The Canadian Geographer
Le Géographe canadien
Canadian Association of Geographers
Association Canadienne des Géographes

# Spatial variation in bicycling risk based on crowdsourced safety data

Jaimy Fischer
Faculty of Health Sciences, Simon Fraser University

Stephanie Sersli
Faculty of Medicine and Health Sciences, Université de Sherbrooke

Trisalyn Nelson
Department of Geography, University of California Santa Barbara

Hanchen Yu
The Center for Geographic AnalysisHarvard University

Karen Laberee
Department of Geography, University of Victoria

Moreno Zanotto
Faculty of Health Sciences, Simon Fraser University

Meghan Winters
Faculty of Health Sciences, Simon Fraser University

## Key Messages

- Falls and single bicycle incidents with infrastructure, roads, and railroads are relatively frequent incidents and lead to injury.
- Crowdsourced data can supplement official reports to improve bicycling safety analysis.
- Narrative and contextual details help inform interventions that can reduce injury for bicyclists.

*Bicycling-related injury data are difficult to obtain from official reports, which capture only about 20% of crashes and often lack coordinates, injury outcomes, and narratives needed for understanding where and why incidents occurred. Crowdsourced data on bicycling safety provides new opportunities for the study of bicycling injury and risk. Our goal was to quantify factors that influence the spatial variation in unsafe bicycling across a city, based on self-reports of bicycling incidents. To meet this goal, we leveraged BikeMaps.org, a global tool for reporting bicycling safety incidents, drawing on data from Metro Vancouver. We summarized incident conditions that led to injury, developed a model to identify predictors of injury using random forest regression, and mapped bicycling incident hot spots. Our results demonstrate that injuries from bicycling incidents are associated with older and younger bicyclists, downhill slopes, parked cars, recreation and weekend rides, falls, and single bicycle incidents with infrastructure, roads, and railroads. The broad range of incidents reported to BikeMaps.org allows us to add evidence that falls and single bicycle collisions are major causes of injury. Also, we demonstrate the value of attributing safety hot spots with contextual details to identify infrastructure interventions that can reduce injury for bicyclists.*

Correspondence to / Adresse de correspondance: Trisalyn Nelson, Department of Geography, University of California Santa Barbara, Santa Barbara, Ellison Hall, Isla Vista, CA 93117, USA.
Email/Courriel: trisalyn@ucsb.edu

*The Canadian Geographer / Le Géographe canadien* 2022, 1–13
DOI: 10.1111/cag.12756

Keywords: crowdsourced data, bicycle safety, GIS, spatial analysis, active transportation

**Les variations spatiales des risques associés au cyclisme à partir des données de type «*crowdsourcing* »**

*Les données sur les blessures reliées à la pratique du vélo sont difficiles à obtenir des rapports officiels qui ne recensent qu'environ 20% des collisions. De plus, il manque souvent dans ces rapports les coordonnées des lieux de la collision, la nature des blessures et le détail des faits requis pour comprendre où et pourquoi les incidents se sont produits. Les données de type « crowdsourcing » sur la sécurité du cyclisme offrent de nouvelles possibilités pour l'étude des risques et des blessures reliés à la pratique du vélo. L'objectif de cette recherche est de quantifier les facteurs qui influencent la variation spatiale des risques associés à la pratique du vélo dans une ville selon les propres déclarations des personnes impliquées dans des incidents. Pour atteindre cet objectif, nous avons misé sur BikeMaps.org, un outil reconnu pour signaler les incidents à vélo, en focalisant sur les données du grand Vancouver. Dans cette optique, nous avons compilé les caractéristiques des incidents, puis développé un modèle de type arbre de décision pour identifier les variables explicatives des blessures et, finalement, nous avons cartographié les points chauds pour les incidents à vélo. Nos résultats démontrent que les blessures en vélo touchent davantage les plus jeunes et les plus vieux cyclistes, elles se localisent dans des pentes descendantes, elles impliquent des voitures stationnées, elles se font lors de balades récréatives et de fin de semaine, elles sont reliées aux chutes, aux infrastructures ainsi qu'aux routes et voies ferrées. Le large éventail d'incidents signalé sur BikeMaps.org nous permet d'ajouter des données probantes sur le fait que les chutes et les collisions individuelles des cyclistes sont des causes majeures de blessures. Au final, nous démontrons la valeur de l'identification des points chauds incluant le contexte de l'incident pour identifier des interventions visant les infrastructures routières qui peuvent réduire les risques et blessures des cyclistes.*

Mots clés : données de type « crowdsourcing », sécurité à bicyclette, SIG, analyse spatiale, transport actif

## Introduction

Bicycling is shown to improve health and relieve urban congestion (Oja et al. 2011; Götschi et al. 2016; Zahabi et al. 2016). Cities around the world are promoting bicycling as a healthy and sustainable form of travel (Buehler and Pucher 2012; Fishman 2016; Pucher and Buehler 2017; Fan et al. 2019; Lee and Pojani 2019; Rosas-Satizábal and Rodriguez-Valencia 2019; Soliz 2021). However, safety concerns pose barriers, especially where people riding bicycles share the road with people driving vehicles (Heinen et al. 2010; Pucher et al. 2010). To overcome safety concerns and increase bicycling, cities are investing in bicycling-specific infrastructure, as separating bicyclists and drivers has been shown to improve real and perceived bicycling safety (Sanders 2015).

Lack of bicycling data, including bicycling safety data, is a challenge when aiming to implement pro-bicycling policy (Winters and Branion-Calles 2017). Traditional bicycling safety data, which include data from police reports, insurance claims, and hospital records, are biased toward crashes that include vehicles and account for only ~20% of all bicycling incidents (Winters and Branion-Calles 2017). These official sources under-sample incidents, including falls and crashes, that do not involve vehicles, incidents on multiuse paths, and near misses.

Crowdsourced datasets are filling gaps in bicycling data, and provide enhanced data on bicycle ridership, infrastructure, perceptions and attitudes toward bicycling, and safety (Nelson, Ferster, et al. 2021). Crowdsourced data on bicycling safety have been shown to increase sampling of incidents in high bicycling ridership areas (e.g., multiuse paths), to capture single bicycle collisions and falls, and are typically the only source of near miss data (Mannering and Bhat 2014; Nelson et al. 2015). As such, crowdsourced safety data can be useful for spatial comparisons of bicycling incidents and can help prioritize investments that will improve bicycling safety. However, effective use of bicycling safety data also requires data on bicycling ridership, as it is important to present safety relative to exposure, which in the case of bicycling is the number of people riding bicycles at a particular place and time.

New approaches to data collection enabled by smartphone technology have expanded possibilities for incident reporting to a much larger sample of bicyclists and a wider variety of incidents. Websites and mobile apps provide a forum for individuals to report locations and details of safety-related incidents (Nelson, Ferster, et al. 2021). Crowdsourcing bicycling safety incidents, which include minor and near miss reports, provide early warning of preventable incidents and increase awareness of risks that can lead to serious incidents (Aldred 2016). Furthermore, minor incidents are also important as they can negatively impact people's perceptions of bicycling, hampering efforts to increase bicycling as a viable mode of transport for the population at large (Aldred and Crosweller 2015). These new data sources can supplement the gaps of traditional datasets, for example by providing more granular incident details and increasing the amount of data available for safety research.

Our goal was to quantify factors that influence the spatial variation in unsafe bicycling across a city based on crowdsourced self-reports of bicycling incidents. To meet this goal, we drew upon data from BikeMaps.org, a global tool for reporting bicycling safety incidents (Nelson et al. 2015) in Metro Vancouver, British Columbia. First, we summarized the proportion of all incidents and those that led to injury by the characteristics of people and places associated with incident; second, we developed a model to identify important predictors of injury using random forest regression; and third, we mapped hot spots of risk using kernel density estimation and labelled hot spots using narrative descriptions from self-reports.

## Methods

### Study area

Our study is set in Metro Vancouver, British Columbia (Figure 1). Metro Vancouver is a federation of 21 municipalities (including the City of

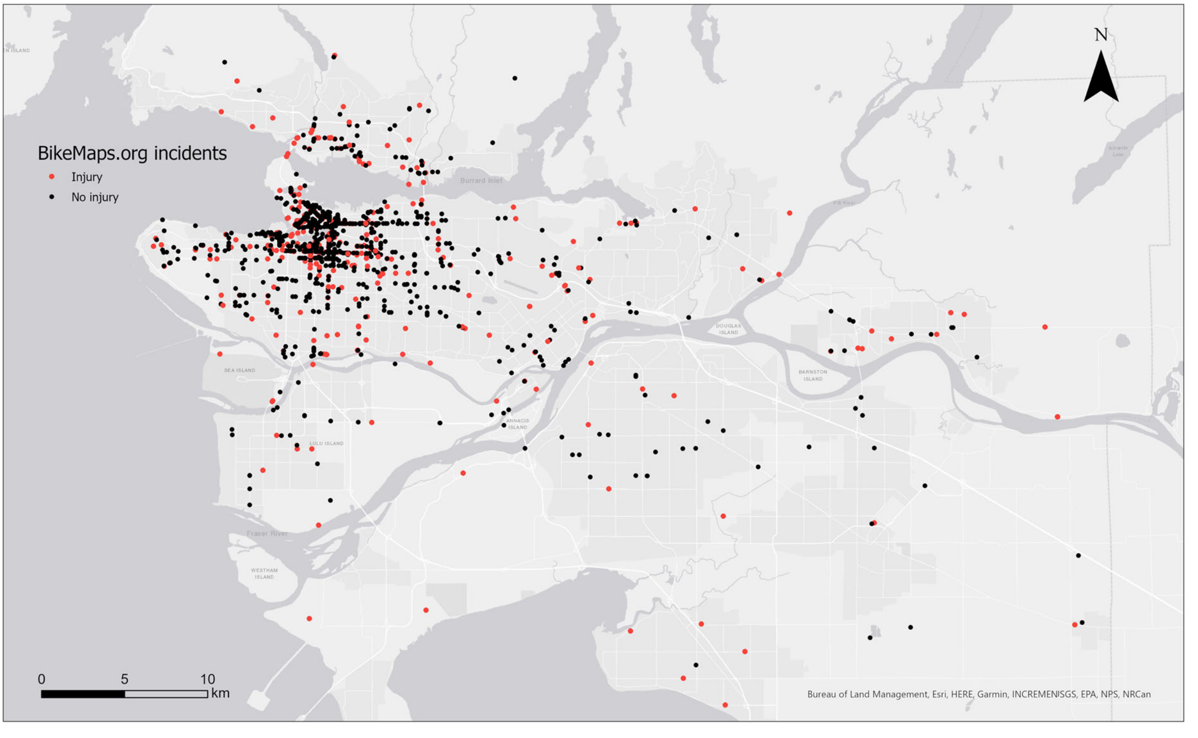


**Figure 1**
Study area and BikeMaps.org incidents.

Vancouver), one Treaty First Nation (Tsawwassen First Nation), and one electoral area with diverse urban form and transportation infrastructure (Metro Vancouver 2021). Metro Vancouver has promoted itself as a place for healthy living, with an emphasis on fitness and an active outdoor lifestyle. Its temperate climate (average 4–18 ℃ (Environment Canada 2021)) is conducive to year-round bicycling. Across Metro Vancouver, the bicycle journey-to-work mode share is above average for Canada, with an overall rate of 2.3% compared to the national rate of 1.4% (Statistics Canada 2019). Across municipalities, bicycling rates range from <0.5% in less central and more sprawling areas to 6.1% in the City of Vancouver (Statistics Canada 2019).

Bicycle journey-to-work mode share in the City of Vancouver has nearly doubled from 3.3% in 1996 to 6.1% in 2016 (Statistics Canada 2019). Ridership doubling reflects the positive outcomes of investments by the city to support bicycling as a mode of transportation in recent decades. In 2020, Vancouver's total bicycle route network was 316 km, comprised of local street bikeways (shared roadway along local streets, typically traffic-calmed, 57%); cycle tracks (roadway lane exclusively for bicyclists, physically separated from motor vehicles and the sidewalk, 10%); multiuse paths (two-way paved path shared by bicyclists and pedestrians, 20%); and painted bike lanes (painted lane along a busy roadway, 13%) (City of Vancouver 2021; Firth et al. 2021). The bicycle route network is the densest in downtown and adjacent neighbourhoods, which corresponds to neighbourhoods with the highest population densities in the region.

## BikeMaps.org data

We used a dataset of nearly 1,300 bicycling incidents (collisions, falls, and near misses) reported to BikeMaps.org over a six-year period (2014–2020, Figure 1). Launched in 2014, BikeMaps.org is a website and mobile app where bicyclists can map crashes, falls, and near misses. For each incident, reporters identify a location using an online map and complete questions using dropdown options in three categories: incident details, conditions, and personal details. Extensive details on the BikeMaps.org reporting template have been reported elsewhere (Nelson et al. 2015). In brief, Table 1 lists questions posed to users. Mandatory questions include "Were you injured?" to which reporters answer "yes" or "no," as well as details related to the incident (e.g., type of incident, when the incident occurred). Reporters to BikeMaps.org may also provide detailed incident descriptions (up to 300–500 characters) in an open-ended text section. We reviewed all narratives to ensure data were correctly classified.

## Explanatory variables

To identify important predictors of bicycling injury, we considered 18 explanatory variables (Table 1). The majority (n = 17) were collected or derived from BikeMaps.org reports, and we included an additional variable representing average daily bicycle volumes at incident locations. To represent bicycling volumes, which are important for quantifying exposure to potential bicycling incidents (Ferster et al. 2021), we used Strava Metro data generated from the popular Strava fitness tracking app. Strava is used by people to track bicycling activities and data are provided free of charge to partner cities using the data for transportation decision making. Strava data have gained momentum as a stand-in for the pervasive lack of exposure data available from traditional bicycle count programs and have been used in a range of bicycle monitoring activities, including to control for exposure (Lee and Sener 2021). Strava data are biased toward people that use the Strava app, although researchers have shown strong correlations between Strava data and all bicycling levels (Jestico et al. 2016; Conrow et al. 2018; Lee and Sener 2021). To reduce the impact of bias while benefiting from the continuous spatial data, we use Strava categorially and define five ridership categories representing average daily ridership: very low (<6), low (6–55), medium (56–230), high (231–4,305), and very high (>4,305). These categories are useful for urban planning purposes and five classes is appropriate given the precision of Strava data (Nelson, Roy, et al. 2021).

## Analysis

We summarized incident characteristics for BikeMaps.org reports within Metro Vancouver from 2014 to 2020 using all incidents and only incidents that led to injury. We calculated the number of BikeMaps.org reports by gender, age,

**Table 1**
Explanatory variables.

| Variable | Completeness in BikeMaps.org Metro Vancouver data (%) | Variable attributes |
|---|---|---|
| **Incident details** | | |
| Were you injured? (outcome) | 100% | Injury \| no injury |
| What type of incident was it? | 100% | Collision with moving object or vehicle \| collision with stationary object or vehicle \| fall |
| What sort of object did you collide with? | 100% | Animal \| another bicyclist \| curb \| other \| pedestrian \| pothole \| roadway \| sign/post \| train tracks \| lane divider \| vehicle, angle \| vehicle, head on \| vehicle, open door \| vehicle, passing \| vehicle, rear end \| vehicle, side \| vehicle, turning left \| vehicle, turning right |
| What was the purpose of your trip? | 88% | Commute \| exercise or recreation \| personal business \| social reason |
| Season incident occurred (When was the incident?) | 100% | Winter \| Spring \| Summer \| Fall |
| Day incident occurred (When was the incident?) | 100% | Weekday \| weekend |
| Time of day incident occurred (When was the incident?) | 100% | AM peak (6–9am, weekdays) \| PM peak (3 – 6 pm, weekdays) \| inter-peak hours |
| **Conditions** | | |
| What were the road conditions? | 77% | Dry \| wet \| loose sand or gravel \| icy \| snowy |
| How were the sight lines? | 76% | No obstructions \| view obstructed \| glare or reflection \| obstruction on road |
| What was the terrain like? | 77% | Uphill \| downhill \| flat |
| How were you moving? | 77% | Heading straight \| turning left \| turning right |
| Were there cars parked on the roadside? | 72% | Yes \| no |
| **Personal details** | | |
| Age category (What is your birth year?) | 61% | Under 30 \| 31 – 50 \| 50+ |
| Gender (Please select your sex) | 63% | Male \| female \| other |
| Do you bike at least once per week? | 67% | Yes \| no |
| Were you wearing a helmet? | 66% | Yes \| no |
| Were you using bike lights? | 71% | Front and back lights \| front lights only \| back lights only \| no lights |
| **Average daily Strava ridership** | 100% | Very high (681–4,305) \| high (231–680) \| medium (56–230) \| low (6–55) \| very low (0–5) |

frequency of bicycling, trip purpose, type of incident and when it occurred, and the type of object that was involved in the incident. To understand what incident characteristics had a greater prevalence of injury in BikeMaps.org reports, we also calculated the proportion of all incidents in each category that resulted in injury.

We trained a balanced random forest classifier on BikeMaps.org data to predict injuries. Random forest is a supervised machine-learning algorithm which can be used to explain the relative importance and impact of each explanatory variable (Breiman 2001). It is an ensemble of decision trees, such that each tree is grown from the values of a random vector sampled independently and with the same distribution for all trees in the forest (Breiman 2001). We used random forests as they are a classification procedure that determines which variables are more relevant for splitting data into classes and provides a list of variables by level of "importance" for the classification. Variable importance measures the relative impact that each variable has on model accuracy, with top ranking variables causing the greatest decrease in accuracy when removed from the model (Branion-Calles et al. 2016; Fischer et al. 2020). Variables with negative importance decrease overall accuracy, thus we removed variables with negative importance scores from the final model. We used partial dependence plots to visualize the relationship between final explanatory variables and injury. Plots show the probability of injury for each of the values of an explanatory variable, holding all other variables constant.

When developing priorities for infrastructure investments, it can be important to consider areas with high incident density, or "hot spots." We converted bicycling incidents to hot spots using kernel density estimation (KDE). KDE is useful for visualizing the spatial variability of point data, and it works by creating a smoothed hot spot surface representing the density of events. The smoothness of the KDE surface is defined by the kernel function bandwidth, and literature suggests that testing KDE surfaces across a range of bandwidths is an accepted method for selecting one that best represents the spatial variability of the data (Atkinson and Unwin 2002; Nelson and Boots 2008). We tested bandwidths from 100 m to 500 m in 50-m increments and selected a bandwidth of 400 m because it provided distinct hot spots and best represented the spatial variability of BikeMaps.org incidents across the Metro Vancouver study area. Following Nelson and Boots (2008), we defined incident hot spots as the top 10% of the KDE values.

To better understand the context of hot spots, narrative data were examined after the random forest model and hot spot mapping were completed. We grouped narrative data by hot spot, type of incident, and type of object involved in the incident. Then we inductively coded narrative data (Thomas 2006) and summarized the common aspects for each object type and hot spot.

## Results

Summary statistics in Table 2 provide contextual information on BikeMaps.org incident reports that we relate to random forest results below. However, given that absolute counts in each category may vary widely, prevalence may not be an appropriate measure for direct comparison. Of the 1,297 BikeMaps.org incident reports for Metro Vancouver, about one quarter (26.5%, n = 343) resulted in injury. Incidents were most often reported by men (44.3%), people aged 30–49 (45.0%), and people who use bicycles at least once a week (70.1%). In Table 2, we also report the proportion of reports that resulted in injury across each characteristic. Results show that people aged 30–49 had the lowest prevalence of injury from reported bicycling incidents (19.9–20.9%). Older adults (50+) and people under 30 had more injuries from incidents (32.1% and 100.0%, respectively). Falls had a greater prevalence of injury compared to collisions (83.3% vs. 74.2%). As well, recreational bicycling (39.3%) and weekend bicycling (35.1%) had high proportions of incidents that resulted in injury, compared to other types of trips and weekday rides (see

**Table 2**
Incident characteristics for BikeMaps.org reports within Metro Vancouver, 2014–2020.

| Characteristic | Total (n) | Injury (n) | No injury (n) | Injury (%) |
|---|---|---|---|---|
| Total | 1297 | 343 | 954 | 26.5 |
| **Gender** | | | | |
| Men | 575 | 145 | 430 | 25.2 |
| Women | 294 | 72 | 222 | 24.5 |
| Non-binary | 14 | 5 | 9 | 35.7 |
| Missing | 414 | 121 | 293 | 29.2 |
| **Age** | | | | |
| <20 | 3 | 3 | 0 | 100.0 |
| 20–29 | 40 | 14 | 26 | 35.0 |
| 30–39 | 302 | 63 | 239 | 20.9 |
| 40–49 | 282 | 56 | 226 | 19.9 |
| 50–59 | 137 | 44 | 93 | 32.1 |
| 60–69 | 81 | 26 | 55 | 32.1 |
| 70+ | 14 | 7 | 7 | 50.0 |
| Missing | 438 | 130 | 308 | 29.7 |
| **Bicycle usage** | | | | |
| Weekly bicyclist, yes | 909 | 225 | 684 | 24.8 |
| Weekly bicyclist, no | 164 | 8 | 6 | 57.1 |
| Missing | 374 | 110 | 264 | 29.4 |
| **Incident type** | | | | |
| Collision | 368 | 273 | 95 | 74.2 |
| Fall | 84 | 70 | 14 | 83.3 |
| Near miss | 845 | 0 | 845 | NA |
| **Trip purpose** | | | | |
| Commute/Work | 886 | 197 | 689 | 22.2 |
| Exercise or recreation | 140 | 55 | 85 | 39.3 |
| Personal/Social | 201 | 42 | 159 | 0.2 |
| Missing | 70 | 49 | 21 | 70.0 |
| **Day of week** | | | | |
| Weekday | 1049 | 256 | 793 | 24.4 |
| Weekend | 248 | 87 | 161 | 35.1 |
| **Incident was with** | | | | |
| Vehicle | 1096 | 244 | 852 | 22.3 |
| Cyclist | 55 | 24 | 31 | 43.6 |
| Pedestrian | 47 | 10 | 37 | 21.3 |
| Infrastructure* | 24 | 20 | 4 | 83.3 |
| Roadway | 19 | 15 | 4 | 79.0 |
| Train tracks | 14 | 11 | 3 | 78.6 |
| Other | 35 | 19 | 16 | 54.3 |
| Animal | 5 | 0 | 5 | 0.0 |
| E-scooter | 2 | 0 | 2 | 0.0 |

*Curb, lane divider, pothole, and sign/post.

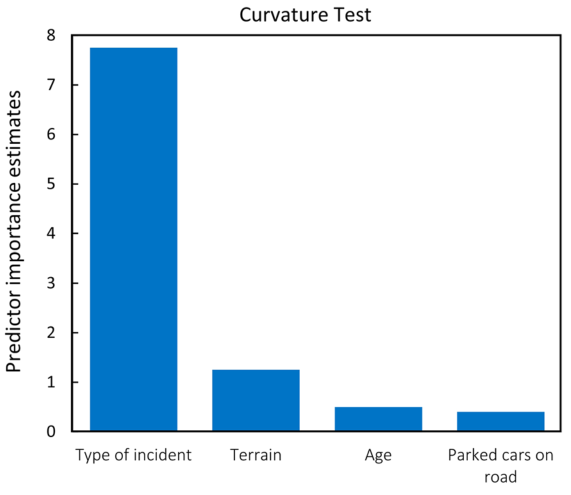


**Figure 2**
Variable importance plot. The plot shows the relative importance of each BikeMaps.org explanatory variable for correctly classifying injury outcomes. Variables causing the largest increase in predictive accuracy rank higher in importance in a random forest model. Type of incident is categorized as collision, fall, or near miss. Terrain variable is categorized as riding uphill, on flat terrain, or downhill.

Table 2). Incidents that occurred as the result of crashing with infrastructure (curb, lane divider, pothole, and sign/post), roads, and train tracks also had high proportions of injuries (78.6–83.3%).

We developed a model to identify important predictors of injury using random forest regression. Our random forest model had an overall accuracy of 92%. In order of importance for prediction, the influence of explanatory variables in the model were: type of incident (collision, fall, or near miss); terrain the bicyclist was riding on (uphill, flat, downhill); age of the bicyclist; and whether there were parked cars on the road (Figure 2). Other variables had minimal predictive power.

Partial dependence plots for the four most important explanatory variables (Figure 3) provide greater detail about the specific relationships between variables and injury prediction. Results of incident type show that collisions with stationary objects and falls both had a slightly higher probability of a resulting injury than collisions with moving objects. Injury was also more likely to be predicted when the bicyclists were on downhill terrain, older (>49 years) or younger (<30 years) in age, and on streets with cars parked on the roadside.

We identified 14 hot spots within the City of Vancouver and evaluated crash characteristics at each location (Figure 4; Table 3). The majority (64.0%) of injuries in hot spot areas resulted from collisions with motor vehicles. Report details and open-text narratives revealed that some hot spots had site-specific characteristics related to the hot spot. Four injury hot spots (E, G, K, M) were exclusively the result of incidents between bicyclists and motor vehicles, involving either vehicles turning or failing to yield. At hot spot M, all injuries involved conflicts with motor vehicles at one specific traffic circle. Driver failure to stop or yield was associated with at least 50% of injuries at two more hot spots (D and J). Conflict with turning vehicles was associated with at least 50% of injuries at two more hot spots (A, L). Only three hot spots (C, H, N) did not involve conflict with motor vehicles. Of these, hot spots H and N were hot spots of injury from collisions with train tracks. The four biggest hot spots (B, F, I, and L) had a diversity of incident types. These hot spots are characterized by high pedestrian, bicyclist, and car traffic and thus reflect greater potential for intermodal conflict.

## Discussion

Crowdsourced data provide a mechanism to study a wider range of bicycling incidents than official reports, including incidents that results in near misses, minor injuries, or are bike-only incidents (e.g., falls). Another benefit of crowdsourced reporting is the ability to have reporters provide narrative details on what happened during the incident. While we do not provide a text or qualitative analysis of BikeMaps.org reports, we draw on this benefit in our discussion below by contextualizing results of the quantitative analysis using quotes and open-text narrative details.

In our study, we found age was associated with bicycling injury. People younger than 30 and older than 49 had a higher likelihood of incidents ending in injury based on both our summary statistics and random forest modelling, despite comprising a lower proportion of total BikeMaps.org reports. Our results corroborate other similar research findings on associations between older age and

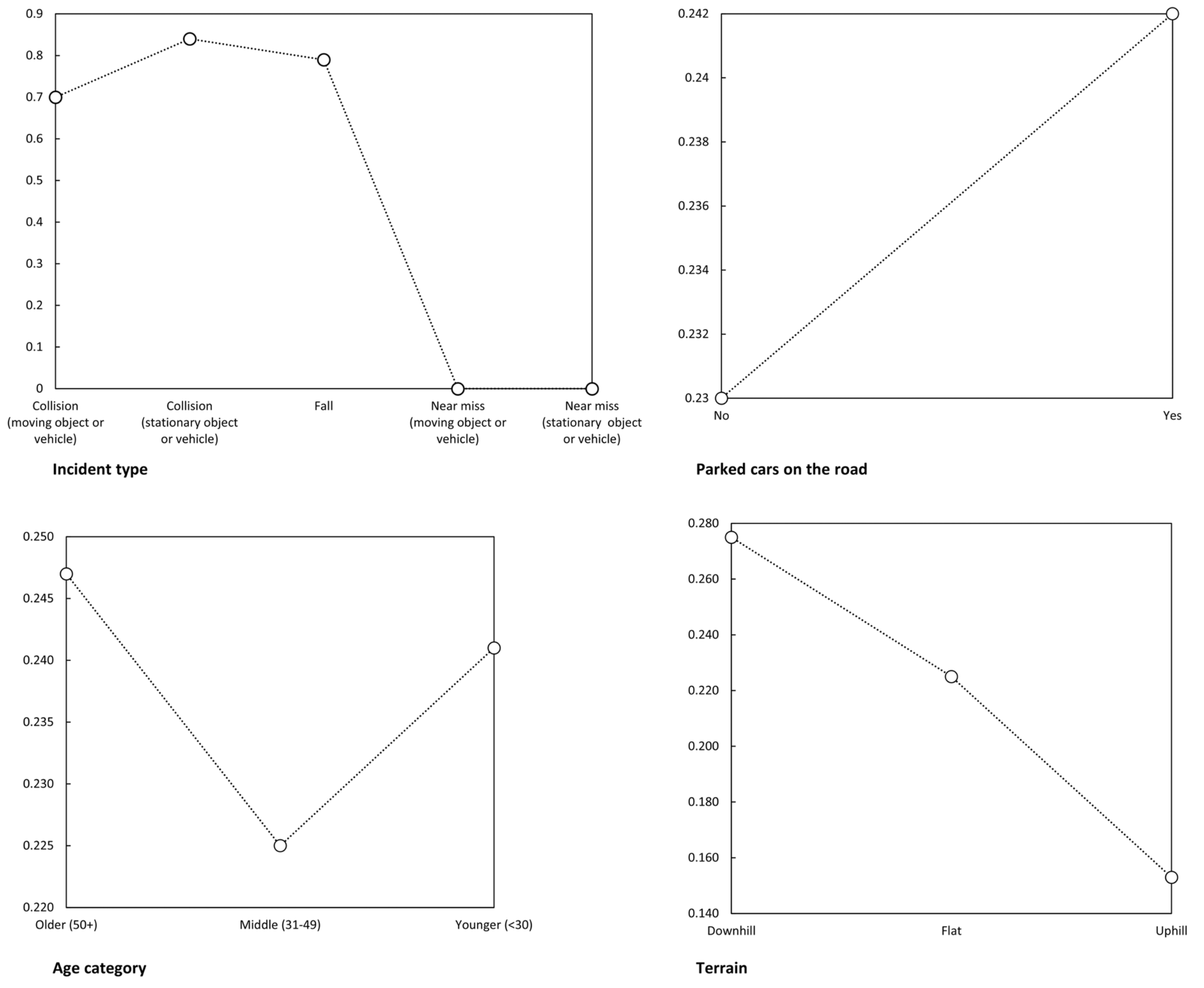


**Figure 3**

Partial dependence plots for the injury class against different variables. The y-axis is the posterior probabilities of the class, and plots are interpreted as the probability of a class prediction over the range of variable attribute values.

bicycling injury (Vanparijs et al. 2015; Prati et al. 2017; Chen and Shen 2019; Liu et al. 2020; Meuleners et al. 2020). For example, an Australian study found that although older bicyclists were less likely to experience a crash than younger bicyclists, they were more likely to be injured (Poulos et al. 2015). The higher risk of older adults could reflect that older bicyclists are more vulnerable to injury due to general aging-associated physiology. There is a lack of research on bicycling injury in young people; however, we found evidence that the overall burden of injury in this population may be underestimated as younger people may be less likely to report crash injuries through official mechanisms (e.g., police, insurance) (Watson et al. 2015). Our analysis does not uncover explanations as to why younger people had a higher likelihood of being injured, but given this population is underrepresented in bicycling injury studies we view this result as important and an example where crowdsourced data can help fill data gaps and support new research opportunities. We also found injury to be associated with topography, notably downhill slopes, and parked

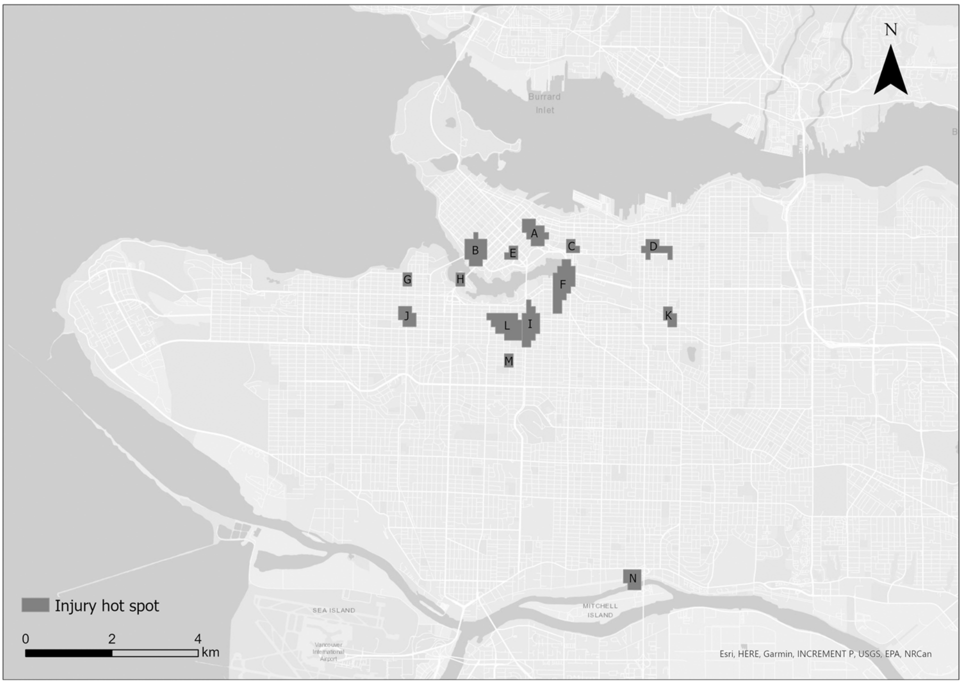


**Figure 4**
Location of 14 injury hot spots. All but one of the hot spots were within 5 km of the downtown core.

cars on the road. Downhill slopes have been linked to increased crash rates and severity in several studies (Teschke et al. 2012; Harris et al. 2013; Meuleners et al. 2019), likely due to higher speed going downhill. Vehicles parked on roads can obstruct sightlines and also increase the potential for doorings, which have been shown to be a substantive source of injuries for bicyclists (Ferenchak and Marshall 2019).

**Table 3**
Hot spot incident and injury characteristics.

| Hot spot | Number of incidents | Number of injuries (%) | Average daily Strava ridership |
|---|---|---|---|
| A | 18 | 7 (38.9) | 231–4305 (high) |
| B | 31 | 15 (48.4) | 231–4305 (high) |
| C | 5 | 6 (60.0) | 231–4305 (high) |
| D | 13 | 6 (46.2) | >4305 (very high) |
| E | 6 | 2 (33.3) | 56–230 (medium) |
| F | 36 | 21 (58.3) | >4305 (very high) |
| G | 5 | 3 (60.0) | 231–4305 (high) |
| H | 3 | 3 (100.0) | 6–55 (low) |
| I | 26 | 10 (38.5) | 231–4305 (high) |
| J | 9 | 6 (66.7) | 231–4305 (high) |
| K | 6 | 3 (50.0) | 231–4305 (high) |
| L | 34 | 11 (32.4) | >4305 (very high) |
| M | 3 | 3 (100.0) | 231–4305 (high) |
| N | 9 | 9 (100.0) | 231–4305 (high) |

Our results indicate that falls were predictive of injury. Single bicycle collisions with infrastructure (curb, lane divider, pothole, and sign/post), roads, and train tracks also had a high likelihood of injury. This adds to evidence from other studies that falls and single bicycle crashes are an under-recognized but significant cause of bicycling injuries (Teschke et al. 2014; Myhrmann et al. 2021). In part, lack of data has made it difficult to quantify the impact of falls and single bicycle incidents (Foley et al. 2020; Meuleners et al. 2020; Utriainen 2020). For example, several reports outline how transitions on and off bicycle infrastructure can be problematic. As an example, one report highlights how curbs and awkward or shallow angles of approach increase bicycling hazard: “decided to enter gas station driveway to east so could cycle to shared path. There was a 3–4 cm lip on the curb which I entered at an angle such that the curb ‘grabbed’ my tire causing me to [fall]” (report #8765, regular cyclist, personal trip, no injury treatment, woman). Bollards, posts, and infrastructure meant to manage transportation flow can also be problematic. For instance, fixed bollards intended to prevent cars from driving onto multi-use paths can be hazardous for bicyclists. “There are pillars, not well visible from bigger distance and without any horizontal warnings or signs. I was checking my gear shifting for few seconds and suddenly hit them, falling off completely uncontrolled” (report #4992, regular cyclist, commute trip, ER visit, man). Not surprisingly, collisions with motor vehicles also showed a high probability of injury consistent with many previous studies (e.g., Teschke et al. 2014; Cripton et al. 2015; Beck et al. 2016).

Several falls and single bicycle incidents were also associated with transient conditions or temporary objects on the road surface. These included rocks, chunks of concrete, and gravel, but also slippery surfaces caused by black ice and leaves (Myhrmann et al. 2021). Roadworks were mentioned in several reports where injuries occurred when a bicyclist lost control due to uneven surfaces or slipped on construction site materials: “I was forced to use the road where there's been excessive road works. My front tire slipped on the metal plates laid over the ditches dug there. There were no signs, tapes or cones to indicate or avoid the hazards” (report #4302, regular cyclist, commute trip, no injury treatment). These reports illuminate the often-overlooked range of experiences that pose safety and injury risk for bicyclists.

Bicycling incidents with motor vehicles had a slightly lower prevalence of injury. There are several potential explanations for this finding. One is the nature of reports to BikeMaps.org, which represent a broad range of bicycling incidents, but a higher proportion of these incidents are those commonly experienced (but rarely reported) by bicyclists such as near misses and crashes with minor or no injury. Near misses are the most common reports on BikeMaps.org. Past analysis has shown that near miss self-reports have a higher probability of being associated with incidents involving motor vehicles than with other incident conditions (Branion-Calles et al. 2017), likely because they are alarming and influence perceptions of risk. Without a doubt, bicycle collisions with motor vehicles are a source of significant injury risk for bicyclists. While BikeMaps.org captures many near misses and minor collisions with motor vehicles, the important outcome of this research is that it adds evidence that incident conditions beyond those involving motor vehicles also lead to injury for bicyclists. Ideally, official sources of data and crowdsourced reports could be combined in widespread injury analysis. But data access, record consistency, and the need to protect privacy make it difficult for official reports to be mapped at a street level and attributed with injury outcomes.

Through hot spot analysis we showed that incidents and outcomes vary spatially, and that there is benefit to being able to contextualize clusters of high incidents using detailed narratives. For example, although traffic circles and roundabouts were only mentioned in approximately 6% of all motor vehicle-related injury reports, one injury hot spot was entirely characterized by traffic circle conflicts. One report describes being struck by a car while rounding the traffic circle: “I had a few injuries, cuts, scrapes and a strained wrist. The driver of the vehicle alleged she couldn't see me but there were no obstacles to the line of sight” (report #367, regular cyclist, recreation trip, family doctor visit, woman). Traffic circles are used on residential streets as a traffic-calming strategy; however, they have been shown to be hazardous for cyclists (Harris et al. 2013). Similarly, in hot spot L, all injuries involving motor vehicles involved either a driver opening a car door (dooring) or a turning into a driveway or alley: “Struck by an oncoming Toyota Corolla when it attempted to turn left into the alley off Heather between 10th Ave & Broadway. No explanation why driver did not see me” (report #4657, regular cyclist, commute trip, family doctor visit, woman).

A strength of our study is that we leverage information that has not typically been available for studying bicycle safety—self-reports. As self-reports include a broad range of incident types and near miss incidents, the data are more plentiful and can be used to assess spatial variation in bicycling safety. Linking incidents to specific locations and infrastructure is helpful for identification of what leads to injury. Further, hot spot identification is useful for prioritizing infrastructure investments, and detailed descriptions of the major issues involved in bicycling incidents can help planners and practitioners to streamline upgrades and interventions.

A limitation of our study is that we do not compare injury patterns to official data, and as such, our results may exhibit selection bias related to BikeMaps.org users. However, there are no complete bicycling incident datasets available, and research demonstrates that official sources—insurance claims, police reports, hospital admissions—underestimate the overall burden of bicycling incidents, especially those that do not involve a motor vehicle (Winters and Branion-Calles 2017). Evaluating crowdsourced data adds depth to our understanding of bicycling incidents as reports detail circumstances around near misses, which may be an early warning for a more serious incident, as well as safety incidents aside from those that occur with motor vehicles. Importantly, our analysis does not consider equity—either among BikeMaps.org users or bicycle infrastructure provision. Earlier analysis using BikeMaps.org data indicates that those who file reports tend to be men and bicycle at least once per week (Ferster et al. 2017). Likewise, Vancouver's investment in bicycle infrastructure has been concentrated in densifying areas which are also marked by affluence (Firth et al. 2021).

## Conclusion

In this paper we examined bicyclist injuries in Metro Vancouver using crowdsourced data and characterized conditions that are more likely to lead to injury when incidents occur. Our results highlight incident characteristics, road conditions, and infrastructure that impact variation in bicycling incidents and injury. Bicycling incident injury is most associated with older and younger bicyclists, downhill slopes, parked cars on roads, falls, and single bicycle incidents with infrastructure, roads, and railroads. The results are biased toward the types of incidents that are reported to BikeMaps.org and highlight the importance of considering the health outcomes of the broad range of incidents that lead to injury. Self-reporting bicycling safety incidents enhances the data available for understanding spatial variation in risk and will help health and planning professionals set priorities for future infrastructure investments.

## References


Aldred, R. 2016. Cycling near misses: Their frequency, impact, and prevention. *Transportation Research Part A: Policy and Practice* 90: 69–83.

Aldred, R., and S. Crosweller. 2015. Investigating the rates and impacts of near misses and related incidents among UK cyclists. *Journal of Transport Health* 2(3): 379–393.

Atkinson, P. J., and D. J. Unwin. 2002. Density and local attribute estimation of an infectious disease using MapInfo. *Computers and Geosciences* 28(9): 1095–1105.

Beck, B., M. Stevenson, S. Newstead, P. Cameron, R. Judson, E. R. Edwards, A. Bucknill, M. Johnson, and B. Gabbe. 2016. Bicycling crash characteristics: An in-depth crash investigation study. *Accident Analysis and Prevention* 96: 219–227. https://doi.org/10.1016/j.aap.2016.08.012

Branion-Calles, M. C., T. A. Nelson, and S. B. Henderson. 2016. A geospatial approach to the prediction of indoor radon vulnerability in British Columbia, Canada. *Journal of Exposure Science and Environmental Epidemiology* 26(6): 554–565.

Branion-Calles, M., T. Nelson, and M. Winters. 2017. Comparing crowdsourced near-miss and collision cycling data and official bike safety reporting. *Transportation Research Record: Journal of the Transportation Research Board* 2662: 1–11.

Breiman, L. 2001. Random forests. *Machine Learning* 45(1): 5–32.

Buehler, R., and J. Pucher. 2012. Cycling to work in 90 large American cities: New evidence on the role of bike paths and lanes. *Transportation* 39(2): 409–432.

Chen, P., and Q. Shen. 2019. Identifying high-risk built environments for severe bicycling injuries. *Journal of Safety Research* 68: 1–7.

City of Vancouver. 2021. *Bikeways—City of Vancouver Open Data Portal.* https://opendata.vancouver.ca/explore/dataset/bikeways/information/?disjunctive.bike_route_namedisjunctive.year_of_construction/

Cripton, P. A., H. Shen, J. R. Brubacher, M. Chipman, S. M. Friedman, M. A. Harris, M. Winters, et al. 2015. Severity of urban cycling injuries and the relationship with personal, trip, route and crash characteristics: Analyses using four severity metrics. *BMJ Open* 5: e006654. https://doi.org/10.1136/bmjopen-2014-006654

Conrow, L., E. Wentz, T. Nelson, and C. Pettit. 2018. Comparing spatial patterns of crowdsourced and conventional bicycling datasets. *Applied Geography* 92: 21–30. https://doi.org/10.1016/j.apgeog.2018.01.009

Environment Canada. 2021. *Canadian climate normals: 1981–2010 climate normals & averages.* https://climate.weather.gc.ca/climate_normals/results_1981_2010_e.html?searchType=stnNametxtStationName=vancouversearchMethod=containstxtCentralLatMin=0txtCentralLatSec=0txtCentralLongMin=0txtCentralLongSec=0stnID=889dispBack=0

Fan, A., X. Chen, and T. Wan. 2019. How have travelers changed mode choices for first/last mile trips after the introduction of bicycle-sharing systems: An empirical study in Beijing, China. *Journal of Advanced Transportation.* https://doi.org/10.1155/2019/5426080

Ferenchak, N. N., and W. E. Marshall. 2019. Advancing healthy cities through safer cycling: An examination of shared lane markings. *International Journal of Transportation Science and Technology* 8(2): 136–145.

Ferster, C. J., T. Nelson, M. Winters, and K. Laberee. 2017. Geographic age and gender representation in volunteered cycling safety data: A case study of BikeMaps.org. *Applied Geography* 88: 144–150.

Ferster, C., T. Nelson, K. Laberee, and M. Winters. 2021. Mapping bicycling exposure and safety risk using Strava Metro. *Applied Geography* 127: 102388. https://doi.org/10.1016/j.apgeog.2021.102388

Firth, C. L., K. Hosford, and M. Winters. 2021. Who were these bike lanes built for? Social-spatial inequities in Vancouver's bikeways, 2001–2016. *Journal of Transport Geography* 94: 103122. https://doi.org/10.1016/j.jtrangeo.2021.103122

Fischer, J., T. Nelson, K. Laberee, and M. Winters. 2020. What does crowdsourced data tell us about bicycling injury? A case study in a mid-sized Canadian city. *Accident Analysis and Prevention* 145: 105695. https://doi.org/10.1016/j.aap.2020.105695

Fishman, E. 2016. Cycling as transport. *Transport Reviews* 36(1): 1–8.

Foley, J., M. Cronin, L. Brent, T. Lawrence, C. Simms, K. Gildea, J. Ryan, C. Deasy, and J. Cronin. 2020. Cycling related major trauma in Ireland. *Injury* 51(5): 1158–1163.

Götschi, T., J. Garrard, and B. Giles-Corti. 2016. Cycling as a part of daily life: A review of health perspectives. *Transport Reviews* 36(1): 45–71.

Harris, M. A., C. C. O. Reynolds, M. Winters, P. A. Cripton, H. Shen, M. L. Chipman, M. D. Cusimano, et al. 2013. Comparing the effects of infrastructure on bicycling injury at intersections and non-intersections using a case-crossover design. *Injury Prevention* 19(5): 303–310.

Heinen, E., B. van Wee, and K. Maat. 2010. Commuting by bicycle: An overview of the literature. *Transport Reviews* 30(1): 59–96.

Jestico, B., T. Nelson, and M. Winters. 2016. Mapping ridership using crowdsourced cycling data. *Journal of Transport Geography* 52: 90–97. https://doi.org/10.1016/j.jtrangeo.2016.03.006

Lee, K., and I. N. Sener. 2021. Strava Metro data for bicycle monitoring: A literature review. *Transport Reviews* 41(1): 27–47. https://doi.org/10.1080/01441647.2020.1798558

Lee, Q. Y., and D. Pojani. 2019. Making cycling irresistible in tropical climates? Views from Singapore. *Policy Design and Practice* 2(4): 359–369.

Liu, J., A. J. Khattak, X. Li, Q. Nie, and Z. Ling. 2020. Bicyclist injury severity in traffic crashes: A spatial approach for geo-referenced crash data to uncover non-stationary correlates. *Journal of Safety Research* 73: 25–35.

Mannering, F. L., and C. R. Bhat. 2014. Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research* 1: 1–22. https://doi.org/10.1016/j.amar.2013.09.001

Metro Vancouver. 2021. *Metro Vancouver*. http://www.metrovancouver.org/

Meuleners, L. B., M. Fraser, M. Johnson, M. Stevenson, G. Rose, and J. Oxley. 2020. Characteristics of the road infrastructure and injurious cyclist crashes resulting in a hospitalisation. *Accident Analysis and Prevention* 136: 105407. https://doi.org/10.1016/j.aap.2019.105407

Meuleners, L. B., M. Stevenson, M. Fraser, J. Oxley, G. Rose, and M. Johnson. 2019. Safer cycling and the urban road environment: A case control study. *Accident Analysis and Prevention* 129: 342–349. https://doi.org/10.1016/j.aap.2019.05.032

Myhrmann, M. S., K. H. Janstrup, M. Møller, and S. E. Mabit. 2021. Factors influencing the injury severity of single-bicycle crashes. *Accident Analysis and Prevention* 149: 105875. https://doi.org/10.1016/j.aap.2020.105875

Nelson, T. A., and B. Boots. 2008. Detecting spatial hot spots in landscape ecology. *Ecography* 31(5): 556–566.

Nelson, T. A., T. Denouden, B. Jestico, K. Laberee, and M. Winters. 2015. BikeMaps.org: A global tool for collision and near miss mapping. *Frontiers in Public Health* 3: 1–8. https://doi.org/10.3389/fpubh.2015.00053

Nelson, T., C. Ferster, K. Laberee, D. Fuller, and M. Winters. 2021. Crowdsourced data for bicycling research and practice. *Transport Reviews* 41(1): 97–114.

Nelson, T., A. Roy, C. Ferster, J. Fischer, V. Brum-Bastos, K. Laberee, H. Yu, and M. Winters. 2021. Generalized model for mapping bicycle ridership with crowdsourced data. *Transportation Research Part C: Emerging Technologies* 125: 102981. https://doi.org/10.1016/j.trc.2021.102981

Oja, P., S. Titze, A. Bauman, B. de Geus, P. Krenn, B. Reger-Nash, and T. Kohlberger. 2011. Health benefits of cycling: A systematic review. *Scandinavian Journal of Medicine and Science in Sports* 21(4): 496–509.

Poulos, R. G., J. Hatfield, C. Rissel, L. K. Flack, S. Murphy, R. Grzebieta, and A. S. McIntosh. 2015. An exposure based study of crash and injury rates in a cohort of transport and recreational cyclists in New South Wales, Australia. *Accident Analysis and Prevention* 78: 29–38. https://doi.org/10.1016/j.aap.2015.02.009

Prati, G., L. Pietrantoni, and F. Fraboni. 2017. Using data mining techniques to predict the severity of bicycle crashes. *Accident Analysis and Prevention* 101: 44–54.

Pucher, J., and R. Buehler. 2017. Cycling towards a more sustainable transport future. *Transport Reviews* 37(6): 689–694.

Pucher, J., J. Dill, and S. Handy. 2010. Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine* 50: 106–125.

Rosas-Satizábal, D., and A. Rodriguez-Valencia. 2019. Factors and policies explaining the emergence of the bicycle commuter in Bogotá. *Case Studies on Transport Policy* 7(1): 138–149.

Sanders, R. L. 2015. Perceived traffic risk for cyclists: The impact of near miss and collision experiences. *Accident Analysis and Prevention* 75: 26–34. https://doi.org/10.1016/j.aap.2014.11.004

Soliz, A. 2021. Divergent infrastructure: Uncovering alternative pathways in urban velomobilities. *Journal of Transport Geography* 90: 102926. https://doi.org/10.1016/j.jtrangeo.2020.102926

Statistics Canada. 2019. Census Profile, 2016 Census – British Columbia [Province] and Canada [Country]. Ottawa, ON: Statistics Canada.

Teschke, K., T. Frendo, H. Shen, M. A. Harris, C. C. Reynolds, P. A. Cripton, J. Brubacher, et al. 2014. Bicycling crash circumstances vary by route type: A cross-sectional analysis. *BMC Public Health* 14(1): 1–10. https://doi.org/10.1186/1471-2458-14-1205

Teschke, K., M. A. Harris, C. C. O. Reynolds, M. Winters, S. Babul, and M. Chipman. 2012. Route infrastructure and the risk of injuries to bicyclists: A case-crossover study. *American Journal of Public Health* 102(12): 2336–2343.

Thomas, D. R. 2006. A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation* 27(2): 237–246. https://doi.org/10.1177/1098214005283748

Utriainen, R. 2020. Characteristics of commuters' single-bicycle crashes in insurance data. *Safety* 6(1): 13. https://doi.org/10.3390/safety6010013

Vanparijs, J., L. I. Panis, R. Meeusen, and B. De Geus. 2015. Exposure measurement in bicycle safety analysis: A review of the literature. *Accident Analysis and Prevention* 84: 9–19. https://doi.org/10.1016/j.aap.2015.08.007

Watson, A., B. Watson, and K. Vallmuur. 2015. Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accident Analysis and Prevention* 83: 18–25. https://doi.org/10.1016/j.aap.2015.06.011

Winters, M., and M. Branion-Calles. 2017. Cycling safety: Quantifying the under reporting of cycling incidents in Vancouver, British Columbia. *Journal of Transport Health* 7: 48–53.

Zahabi, S. A. H., A. Chang, L. F. Miranda-Moreno, and Z. Patterson. 2016. Exploring the link between the neighborhood typologies, bicycle infrastructure and commuting cycling over time and the potential impact on commuter GHG emissions. *Transportation Research Part D: Transport and Environment* 47: 89–103.