Terps for Bike Lanes Improved Data Collection and Aggregation Strategy

STUDENTS
Noah Hanssen, Felipe Veras, and Simone Volman

UNDER THE SUPERVISION OF
Professor Martin Sanders

IN PARTNERSHIP WITH
Terps For Bike Lanes (T4BL) and Partnership for Action Learning in Sustainability (PALS)

COURSE
PLCY400: Senior Capstone

PALS
An initiative of the National Center for Smart Growth
Kathryn Howell, NCSG Executive Director
Kim Fisher, PALS Director

Table of Contents

Abstract

At the University of Maryland (UMD), Terps for Bike Lanes is a student-led organization that advocates for on-campus biking policy and infrastructure changes through conversations with UMD administration and the Student Government Association. Their advocacy efforts have been inhibited by challenges in collecting reliable and statistically significant data regarding on-campus bicycle incidents.

The goal of this research is to identify the best data collection practices for bicycle incidents on college campuses. Prior research has shown that a mixed-methods approach—using a combination of surveys, formal reports, and interviews—is the best practice for collecting comprehensive quantifiable data while controlling for the effects of underreporting and selection bias.

This report identifies three categories of relevant actors: bike advocates, those who collect information on bicycle incidents at UMD, and those who collect data from UMD students. We investigated their data collection methodologies and found that all used reporting forms with key similarities: multiple-choice questions for key data points, broad answer choices with the option to increase specificity, and an option for respondents to provide additional detail or follow-up. At UMD, census-sampled surveys with incentives for respondents were also used to supplement research on a particular topic.

These findings suggest Terps for Bike Lanes can improve their advocacy efforts by adopting a holistic mixed-methods data collection approach that includes an updated reporting form, conducting a campus-wide survey about biking and safety, and conducting interviews and focus groups to gain further information.

Introduction

The Terps for Bike Lanes (TFBL) ARCGIS webpage links and its in-house developed Qualtrics crash reporting form are both flawed and incomplete. This report investigates the survey and data collection software and strategies that would allow TFBL to best obtain, organize and analyze accurate data on biking incidents on-campus for the purposes of policy advocacy.

Bike crash data and university data collection practices have been reviewed and organized by recurring methodologies of self-reporting, naturalistic techniques, and harder-to-quantify technological interventions, all fairly standardized in the field. While the team investigated other technological interventions, the feasibility of implementing them appeared too costly and complex to be useful for TFBL. However, a mix of self-reporting and naturalistic data collection techniques appeared to yield strong results. Knowing the best practices, we investigated how to best implement them for TFBL, with input from experts.

Andrew Bernish is a UMD professor and GIS Analyst for the Maryland Department of Transportation (MDOT). He had previously consulted with Terps for Bike Lanes to develop

their current ArcGIS crash report form. He stressed that the original design and intent of that software was for internal use to organize the crash data that TFBL already had access to, not be a publicly facing and widely used reporting form. He further noted that formatting changes would be easy for someone experienced in the field but would still require organized consulting and coordination (A. Bernish, personal communication, April 24, 2024).

The team also worked with Jeremiah Lowery, the Washington Area Bicyclist Association's (WABA) Director of Advocacy. WABA had recently developed its own crash reporting form, and Lowery shared the process for developing their crash reporting survey web page, the results and benefits in using it, maintaining the website, and how the survey fits into their overall data collection strategy.

Lastly, Dr. Alana Hackshaw, UMD School of Public Policy's Chief Diversity and Inclusion Officer, explained how surveys are conducted at the University, often in a decentralized way.

Based on these interviews, the team developed strategies that would result in better data collection for Terps for Bike Lanes. Each strategy can be integrated independently with its own set of advantages and challenges, while also using TFBL's different tools and opportunities. Knowing that the TFBL reporting form was developed for internal data organization and collection, we recommend continued use of this form but strategically for independently run site studies of pre-identified high and low-risk parts of campus. A new public-facing crash reporting form could be modeled on the WABA form and should be developed from the ground up using the Qualtrix survey development software.

We also recommend developing and conducting student experiential surveys each semester to further understand the broader student perspective on the high biking risk parts of campus and to what degree students feel biking is feasible or safe.

Literature Review

Terps for Bike Lanes is a student organization at the University of Maryland that aims to improve the university's biking policies and infrastructure through meetings with UMD and College Park stakeholders. The success of their advocacy efforts depends on the quality and reliability of the evidence they present regarding on-campus biking incidents.

Unfortunately, they have struggled to find a way to collect data that is quantifiable and includes the details of incidents. On-campus bike incident data depends on voluntary reports from students. As a result, TFBL has struggled to determine if that data is statistically relevant. The group has also struggled to find affordable and accessible software that can collect and analyze this data (Terps for Bike Lanes, personal communication, February 6, 2024).

To answer the research question, how can reliable and statistically significant data regarding on-campus bike incidents be collected, it's necessary to explore both how to collect reliable data regarding bike incidents, and how reliable data can be collected by surveying college students.

In the field of bike safety, three data collection techniques have been identified: self-report studies, naturalistic studies, and video analysis (Ibrahim et.al., 2021). Self-report studies use information such as survey data and police crash reports.

Wisconsin's Department of Transportation developed a new and significantly more detailed form for collecting crash data and found that they could better identify factors contributing to pedestrian and bicycle crashes (Wisconsin DOT, 2021). However, this more detailed form resulted in inconsistencies, as one location or factor could be described many ways. The added detail also posed issues of statistical significance because the sample size wasn't sufficiently large. There was a tradeoff between collecting valuable additional information and increasing the work collecting and analyzing the data.

Police crash reports have also been used to identify potential high-risk areas for pedestrians and bicyclists (Tao et al., 2021). Police crash reports have consistent variables and are relatively easy to transform into numerical figures. However, police crash reports alone don't provide a complete picture of this issue due to an underrepresentation of the frequency of crashes, the underreporting of low-impact or low-injury crashes, and no accounting for near-misses or crash risk (Chaurand & Delhomme, 2012).

Police crash reports also fail to provide enough detail regarding the nature, direction, and impact of bike incidents (Lusk et al., 2015). Given these limitations, police crash reports have been supplemented by researchers interacting directly with subjects, such as focus groups and expert interviews. (Cohen et al., 2016; Dobbs et al., 2009; Ibrahim et.al, 2021). Combining these methodologies allows researchers to gather quantitative data in crash reports and detailed insight offered by direct contact with people.

Technologies have been developed or proposed to collect bicycle incident data. One prototype would be attached to a bicycle to detect and record any crashes (Lin et. al, 2019). Another would use smartphones for voice-assisted crash reporting (Williams, 2018). GPS and bike-share system data have been used to analyze riding patterns and how environmental factors impact them (Luan et al., 2020). Video analysis has been proposed to capture near-crashes and misses and to avoid individual interpretations in self-reports (Ibrahim et.al, 2021).

However, these methodologies require substantial data processing, particularly if multiple cameras are used, which limit feasibility and scalability. The use of site sensors presents similar challenges; the existing technology makes it difficult to compare data from different locations and environments (Ibrahim et al., 2021). While using different technologies and video analysis can provide a new perspective and would eliminate human error in reporting, the costs and labor associated with their use make it difficult to implement.

Though it is helpful to understand how data regarding bicycle incidents and crashes is collected in general, it is also important to consider how relying on an undergraduate student population for this data would impact its collection and usage. We must understand what makes a college student population unique when it comes to data collection, what the challenges are with relying on self-reported data from college students, and how these challenges can be mitigated to ensure that the final data collected is reliable and statistically significant.

In survey research, nonresponse bias is a risk factor in the reliability of survey data; it's assumed that high response rates are necessary to achieve unbiased results (Fosnacht et al., 2017). But the impact of nonresponse bias, defined as "how much survey responders and nonresponders differ on survey variables of interest," may not be as severe as previous researchers have assumed (Keeter, as cited in Fosnacht et.al, 2017, p. 246). *The Review of Higher Education* compared simulated low response rate survey estimates and high response rates and full sample estimates, finding that survey estimates were reliable at a response rate of 5-10% given a large sampling of at least 500 students and at a response rate of 20-25% for smaller sampling frames (Fosnacht et.al, 2017).

Additionally, higher response rates were found to have a trivial impact on reliability once a survey had received 50 to 75 responses. Similarly, a University of Michigan-Dearborn study examined the impact of low response rates on the accuracy of student course evaluations and found that a 50% response rate was sufficient when the class size was greater than 50 (He and Freeman, 2021). Responses to course evaluations are more likely to have a skewed rather than normal distribution and are more prone to bias. However, the deviation of an observed teaching evaluation score from the true teaching evaluation score was reduced by an average of 0.02 for every 10% increase in response rate, which suggests an improvement in accuracy. Overall, lower response rates appear to be capable of yielding accurate and reliable results, even in cases where the collected data has a high likelihood of being biased.

Research and survey design is an additional factor to consider. When selecting a research method it's important to weigh what each method can provide and the costs of using them.

At universities, surveys appear to be the most frequently used method of collecting quantitative data. Research into data collection and effective survey methodology on college campuses regarding on-campus incidents has been limited. Most of the literature has focused on sexual assault cases and reports on college campuses. While sexual assault incidents and bicycle crashes aren't comparable in their nature or severity, the literature on effective and accurate sexual assault data collection can still provide relevant insight for our research. The difficulties that TFBL has identified in their data collection—underreporting of bicycle crashes, low response rates on optional surveys, and selection bias among respondents—are the same ones that colleges have identified in their collection of sexual assault data.

In cases where there may be a gap between the actual number of incidents and the number reported, surveys can help researchers better understand the scope of an issue. A Stockton University study used a mixed methods approach to identify sexual assault concerns on campus. Comparing data from survey responses to campus police and Clery Act reports, that study found that 10% or less of sexual assault or harassment incidents were formally reported (Shah & Gu, 2020).

Many of the reporting barriers for campus sexual violence survivors may also exist for victims of bicycle incidents; surveys can help researchers fill in data gaps. Universities typically employ census sampling, sending surveys to every student, which is more cost-effective and inclusive than random sampling (Dillman, as cited in Jeffrey et. al, 2022). While random sampling is

typically considered the “gold standard” for data gathering, it’s costly and requires more administrative labor than census sampling (McMahon, as cited in Jeffrey et al., 2022).
A comparison of sexual violence rates on campus from census-sampled climate surveys gave a response rate of 12.8%, and a randomly sampled survey gave a response rate of 53.7%. Both surveys were found to produce similar incident rates despite significant differences in response rate and the respondent demographics (Jeffrey et. al, 2022).

The literature review demonstrates a variety of data collection methods for bicycle incidents and safety, as well as important characteristics and considerations for conducting research on university campuses. The insights gained from the two different research areas is valuable in considering the methods that can be used by Terps for Bike Lanes.

Findings and Discussion

To better understand best practices for collecting and organizing bike incident data for Terps for Bike Lanes, we interviewed relevant actors in cycling and data collection: bike advocacy organizations, university staff involved in student data collection, local police, and faculty involved in data analysis and Geographic Information Systems (GIS).

In the interview with Andrew Bernish, professor and GIS Analyst for the Maryland State Department of Transportation (MDOT) he noted that his original goal and his understanding of previous work with TFBL was to design ArcGIS reporting software that could be used internally to manage the crash data they had already collected. The crash reporting site was not intended to be public facing, only to assist in the aggregation of existing data.

He believes reorienting the crash reporting website isn’t particularly difficult, a GIS professional could likely implement necessary changes “in a few hours,” but that he is unable to be a consultant for these changes (A. Bernish, personal communication, April 24, 2024). He suggested using the UMD GIS Department or Library to accomplish software and design changes, which could complete the required work product at a reasonable cost. However, there are limitations to altering the ArcGIS webpage, since problems in UI are a part of using ArcGIS.

The Washington Area Bicyclist Association (WABA) recently developed and released a bike crash reporting form. An interview with Jeremiah Lowery, WABA’s director of advocacy covered the process of developing a new site, the kind of work that may be needed to make the transitions TFBL were considering, the results WABA has achieves, and the survey’s role in WABA’s overall data collection and advocacy strategy.

Noting the changes necessary for the functionality of TFBL highlighted those elements of WABA’s crash report form, particularly improved functionality of the interactable crash map and reprogramming prompt boxes to preset multiple-choice variables, that would be relevant.

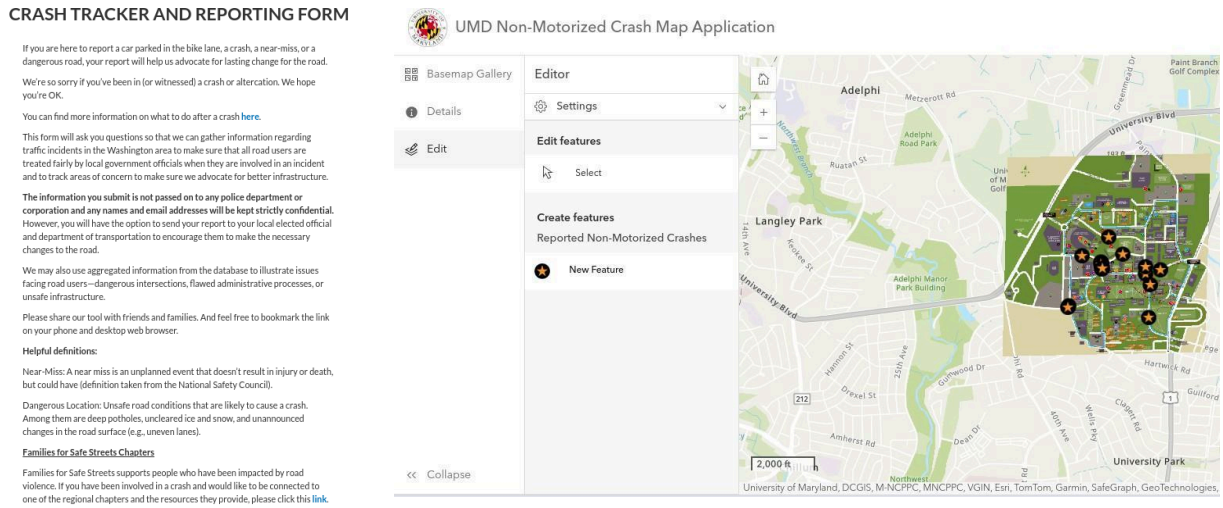
The WABA crash reporting website was developed using a professional contractor, with any necessary edits or updates now handled by Lowery. The full cost of website development was

approximately $12,000, which doesn't include maintenance or updates that expand the jurisdictions covered by the reporting form or costs of analyzing the collected data (J. Lowery, personal communication, April 24, 2024). Lowery mentioned that maintenance and changes are, considering he is not trained in web development or programming.

According to Lowery, WABA's crash reporting survey has netted significant results with strong use. Even incidents less likely to be reported, like "near-misses," are being reported with strong frequency. However, he also emphasized that the survey alone doesn't provide all the feedback they're after (J. Lowery, personal communication, March 29, 2024). The survey has become a point of emphasis when WABA engages in other forms of on-the-ground advocacy. In classes or events, they advertise the survey and reiterate its importance in reporting crashes, near misses, dangerous biking locations, or parked cars.

Additionally, WABA's reporting website includes well-written instructions and definitions for users to ensure accurate reporting. Lowery felt that part of WABA's success comes from the website and in-person work, both of which effectively communicate user instructions and information (J. Lowery, personal communication, March 29, 2024).

Figure 1. WABA crash tracker and reporting form



Source: Washington Area Bicyclist Association. (2024, April 6). https://waba.org/crashtracker/, Bernish, A. (n.d.). UMD Non-Motorized Crash Map Application. Uofmd.maps.arcgis.com. https://uofmd.maps.arcgis.com/apps/instant/sidebar/index.html?appid=6a009d7ce55d4d00bded43372c523cf5

The team had also communicated with the University of Maryland Police Department about their reporting process and dataset. The data they could provide is organized using a standard process across precincts and was information we have access to, the variables and relevant information obtained after a crash.

The Chief Diversity and Inclusion Officer for the School of Policy, Dr. Alana Hackshaw, shared that survey and student experience investigation methodologies are inconsistent across

departments. For example, the Community and Belonging Survey is conducted by a professional firm contracted by UMD, while other surveys are developed and distributed internally by individual departments.
In further discussion with both Bernish and another group consulting with TFBL on creating a Bike Ambassador Program, we learned of Bernish's current contracted work with the City of College Park in developing a city-wide bike crash reporting form. Like us, he had examined WABAs reporting form and used it as a model for his own (A. Bernish, personal communication, May 8, 2024).

Recommendations

The following three recommendations are separate but related data collection strategies that Terps for Bike Lanes can implement. Bernish's crash reporting form has similar specifications, and we recommend further communication with him and the City of College Park to adopt that form. A second strategy will allow TFBL to advertise and distribute this form, and a third strategy is data collection through an even broader student experiential survey.

Adopt the College Park Reporting Form

Given TFBL's constraints as a student-led organization, developing an effective crash reporting form would require significant time and money. It would be more efficient to adopt the crash reporting form nearly completed by Professor Bernish, with whom TFBL previously worked and would cover the City of College Park. Bernish's form is modeled on the same form we would have recommended with the added benefit of tracking beyond campus, allowing potential identification of off-campus student commuters. Additionally, adopting this form would eliminate significant data redundancies.

However, adopting this form won't collect the data TFBL would like for their advocacy. The form must become a promotional centerpiece for TFBL. Like WABA, there must be a consistent message about the form's utility. As Lowery stated, in events, information sessions, classes, and WABA advocacy, the crash reporting form is mentioned and advertised (J. Lowery, personal communication, March 29, 2024). This promotion could be further supplemented and supported by the Bike Ambassadors program TFBL is developing.

The additional benefit of using Bernish's form is the potential assistance in aggregating and organizing datasets stemming from a professional grade reporting form. While it's unlikely the data analysis will be done for TFBL, the ability to be a part of a city-wide data collection process will likely give access to better data and better consultation.

Adopt a Holistic Data Approach

As mentioned in the Literature Review, in addition to adopting and distributing a crash reporting form for quantitative data, other mixed methods of collecting campus data can be used to supplement the drawbacks of exclusively using a self-report form. Interviews and focus groups should be strategized to collect supplementary data and advocacy material. Cyclists perceive

infrastructure risks and identify areas of low and high safety risk.

“Perceived risk” can impact ridership as much as areas where danger is objectively higher. Interviewing cyclists, conducting focus groups of campus cyclists, and allowing “expert” input from a Bike Ambassadors Program could all assist TFBL in identifying areas of campus that are considered high or low risk, and gathering insight into why areas of campus are perceived that way.

Survey Students each Semester

Retooling the crash reporting as a more generalized “campus traversal experience” survey would gather more transportation information and create a more robust dataset. The survey would prioritize biking, but ask students about their usual mode of transportation, where they experience difficulties on campus, where they perceive higher risks to cyclists, and their interest in a follow-up interview.

Map interface changes would be required, but rather than updating the survey website to be better equipped to track crashes or near-crashes, it would gather information from a broader student body about the perceived risk of different parts of campus. As mentioned in the literature review, perceived risk is representative and has a similar impact as actual risk, providing an evidence base for policy changes.

This type of survey would require coordination with the UMD Department of Transportation, which can widely disseminate student surveys. During a Fall or Spring Semester, TFBL would open the survey to students during a chosen time coordinated with the University or departments for adoption and dissemination of the survey.

Regarding feasibility, retooling would require similar consulting as would changing the crash reporting website, but the follow-up and dissemination process would require a different kind of work to manage effectively. TFBL would need to coordinate with the UMD Department of Transportation, and other University departments, to distribute the survey each semester. However, once that work is done, TFBL would only be responsible for maintaining the website and helping advertise the survey, rather than the alternative constant advocacy and an integration strategy.

Conclusion

To best advocate for policy changes supporting bike safety and infrastructure on campus, Terps for Bike Lanes needs strong evidence, namely, the ability to find locations that are the greatest risk to cyclists. As the number of cyclists increases, finding ways to reduce the risks of biking on campus creates a positive feedback loop to increase biking viability and popularity.

Currently, TFBL struggles to collect the data required using only their ArcGIS report form. A more holistic and defined data collection and aggregation strategy using a mixture of quantitative and qualitative data could provide this evidence base.

Three separate but related strategies can accomplish TFBL's goal:

- integrate ArcGIS into internal site study data collections built on focus group interviews
- develop a qualtrics reporting form built from the ground up to be public facing, modeled on WABA's Crash Reporting Form
- develop a Qualtrics student experiential survey, using WABA's form as a model, but sampling the broader student body for areas they define as high risk.

References

Chaurand, N., & Delhomme, P. (2012, September 25). Cyclists and drivers in road interactions: A comparison of perceived crash risk. Accident Analysis & Prevention. https://www.sciencedirect.com/science/article/pii/S0001457512003120?casa_token=S5SEct5rTrAAAAAA%3Al4cdoViwZrjgcrFtK_l30CFA9KLibuJKDWM88WAcXp9tVMkg VeqTHNt9KLnBZ_5x3A3Ayz2g_Q#bib0290

Dobbs, Gabe, "Pedestrian and Bicycle Safety On A College Campus: Crash and Conflict Analyses With Recommended Design Alternatives For Clemson University" (2009). All Theses. 552. https://tigerprints.clemson.edu/all_theses/552

Fosnacht, K., Sarraf, S., Howe, E., & Peck, L.K. (2017). How Important are High Response Rates for College Surveys?The Review of Higher Education 40(2), 245-265. https://doi.org/10.1353/rhe.2017.0003.

He, J., & Freeman, L. A. (2021). Can we trust teaching evaluations when response rates are not high? implications from a monte carlo simulation. Studies in Higher Education, 46(9), 1934–1948. https://doi.org/10.1080/03075079.2019.1711046

Ibrahim, M. R., Haworth, J., Christie, N., Cheng, T., & Hailes, S. (2021). Cycling near misses: a review of the current methods, challenges and the potential of an AI-embedded system. Transport Reviews, 41(3), 304–328. https://doi-org.proxy-um.researchport.umd.edu/10.1080/01441647.2020.1840456

Jaggars, S. S., Fletcher, J., Stacey, G. W., & Little, J. M. (2014). Simplifying complexity in the student experience: Gathering data. New York, NY: Columbia University, Teachers College, Community College Research Center https://ccrc.tc.columbia.edu/media/k2/attachments/simplifying-complexity-gathering-data.pdf

Jeffrey, N. K., Senn, C. Y., Krieger, M. A., & Forrest, A. (2022). Sexual violence among postsecondary students: no evidence that a low response rate biases victimization or perpetration rates in a well-designed climate survey. *Journal of Evidence-Based Social Work*, *19*(5), 537–556. https://doi.org/10.1080/26408066.2022.2086443

Lin, B., Marthur, D. M., & Tam, A. (2019, October 23). Bike Crash Detection. The Grainger College of Engineering. https://courses.engr.illinois.edu/ece445/getfile.asp?id=16233]

Louisiana Transportation Research Center. (2019, August). PEDESTRIAN AND BICYCLE COUN COLLECTION AND USE: A GUIDE FOR LOUISIANA. https://www.ltrc.lsu.edu/pdf/2019/Appendix%20D.pdf

Luan, S., Li, M., Li, X., & Ma, X. (2020, June 13). Effects of built environment on bicycle wrong way riding behavior: A data-driven approach. Accident Analysis & Prevention. https://www.sciencedirect.com/science/article/pii/S0001457519314241?casa_token=w

G0sI4j1VOEAAAAA%3Avd36ON1xsaJNHHJEhHGC6RMbHV2vxpOlXp9oHfQc8fvlRUc 49vOR5U2Jc6envQ8jX-DJWyL79A#bbib0230
Lusk AC, Asgarzadeh M, Farvid MS Database improvements for motor vehicle/bicycle crash analysis Injury Prevention (2015); 21:221-230. https://injuryprevention.bmj.com/content/21/4/221
Martin, E., Cohen, A., Botha, J., & Shaheen, S. (2016, March). Bikesharing and Bicycle Safety. Bikesharing and Bicycle Safety. https://dot.ca.gov/-/media/dot-media/programs/research-innovation-system-information/documents/f0016782-ca16-2635-finalreport.pdf
Tao, T., Lindsey, G., Cao, J., & Wang, J. (2021, May 11). The effects of pedestrian and bicycle exposure on crash risk in Minneapolis. Journal of Transport and Land Use. https://www.jtlu.org/index.php/jtlu/article/view/1980
Williams, B. (2018, June). Bicycle Crash Detection: Using A Voice-Assistant for More Accurate Reporting. Graduate School of the University of Oregon. https://www.cs.uoregon.edu/Reports/MS-201806-Williams.pdf
Wisconsin DOT. (2021, September). Comprehensive evaluation of DT4000 data quality for pedestrian and bike crashes. Institute For Physical Infrastructure and Transportation. https://uwm.edu/ipit/projects/comprehensive-evaluation-of-dt4000-data-quality-for-pedestrian-and-bike-crashes/