

Equitable Estimation of Accurate High Injury Networks (HINs) for Vulnerable Road Users

Project 2459
April 2026

Anurag Pande, PhD

Anandamayee Majumdar, PhD

Lucas Roman Kantorowski



Introduction

Vision Zero policies depend on accurate identification of where and how people are being seriously injured in traffic incidents. In the United States, High-Injury Networks (HINs) are typically derived from police-reported crash data, yet growing evidence suggests that these records substantially undercount injuries involving vulnerable road users (VRUs), including pedestrians and bicyclists. Incomplete data can distort safety priorities, delay interventions, and undermine equity goals.

This research examines whether police-reported crash data alone are sufficient for developing equitable and accurate HINs. The study evaluates the extent of underreporting by comparing police crash records with Emergency Medical Services (EMS) data and real-time 911 call data. Overall, the research assesses whether integrating alternative data sources can provide a more complete, timely, and representative picture of traffic injury risk to better support Vision Zero implementation.

Study Methods

The study used a multi-source comparative approach across two California case studies. First, EMS-reported bicycle and pedestrian injury data from the City of San Luis Obispo were compared with police-reported crash records from the Statewide Integrated Traffic Records System (SWITRS). Spatial analyses and heatmaps were developed to examine where the datasets overlapped and diverged.

Second, the research team collected and archived real-time 911 call data from PulsePoint for the City of San Francisco using an automated scraper. These records were matched to official city traffic collision data using time, location, and street name criteria. Incidents were analyzed at both census tract and police district levels. Statistical analyses, including stepwise logistic regression, were used to examine which factors influenced whether 911-reported incidents appeared in official crash databases. Variables included population density, time of occurrence, incident type, and neighborhood demographics.

Findings

The results show that police-reported crash data alone do not capture the full extent of traffic injuries, particularly for VRUs. In San Luis Obispo, EMS data identified numerous pedestrian and bicycle injury incidents—especially near the Cal Poly campus—that were absent from SWITRS records. Police data were more prevalent on major arterials and in the downtown core, highlighting how roadway context influences reporting. In San Francisco, a substantial share of traffic collisions reported through 911 calls did not appear in official crash records. Matching rates varied significantly across police districts. Dense, institutional districts such as Mission, Central, Northern, and Tenderloin showed higher overlap, while more residential districts—including Richmond, Ingleside, Taraval, Southern, and Bayview—had lower matching rates. Regression results indicated that population density, time of day, and incident type affect whether a 911-reported collision is captured in police data. Neighborhood income and racial composition were not statistically significant predictors.

More than half of traffic collisions reported through 911 calls in some San Francisco districts did not appear in official police crash databases.

Policy Recommendations

Transportation agencies implementing Vision Zero should not rely exclusively on police-reported crash data to identify High-Injury Networks. Integrating EMS and 911 call data can significantly improve the completeness and timeliness of injury surveillance. Agencies are encouraged to develop data-sharing partnerships with EMS providers and explore automated collection of 911 call data where feasible. HIN methodologies should explicitly acknowledge data limitations and consider multi-source validation, particularly in residential areas where underreporting appears most pronounced.

State and regional agencies should invest in standardized data systems and guidance that facilitate

integration of health and emergency response data into safety planning that effectively reflects and addresses real-world risks.

About the Authors

Anurag Pande, PhD

Dr. Pande is a Professor of Civil Engineering at Cal Poly. His research interests include traffic simulation, data mining applications, and observational data analysis, including in the areas of traffic safety and crashes, driver behavior, transportation resilience, and emergency evacuation.

Anandamayee Majumdar, PhD

Dr. Majumdar is an Assistant Professor of Statistics at San Francisco State University. Her research interests include spatial and temporal modeling and simulation, with applications to geographically referenced data, and observational data analysis.

Lucas Roman Kantorowski is a graduate student at Cal Poly. Anurag Pande (PI) and Anandamayee Majumdar (Co-PI) are faculty researchers specializing in traffic safety and in statistical data analysis, respectively.

To Learn More

For more details about the study, download the full report at transweb.sjsu.edu/research/2459



MTI is a University Transportation Center sponsored by the U.S. Department of Transportation's Office of the Assistant Secretary for Research and Technology and by Caltrans. The Institute is located within San José State University's Lucas Graduate School of Business.