# Predicting Safety: Development and Application of Bicycle Safety Performance Functions in Seattle, WA

**Libby Thomas[*], Bo Lan[*], Rebecca L. Sanders[#1], Alexandra Frackelton[#2], Spencer Gardner[#3], Michael Hintze[#4]**

[*] Highway Safety Research Center,
University of North Carolina
730 Martin Luther King Jr Blvd, Suite 300
Chapel Hill, NC 27599-3430
email: thomas@hsrc.unc.edu

[*] Highway Safety Research Center,
University of North Carolina
730 Martin Luther King Jr Blvd, Suite 300
Chapel Hill, NC 27599-3430
email: lan@hsrc.unc.edu

[#1] Toole Design Group
319 SW Washington Street, Suite 800
Portland, OR 97214
Email: rsanders@tooledesign.com

[#2] Toole Design Group
8484 Georgia Ave, Suite 800
Silver Spring, MD 20910
Email: afrackelton@tooledesign.com

[#3] Toole Design Group
16 N. Carroll Street, Suite 200
Madison, WI 53703
Email: sgardner@tooledesign.com

[#4] Toole Design Group
811 1st Avenue, Suite 500
Seattle, WA 98104
Email: mhintze@tooledesign.com

**Keywords:** bicycle safety, safety performance function, left hook, bicycle lane safety.

## 1 INTRODUCTION

Bicycle collisions, particularly more severe collisions, typically occur in low numbers at any one location. Even if there are significant numbers of crashes, high crash locations tend to shift around over time due to a certain amount of randomness in crashes, and a tendency to regress toward the mean (RTM). Yet many agencies default to prioritizing improvements by examining patterns of crash frequency because they lack the experience or data needed to conduct more robust safety analyses. Safety analysis that only identifies prior crash hot spots is unable to consistently predict future high crash locations. However, analysts and practitioners can work to predict future crash locations (or risk of crashes) by identifying and analyzing data associated with past crashes across the entire network rather than relying on apparent patterns of a few crash hot spots. Factors or risks associated with crashes across the system can then be used to predict other risky locations for a more proactive treatment or "systemic" safety approach. These crash prediction analyses produce Safety Performance Functions (SPFs), which may then be used to screen and rank locations by overall risk rather than by frequency-based rankings of where crashes have occurred.

This paper describes the development and use of SPFs for three bicycle crash types to help prioritize locations for potential bicycling safety improvements in Seattle, WA. We also describe how the model results are being used as an additional tool by the city in their efforts to improve bicycling safety and achieve Vision Zero. This approach assumes that future crashes can be predicted, in part, based on relationships associated with prior

This work is licensed under the Creative Commons Attribution 4.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

crashes across the whole system, and that these relationships can be used to help prioritize treatable locations, even if crashes have not occurred there previously.

## 2 LITERATURE REVIEW

Due to issues with randomness, RTM, and the non-linear relationship between exposure and crashes, the Highway Safety Manual and emerging literature have focused on safety analysis procedures that account for RTM as well as exposure related to traffic and user volumes. Predictive models can be used to address the issue of too few crashes and RTM by using data from many locations or system-wide and sufficient study years. Additionally, various ranking techniques can use the model predictions to screen the network to identify locations with potential safety issues *(1,2,3)*. However, despite the potential of multivariate safety analysis to illuminate important relationships affecting bicycle safety, issues with low crash numbers and a general lack of resources for study in the industry have resulted in only limited examples of crash prediction models related to bicycle crash analysis in the literature *(4,5,6)* This paper builds on prior research by developing Safety Performance Functions for bicyclist crashes in Seattle, Washington, as part of the City's Vision Zero effort. Further, this paper demonstrates the practicality of these predictive models for prioritizing locations for bicycle safety improvements in a more proactive, systemic approach to bicycling safety.

## 3 METHODS

The data for this project were analyzed using both descriptive analysis and multivariate analysis. For the descriptive analysis, we examined where and when bicycle crashes were occurring and what factors seemed to be over-represented in fatal and serious injury crashes, and we identified high frequency and severity crash type and location scenarios for additional analysis in the second phase. For the multivariate analysis, the subject of this paper, we developed a comprehensive intersections database using GIS, crash, Census, and local economic data. We then tallied frequencies of multiple types of bicycle crashes that occurred at each intersection, developed counts, and analyzed factors related to three bicycle intersection crash types. Our analysis groups included:

- Btot_int: All bicycle intersection crashes (regardless of driver or bicyclist movement). Intersection crashes (n=1753) represented more than half of all bicycle crashes (56.2%).

- BOD_int: Bicyclist opposite direction crashes (bicyclist and driver traveling in the opposite direction, whether turning or going straight). This group (n=462) accounted for 14.8% of all bicycle collisions; these collisions also tended to be more severe.

- BAng_int: Bicyclist angle crashes (driver and bicyclist traveling perpendicularly to one another). These crashes (n=870) accounted for 27.9% of all bicycle collisions.

The multivariate analysis of the intersections database first used Conditional Random Forest (CRF) to reduce the number of potential predictors to a more manageable and likely significant set. We then used Negative binomial (NB) modeling to determine the combinations of features associated with bicycle crash frequency in each scenario (controlling for bicycle volumes). Only variables significant at $p < 0.05$ were retained in the models,

This work is licensed under the Creative Commons Attribution 4.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

and the best model was selected based on statistics that assess the goodness of fit and model parsimony (specifically, AIC and BIC). All analyses were conducted using SAS PROC GLIMMIX. The model results (SPF equations) were then used to predict expected crashes using SPF, Empirical Bayes (EB) estimates, and Potential Safety Improvements (PSI) for all locations and to rank locations according to the predictions.

## 4 RESULTS AND DISCUSSION

The models revealed multiple built environment and land use factors associated with overall crashes as well as different bicycle crash types, although the results varied somewhat by crash type. In particular, variables indicating more potential for conflicts, such as higher numbers of intersection legs, parking, and higher street classification (closely aligned with AADT) were all positively associated with the likelihood of bicycle crashes. The results of these SPFs, when applied system-wide, suggest intersections where crashes may be likely to occur, but have not yet due to randomness. In fact, when the results from the three predictive methods (model predictions, an EB-adjusted prediction, and PSI) were applied to the network, several of the top intersections had not experienced any crashes for the last eight years—but were immediately recognized by SDOT staff as locations where people felt unsafe walking and bicycling. Ideally, the model results will provide data-driven evidence to address these locations before a crash occurs.

## 5 REFERENCES

[1] Lyon, C., & Persaud, B. Pedestrian Collision Prediction Models for Urban Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1818, Transportation Research Board of National Academis, Washington, D.C., 2002, pp. 102-107 .

[2] Persaud, B., Lyon, C., & Nguyen, T. Empirical Bayes Procedure for Ranking Sites for Safety Investigation by Potential for Safety Improvement. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1665, Transportation Research Board of National Academies, Washington, D.C., 1999, pp. 7-12.

[3] Lan, B., & Persaud, B.N. Investigation and Evaluation of Ranking Criteria for Hot Spot Identification using Bayesian Approach. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2237, Transportation Research Board of National Academies, Washington, D.C., 2011, pp. 117-125.

[4] Nordback, K., Marshall, W.E., and Janson, B.N. Bicyclist safety performance functions for a U.S. city. *Accident Analysis and Prevention*, Vol. 65, 2014, pp.114-22.

[5] Strauss, J., Miranda-Moreno, L., & and Morency, P. Multimodal injury risk analysis of road users at signalized and non-signalized intersections. *Accident Analysis and Prevention, 71*, 2014, pp. 201-209.

[6] Turner, S., Wood, G., Hughes, T., and Singh, R. Safety performance functions for bicycle crashes in New Zealand and Australia. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2236, Transportation Research Board of National Academies, Washington, D.C., 2011, pp. 66-73.

This work is licensed under the Creative Commons Attribution 4.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.