domain knowledge to enhance predictive power, (3) leveraging existing CMF patterns to generate confidence intervals that bracket likely variation around predicted outcomes, and (4) taking advantage of the additional information in case-specific CMF records.

## Policy Recommendations

Based on this study, we recommend that the proposed machine learning-based approach can be used for quick and cost-effective CMF estimation for initial screening and prioritization of critical traffic safety countermeasures, which complements the traditional approaches.

## More Information

This policy brief is drawn from the report "Developing a Safety Effectiveness Evaluation Tool for California" prepared by Yanlin Qi, Jia Li, and Michael Zhang with the University of California, Davis. The report can be found at https://www.ucits.org/research-project/2022-04. For more information about findings presented in this brief, please contact Jia Li (jia.li1@wsu.edu) or Michael Zhang (hmzhang@ucdavis.edu).

*Research presented in this policy brief was made possible through funding received by the University of California Institute of Transportation Studies (UC ITS) from the State of California through the Public Transportation Account and the Road Repair and Accountability Act of 2017 (Senate Bill 1). The UC ITS is a network of faculty, research and administrative staff, and students dedicated to advancing the state of the art in transportation engineering, planning, and policy for the people of California. Established by the Legislature in 1947, the UC ITS has branches at UC Berkeley, UC Davis, UC Irvine, and UCLA.*

Project ID UC-ITS-2022-04 | DOI: 10.7922/G2VD6WVF