https://dx.doi.org/10.1016/j.aap.2020.105652

Roles of infrastructure and land use in bicycle crash exposure and frequency: A case study using Greater London bike sharing data

Hongliang Ding

[a] Department of Civil and Environmental Engineering

The Hong Kong Polytechnic University

Hung Hom, Kowloon, Hong Kong

Email: hongliang.ding@connect.polyu.hk

N.N. Sze (Corresponding author)

[a] Department of Civil and Environmental Engineering

The Hong Kong Polytechnic University

Hung Hom, Kowloon, Hong Kong

Tel: +852 2766-6062; Email: tony.nn.sze@polyu.edu.hk

Haojie Li

[b] School of Transportation, Southeast University, China

[c] Jiangsu Key Laboratory of Urban ITS, China

[d] Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, China

Email: h.li@seu.edu.cn

Yanyong Guo

[b] School of Transportation, Southeast University, China

[c] Jiangsu Key Laboratory of Urban ITS, China

[d] Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, China

Email: guoyanyong@seu.edu.cn

ABSTRACT

Cycling is increasingly promoted as a sustainable transport mode. However, bicyclists are more vulnerable to fatality and severe injury in road crashes, compared to vehicle occupants. It is necessary to identify the contributory factors to crashes and injuries involving bicyclists. For the prediction of motor vehicle crashes, comprehensive traffic count data, i.e. AADT and vehicle kilometer traveled (VKT), are commonly available to proxy the exposure. However, extensive bicycle count data are usually not available. In this study, revealed bicycle trip data of a public bicycle rental system in the Greater London is used to proxy the bicycle crash exposure. Random parameter negative binomial models are developed to measure the relationship between possible risk factors and bicycle crash frequency at the zonal level, based on the crash data in the Greater London in 2012- 2013. Results indicate that model taking the bicycle use time as the exposure measure is superior to the other counterparts with the lowest AIC (Akaike information criterion) and BIC (Bayesian information criterion). Bicycle crash frequency is positively correlated to road density, commercial area, proportion of elderly, male and white race, and median household income. Additionally, separate bicycle crash prediction models are developed for different seasons. Effects of the presence of Cycle Superhighway and proportion of green area on bicycle crash frequency can vary across seasons. Findings of this study are indicative to the development of bicycle infrastructures, traffic management and control, and education and enforcement strategies that can enhance the safety awareness of bicyclists and reduce their crash risk in the long run.

Keywords: Bicycle safety, exposure, random parameter negative binomial model, land use, travel behavior

1. INTRODUCTION

Cycling has been receiving more attention in sustainable urban and transport development (Guo et al., 2019). As an active transport mode, cycling does not only relieve the traffic congestion and emission problems, but also improves the physical and mental well-being of people. In recent year, number of cyclists has been increasing remarkably in many cities around the world. Take the Greater London as an example, average daily bicycle trip was increased by 3.9% in 2015-2017. In particular, about 25% of bicycle trips were made in the Central London. In some parts of the Central London, bicycle constituted to a very high share of commuter trips. In 2016, 65% of commuter trips at Torrington Place were made by bicycle (55% at Tooley Street and 48% at Southwark Bridge) (Transport for London - TfL, 2018).

However, cyclists are vulnerable to severe injury and mortality in road crashes. In the Greater London, 93% of bicycle casualties were related to the collisions with motor vehicles (TfL, 2018). In European Union, bicyclists constituted to 8% of overall road fatalities. Unfortunately, 24% of road fatalities in the Netherlands were cyclists (Lajunen et al., 2016). To mitigate the safety hazards of bicyclists on roads, it is of paramount importance to explore the factors that contribute to the high crash and injury risk of bicyclists. Therefore, effective engineering countermeasures can be developed to enhance the overall safety of bicyclists and promote the bicycle mode.

Many studies have investigated the effects of built environment and bicycle facilities on the bicycle crash frequency at macro-level using cross-sectional model (Siddiqui et al., 2012; Narayanamoorthy et al., 2013; Chen et al., 2015; Guo et al., 2018a). To evaluate the bicycle crash risk, it is necessary to estimate the exposure (i.e. quantifying the crash potential of bicyclists). For vehicle crashes, annual average traffic flow (AADT) and vehicle kilometer traveled (VKT), based on comprehensive traffic count data, can be used to estimate the exposure (Pei et al., 2012). However, it is rare that bicycle count data are available. Bicycle crash exposure may be measured using retrospective and prospective approaches based on self-report data. They are however subjected to self-selection problem. This study addresses the problem of how to accurately measure the bicycle crash exposure, based on the revealed bicycle trip data of a public bicycle rental system. Additionally, effects of possible land use, built environment and bicycle infrastructure attributes on bicycle crash incidence are investigated. For example, Cycle Superhighway ('Superhighway') was introduced in London in early 2010s, targeted to provide cyclists with safer, faster and more direct journeys through the city (Li et al., 2018). In this study, we aim to measure the association between possible factors and bicycle crash frequency at the zonal level, using the integrated crash, environment, population profile and traffic data of London in 2012-2013 (TfL, 2013-2014). A random parameter negative binomial model would be developed to measure the association. Also, effect of the presence of Cycle Superhighway on bicycle crash risk will be considered. Moreover, separate bicycle crash prediction models would be developed for different seasons, i.e. from May to October and from November to April, taking into account the bicyclists' behavior under different weather conditions.

The paper is organized as follows. Literature review is presented in Section 2. Section 3 describes the study area, method of data collection and model formulation. Analysis results are presented in Section 4. Section 5 discusses the policy implications. Finally, Section 6 summarizes the findings and suggests the way forward.

2. LITERATURE REVIEW

Bicycle is a popular transport mode for short distance trips, both commuting and leisure travel, especially for people who do not have access to private car, e.g. adolescents, children and elderly, etc. (Wang et al., 2016; Lajunen et al., 2016; Ji et al., 2016; Vanparijs et al., 2015). Many studies have been carried out to identify the possible risk factors to bicycle crashes. Factors considered are land use, built environment, traffic attributes, and population and household characteristics (Siddiqui et al., 2012; Wei et al., 2013; Chen, 2015; Pulugurtha and Thakur, 2015; Guo et al., 2018a; Guo et al., 2018b).

For the effect of land use, Chen's study (2015) indicated that mixed land use can increase the bicycle crash risk. In particular, likelihood of bicycle crash in the industrial and commercial area is higher. It could be attributed to the conflicts between motor vehicles, bicycles and pedestrian (Narayanamoorthy et al., 2013). For the effect of traffic management and control attributes, increase in bicycle crash frequency was found correlated to the increase in the density of intersection (Siddiqui et al., 2012; Wei et al., 2013; Pulugurtha and Thakur, 2015), traffic signal (Guo et al., 2018c; de Geus et al., 2012; Chen, 2015), presence of cycle lanes (Reynolds et al., 2009; Hamann and Peek-Asa, 2013; Wei et al., 2013; Chen et al.,2016), and presence of on-street parking (Wei and Lovegrove, 2013; Vandenbulcke et al., 2014). Furthermore, environmental factors including landscape and weather conditions can also affect the level of service and safety of bicyclists (Vanparijs et al., 2015; Xing et al., 2019; Zhai et al., 2019; Fournier et al., 2017; El-Assi et al., 2015).

Personal demographic, socioeconomics, household characteristics and population profiles all affect the bicycle crash frequency. In particular, bicycle crash involvement rates of adolescents, children and elderly are all higher than that of other bicyclists. Additionally, their involvement rates of single bicycle crashes are particularly high (Rodgers, 1995; Tin Tin et al., 2010; Siddiqui et al., 2012; Ghekiere et al., 2014). Lack of sufficient skills and non-compliance to relevant guidelines were correlated to the high accident rates of adolescents and children (Mandic et al., 2018; Chong et al., 2017). For the older bicyclists, elevated crash rate could be attributed to the degradation of cognitive performance and mobility (Noland and Quddus, 2004; Vanparijs et al., 2015). For the effect of gender, studies indicate that fatality rate of male cyclist is higher than that of the female counterpart (Rodgers, 1995; Beck et al.,2007; Mindell et al., 2012; Wei et al., 2013; Vanparijs et al., 2015; Guo et al., 2018b). For the socioeconomics and household characteristics, studies indicate that household income could affect the bicycle ownership, travel behavior and therefore bicycle crash involvement (Siddiqui et al., 2012; Guo et al.; 2018a).

To estimate the crash rate, population and/or population density are often used to proxy the exposure at the macroscopic level, especially for active transportation modes like pedestrian and bicycle (Cottrill and Thakuriah., 2010; Siddiqui et al., 2012; Lee et al., 2015; Wang et al., 2017; Sze et al., 2019). Also, some studies have applied the total bicycle track length to proxy the bicycle crash exposure (Wei et al., 2013; Siddiqui et al., 2012). However, they do not account for the difference in the amount of bicycle travel among individuals. Alternately, studies have adopted bicycle count (Miranda-Moreno et al., 2011; Blaizot et al., 2013; Guo et al., 2018a; Nordback et al., 2013), vehicular traffic volume (Beck et al., 2007; Hamann and Peek-Asa, 2013; Wei et al., 2013), bicycle travel time and distance (Mindell et al., 2012; Blizot et al., 2013; Poulos et al., 2015) as the exposure measures for bicycle crash prediction models.

Above studies indicate the effects of land use, built environment, population demographic and socioeconomics, and household characteristics on the frequency and severity of bicycle crash at the macro level. However, it is rare that the effects of highway infrastructure and bicycle facilities on the bicycle crash incidence are considered. Also, the effects of safety perception and travel behavior bicyclist should be investigated. This study attempts to examine the effects of Cycle Superhighway and other road facilities on the frequency of bicycle crash using the crash and traffic data of Greater London. Also, ridership data (frequency and time) of a public bicycle rental system will be used to estimate the bicycle crash exposure. Therefore, effects of travel behavior on bicycle crash involvement can be assessed.

3. METHOD

3.1 Study area

Figure 1 illustrates the boundary of the study area under investigation. The study area covers several Inner London Boroughs like City of London, Islington, Hackney, Tower Hamlets and Westminster, etc. The geographical area was 49.1 km^2 and total population was 0.76 million respectively in 2017. Same as other global cities, cycling is increasingly popular in London in recent years. In 2017, average daily bicycle trip was 30,170 in the study area. It constituted to 2% of overall trips (TfL, 2018). Bicycles in London can generally be classified into three types: (i) privately owned bicycle; (ii) public bicycle rental system (with docking stations), e.g.

Santander Bike; and (iii) dockless bicycle sharing system, e.g. moBike, Ofo and Urbo (Li et al., 2019). Santander Bike constituted to 74% of overall bicycle trips in the study area (TfL, 2018). There are over 750 docking stations of Santander Bike in the study area. Therefore, bicycle exposure is estimated based on the data of Santander Bike. Locations of the docking stations of Santander Bike are shown in **Figure 2**. Four Superhighways, e.g. CS2, CS3, CS7 and CS8, were opened in the study area during the period 2010-2013 (Li et al., 2018). There are three major cycle track or lane types: (i) segregated one-way cycle track; (ii) segregated two-way cycle track; and (iii) non segregated (one-way) cycle lane. The network map and typical layouts of Superhighways are illustrated in **Figure 1** and **Figure 3** respectively**.**

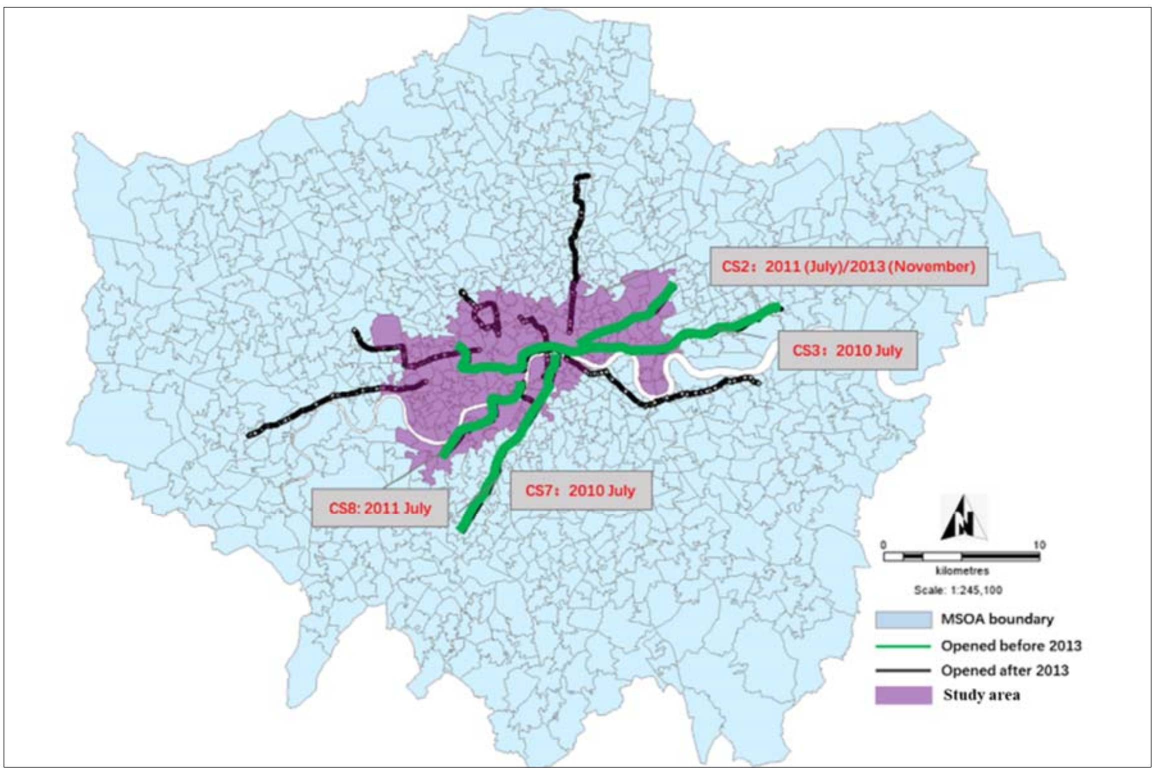

Figure 1. Location of the study area

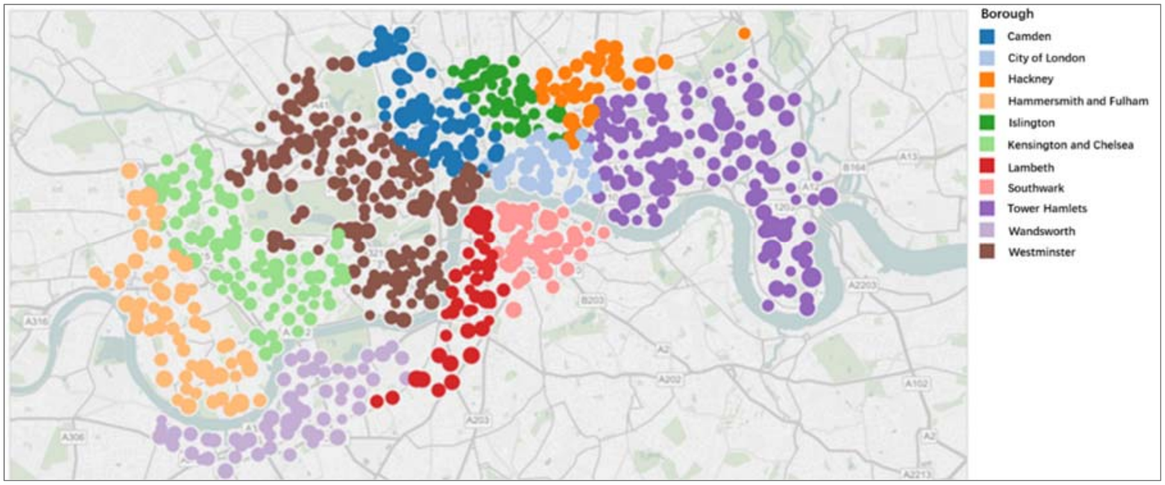

Figure 2. Locations of bicycle docking stations in the study area

(Source: *https://kitchen2018blog.blogspot.com/2018/02/boris-bikes-map.html*)

(a) Segregated one-way cycle track

(b) Segregated two-way cycle track

(c) Non-segregated cycle lane

Figure 3. Illustrations of Cycle Superhighway

(Source: *https://en.wikipedia.org/wiki/Cycle_Superhighway_3* ; *https://www.newcivilengineer.com/archive/new-cycle-superhighway-mooted-21-09-2017/*; *https://www.geograph.org.uk/photo/2372620*)

3.2 Data

In this study, bicycle crash data during the period from 2012 to 2013 was obtained from the Greater London Authority (GLA) collision data extract. The collision data extract consists of two profiles: (i) attendant profile, and (ii) casualty profile. The former records the information on road characteristics (i.e. road type and junction control), speed limit, light condition and weather condition, and information on accident location, casualty age and gender, injury severity and transport mode (i.e. bus, private car, bicycle and pedestrian) is available in the latter. **Figure 4** illustrates the geographic locations of the bicycle crashes occurred in 2012.

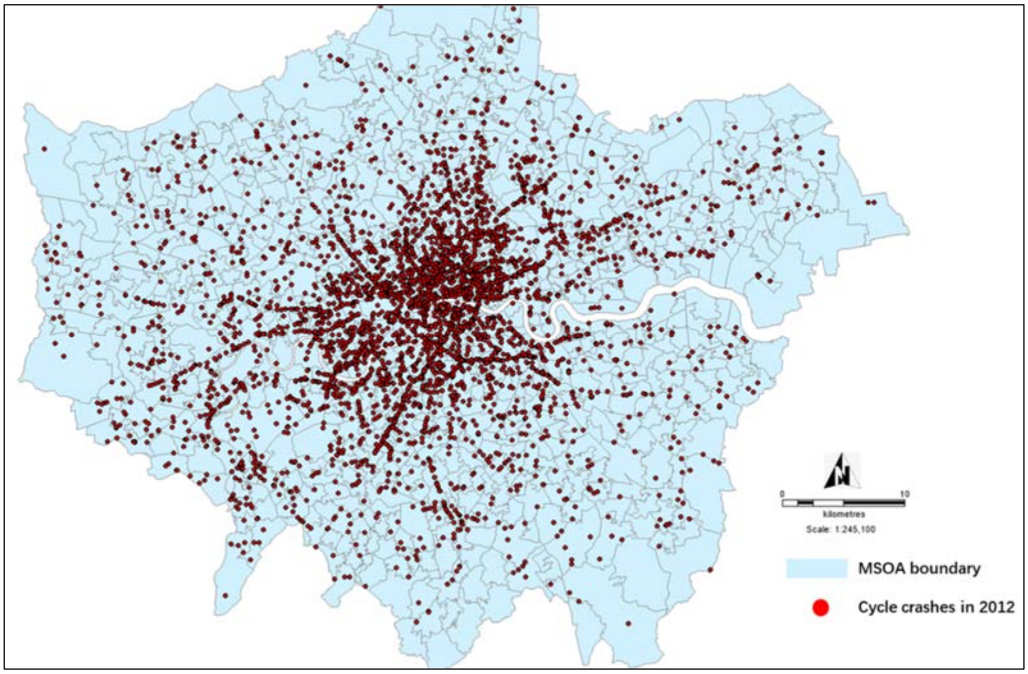

Figure 4. Location of bicycle crashes in the Greater London in 2012

Information on land use and population profiles in the Greater London are available from the census dataset of Office for National Statistics (ONS). In the UK, census data is tabulated using a hierarchical geocoding system, known as Super Output Area (SOA). In this study, the population demographic, socioeconomics and household characteristics data are aggregated at the Middle Layer Super Output Area (MSOA) level. Each MSOA on average has a population of 7,200. Variables considered are land use (i.e. residential, commercial and green area, etc.), property price, number of school, number of employment, gender, numbers of birth and death, race (i.e. white, black, Asian, and minority ethnic, etc.), health condition, household income and household composition (i.e. couple, and couple with dependent children, etc.). The latest updated land use and population census data in 2012 and 2013 are available in this study.

For the road network and traffic count data, information on road characteristics (e.g. major road, minor road, and roadway length), AADT and traffic composition (i.e. private car, taxi, bus and goods vehicle, etc.) of all major road links and some minor road links are available. **Figure 5** illustrates the locations of traffic count stations in the study area. To estimate the bicycle crash exposure, information on bicycle use (i.e. frequency and time duration) in every MSOA obtained from Santander Bike transaction records are used.

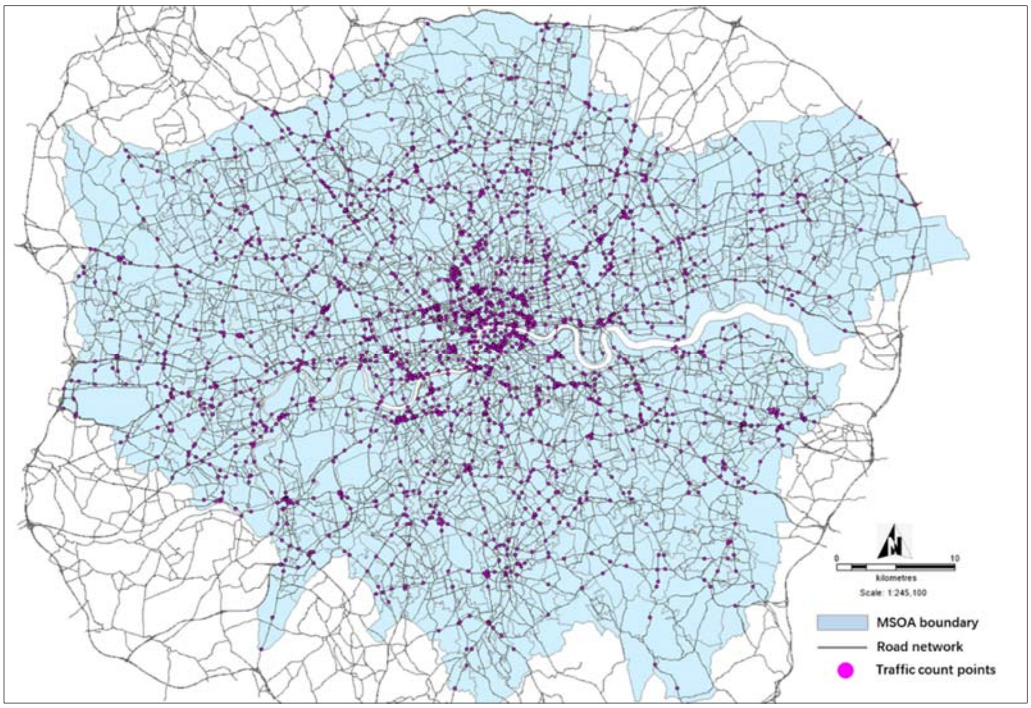

Figure 5. Location of traffic count stations in the Greater London

Information on bicycle crash incidence, crash severity, bicycle exposure, land use, demographic and socioeconomics, and household attributes are matched into the corresponding MSOA using the Geographical Information System (GIS) technique. In this study, frequencies of bicycle crash of 88 MSOAs in 2012-2013 are modeled. Sample size of proposed model is 176. There were 2,795 bicycle crashes in the study area in the observation period. To consider the effect of bicycle infrastructure on bicycle crash incidence, length of Superhighways in every MSOA is also included in the model. **Table 1** summarizes the distribution of the variables considered.

Table 1. Summary statistics of the sample

Category	Factor	Attribute	Mean	Std. Dev	Min	Max
Outcome	Frequency of bicycle crash		15.88	18.17	2	144
Road infrastructure	Road density (km per km^2)		6.14	2.59	0.86	13.09
	Cycle superhighway (km)		0.37	0.52	0	1.87
Land use	Proportion for residential		0.20	0.09	0.05	0.42
	Proportion for commercial		0.23	0.10	0.05	0.51
	Proportion for green area		0.20	0.12	0.02	0.60
	Proportion for transport facilities		0.37	0.06	0.24	0.55
Demographics	Population density (per km^2)		19.52	7.23	3.05	35.8
	Gender	Proportion of male	0.51	0.025	0.454	0.61
		Proportion of female	0.49	0.025	0.39	0.55
	Age	Proportion of age above 64	0.10	0.03	0.03	0.24
		Proportion of others	0.90	0.03	0.76	0.96
Socio-economics	Race	Proportion of white	0.53	0.08	0.34	0.69
		Proportion of others	0.47	0.08	0.31	0.66
	Median annual household income (€)		71369	28673	37130	174960
	Household type	Proportion of couple with children	0.11	0.03	0.05	0.20
		Proportion of others	0.89	0.03	0.80	0.95
Exposure	Total annual bicycle usage time (hour)	Overall	31141	31483	2498	165577
		May to October only	20500	21082	1006	110520
		November to April only	10641	10802	880	59813
	Total annual bicycle use frequency	Overall	87946	81980	7049	476329
		May to October only	54768	51755	5014	289498
		November to April only	33178	31803	2016	186831
	AADT		20365	11404	4306	62889

Note 1: Observation unit is 'Middle Super Output Area' (MOSA)
Note 2: Sample size is 176

3.3 Model formulation

The aim of this paper is to measure the association between bicycle crash frequency, exposure and possible risk factors. Several crash prediction models have been adopted in the bicycle safety literature. For instances, Poisson and Negative Binomial regression approaches are commonly used (Yao et al., 2016; Wong et al., 2007; Turner et al., 2006). To allow for parsimonious specification, a panel mixed negative binomial model (PMNB) was proposed (Bhowmik et al, 2019a). PMNB has good model fit as conventional multivariate negative binomial model. Additionally, to account for endogeneity and unobserved heterogeneities (i.e., spatial spillover effects and weather conditions), methods including random parameter negative binomial model (Bhowmik et al., 2019b), zero-inflated negative binomial model, hurdle negative binomial model (Yasmin et al., 2018; Cai et al 2016), latent segmentation based Poisson and negative binomial models (Yasmin et al., 2016), two-equation Bayesian modeling approach and spatial lag regression model (Strauss et al., 2013 a,b) were proposed. In this study, random parameter approach would be adopted to account for the effect of unobserved heterogeneity (Mannering et al., 2016; Sze et al., 2019).

For the Poisson regression model, probability of having *y* bicycle crash in the i^{th} MSOA and period *t* can be written as,

$$y_{it} \sim f(y_{it}|\mu_{it}, \tau_{it}) = \frac{\exp(-\mu_{it} \cdot \tau_{it})(\mu_{it} \cdot \tau_{it})^{y_{it}}}{y_{it}!} \qquad i, t = 0,1,2 \cdots n,$$

where μ_{it} denotes the expected number of bicycle crash.

Also,

$$\ln(\mu_{it}) = \beta_0 + x_{it}^T \cdot \beta$$

where x_{it} is the column vector of exogenous variables corresponding to the i^{th} MSOA and period *t*, β is a vector of parameters, and β_0 is the constant term.

For the negative binomial model, probability of having *y* bicycle crash can be written as,

$$P(y_{it}) = \frac{\Gamma(y_{it}+\alpha^{-1})}{y_{it}\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_{it}}\right)^{\alpha^{-1}}\left(\frac{\mu_{it}}{\alpha^{-1}+\mu_{it}}\right)^{y_{it}},$$

where $\Gamma(\cdot)$ is subjected to Gamma distribution and α denotes the over-dispersion parameter.

To account for the effect of unobserved heterogeneity, random parameter negative binomial model can be applied, where parameter β is assumed to be randomly distributed. The probability density function then becomes,

$$P(y_{it}) = \int P(y_{it}|\beta) f(\beta) d\beta$$

To assess the goodness-of-fit of proposed models, two indicators, Akaike information criterion (AIC) and Bayesian information criterion (BIC), can be applied (Train, 2001). AIC and BIC can be written as,

$$AIC = -2\ln(L) + 2k, BIC = \ln(n)k - 2\ln(L)$$

where L is the maximum likelihood function, n is the number of observations and k is the number of parameters considered respectively.

4. RESULTS

4.1 Overall Model

In this study, random parameter negative binomial model is applied to measure the association between bicycle crash frequency and possible risk factors, considering the effect of bicycle exposure. **Table 2** shows the results of parameter estimation. Three exposure measures considered are population (Model 0), bicycle use time (model 1) and bicycle use frequency (Model 2).

As shown in Table 2, AIC and BIC of Model 1 are the lowest among the three models, though not statistically significance. Hence, model using bicycle use time as the exposure measure was considered. Model using population as exposure is underperformed since it does not account for the difference in travel pattern among individuals (Guo et al., 2018a; Wang et al., 2017; Lee et al., 2015). On the other hand, studies also indicated that it was appropriate to indicate bicycle safety using relative risk (RR) with respect to travel distance and travel time (Mindell et al., 2012; Vanparijs et al., 2015). Hence, it can be expected model using bicycle use time as the exposure can achieve better goodness of fit.

Results indicate that road density, green area and commercial area, population, age, gender, household income and race all contribute to bicycle crash frequency at the 5% level of significance. For instances, increases in road density (parameter = 0.04), proportion of green area (1.14), proportion of commercial area (1.60), proportion of elderly (6.49), proportion of male (13.58), median annual household income (0.001), and proportion of white (0.03) are correlated to the increase in bicycle crash frequency. Also, the random effects of demographic and socioeconomic characteristics on crash incidence are significant at the 5% level. However, no evidence can be established for the association between bicycle crash, Cycle Superhighway, household composition and traffic volume. No obvious association between bicycle crash and traffic volume was revealed could be because of the "safety in number" effect. In other word, number of bicycle crash does not necessarily increase proportionately with the increase in traffic volume (Bjornskau et al., 2015).

Table 2. Results of parameter estimation of overall bicycle crash prediction model

Category	Factor		Model 0		Model 1		Model 2	
			Coefficients	**T-stat**	**Coefficients**	**T-stat**	**Coefficients**	**T-stat**
Constant			-21.05**	-6.65	-17.59**	-6.40	-18.50**	-6.79
Road infrastructure	Cycle Superhighway		IS	--	IS	--	IS	--
	Road density		0.07**	3.62	0.04*	2.31	0.05*	2.19
Land use	Proportion of green area		1.25**	3.32	1.14**	3.43	1.56**	3.27
	Proportion of commercial area	Mean	1.50**	3.60	1.60**	3.89	1.87**	3.55
		S.D.	*(9.86**)*					
Demographic	Log (population)	Mean	9.02**	4.39	1.80**	4.19	1.39*	2.56
		S.D.			*(12.08**)*		*(6.76**)*	
	Proportion of age above 64	Mean	5.40**	3.11	6.49**	4.39	6.19**	3.55
		S.D.					*(0.72^)*	
	Proportion of male	Mean	14.57**	8.51	13.58**	8.30	13.74**	6.81
		S.D.	*(0.36**)*		*(0.38**)*			
Socio-economics	Median annual household income		<0.001*	2.36	<0.001*	2.03	<0.001^	1.83
	Proportion of white	Mean	0.03**	4.64	0.03**	5.24	0.03**	3.64
		S.D.			*(<0.001^)*		*(0.002**)*	
	Proportion of couple with children		-0.04**	-2.61	IS	--	IS	--
Exposure	Total annual bicycle use frequency						2.98**	4.63
	Total annual bicycle use time (hour)				2.00**	4.38		
	Log (AADT)		IS	--	IS	--	IS	--
Goodness-of-fit	AIC		1150.43		1136.17		1144.64	
	BIC		1194.99		1186.19		1194.66	

Notes: *^, * and ** denote statistical significance at the 10%, 5% and 1% levels respectively.*

IS denotes insignificant.

4.2 Segregated Models

Consider the possible interferences by seasonal effects on the association between bicycle crash incidence and contributory factors, separate bicycle crash prediction models for different seasons: (i) warm season, i.e. May to October; and (ii) cold season, i.e. November to April, are developed. **Table 3** presents the results of parameter estimation of separate models. Consistent to the results of overall model, factors including commercial area, population, elderly, gender and race are significantly correlated to bicycle crash frequency, at the 5% level, in both the warm and cold seasons. Again, the random effects of demographic and socioeconomic characteristics on crash incidence are significant at the 5% level. However, 'Superhighway' is significant only in the cold season (parameter = -0.18), and 'green area' is significant only in the warm season (1.22) respectively.

Table 3 Results of parameter estimation results of separate model

Category	Factor		Warm Season		Cold Season	
			Coefficients	T-stat	Coefficients	T-stat
Constant			-18.57**	-6.11	-17.59**	-6.40
Road infrastructure	Cycle Superhighway		IS	--	-0.18*	-2.43
	Road density		0.08**	2.91	0.08**	3.37
Land use	Proportion of green area		1.22**	3.01	IS	--
	Proportion of commercial area		1.88**	3.23	1.53**	3.37
Demographic	Log (population)	Mean	2.52**	3.75	1.29*	2.24
		S.D.	*(2.92**)*		*(22.99**)*	
	Proportion of age above 64	Mean	6.75**	3.51	4.84*	2.57
		S.D.	*(0.13*)*			
	Proportion of male	Mean	12.99**	5.57	13.28**	5.67
		S.D.	*(0.17*)*		*(0.17*)*	
Socio-economics	Median annual household income		IS	--	IS	--
	Proportion of white	Mean	0.03**	3.94	0.03**	4.35
		S.D.			*(0.01**)*	
	Proportion of couple with children		IS	--	IS	--
Exposure	Log (AADT)		IS	--	IS	--
	Bicycle use time (hour)		1.13*	1.99	1.81**	3.22
Goodness-of-fit	AIC		1014.47		897.48	
	BIC		1064.50		947.51	

Notes: ** and ** denote statistical significance at the 5% and 1% levels respectively.*

IS denotes insignificant.

5. DISCUSSION

5.1 Seasonal Effect

There is a remarkable variation in the usage of rental bicycle (Santander Bike) across months. **Figure 6** illustrates the monthly bicycle rental counts in the study area in 2012 and 2013. As shown in Figure 6, the distribution of bicycle usage is similar to that of mean daily maximum temperature. In particular, average monthly bicycle usage (ranging from 557,142 to 771,428) in the period from May to October (with mean daily maximum temperature ranging from 61°F to 73°F) was remarkably higher than that (ranging from 282,857 to 454,285) in the period from November to April (with mean daily maximum temperature ranging from 48°F to 59°F). We may consider the commuters who cycle even in the cold season as regular bicyclists, while those who only cycle in the warm season as casual bicyclists. As revealed in the crash statistics in 2012-2013, total number of bicycle crashes in the warm season (1,680) was remarkably higher than that in the cold season (1,115). A possible reason is that there are more bicyclists in the warm season. Indeed, it is believed that casual bicyclists, who are expected to ride more in the warm season, usually ride for leisure purpose (TfL, 2018). That is why proportion of green area is positively correlated to the bicycle crash frequency in the warm season only (as revealed in Table 3). Therefore, it is necessary to implement effective education and promotion measures that can enhance the safety awareness and perception of casual bicyclists, particularly children, adolescent and elderly. On the other hand, presence of Cycle Superhighway is negatively correlated to the bicycle crash frequency in the cold season only. This justifies the safety benefit of upgrading bicycle infrastructure, particularly for more skillful regular bicyclists. Nevertheless, it is worth exploring the seasonal trend using disaggregated model for every month or season when comprehensive data is available in the extended study.

Figure 6. Monthly bicycle use frequency and daily maximum temperature in London

5.2 Road Infrastructure

For the road network characteristic, road density is found positively correlated to bicycle crash frequency both in overall and separate models. For instance, number of bicycle crash would be increased by 28% when road density is increased by 100%. This could be attributed to the increase in potential interactions between bicycle and motor vehicle (Wong et al., 2007; Li et al., 2018). For the presence of Cycle Superhighway, it has favorable effect on bicycle crash frequency in the cold season. It could be attributed to the increase in driver awareness and safety perception when travelling through the Cycle Superhighway. As illustrated in Figure 3(b), 3(c) and 3(d), colored asphalt pavements are applied for the cycle track and cycle lane along the Cycle Superhighway. However, no evidence can be established for significant correlation between bicycle crash frequency and presence of Cycle Superhighway for the overall and warm season models. It could be because majority of Cycle Superhighway are non-segregated (see Figure 3(d)). Also, casual bicyclists, who are usually less skillful, are expected to ride more in the warm seasons (Sze et al., 2011). The favorable effect of colored pavement on drivers' safety awareness could be offset. This finding implies that better design of Cycle

Superhighway, such as physical separation between bicycle lane and (motor) traffic lane, would be essential to enhance bicycle safety.

5.3 Land Use

Results indicate that proportion of commercial area (elasticity of 0.41) is positively correlated to bicycle crash frequency. This could be attributed to the frequent pick-up and drop-off activities on the roadsides in the commercial area. Therefore, potential bicycle crash risk could increase (Wong et al., 2007). Additionally, increase in the proportion of green area is correlated to the increase in bicycle crash frequency, particularly in the warm season (elasticity of 0.25). It could be attributed to the access to green area for recreational purpose of casual bicyclists in the warm season (Chen, 2015; Guo et al., 2018a). Moreover, commercial (commercial, office and shopping), green area (public park and plantation) and utility (highway) could constitute to 80% of the study area (Lubbock, 1963). It is importance to enhance the safety level in these areas in which pedestrian, bicycle and vehicular traffic flows are high. As shown in **Figure 4**, bicycle crashes are widely distributed in the entire study area. Bicycle facilities including segregated bicycle track, designed crossing and bicycle signal could have been introduced at the hot spots of bicycle crashes, especially in the commercial and green areas. Yet, it is worth exploring the effects of weather conditions, e.g. rain, strong wind, fog and snow, etc., on the frequency and severity of bicycle crash, when comprehensive real-time weather data are available (Wen et al., 2019; Zhai et al., 2019; Xing et al., 2019).

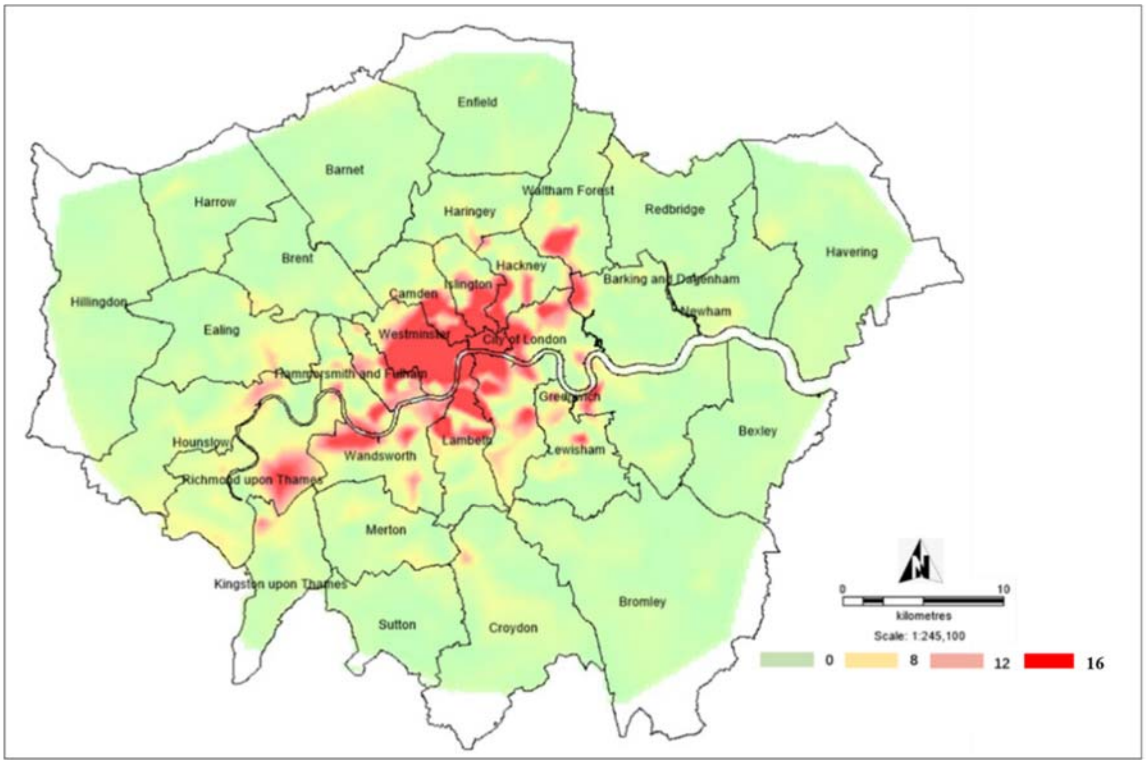

Figure 7. Distribution of bicycle crash by MSOA in the analysis period

5.4 Demographic and Socioeconomics

We also consider the safety effects of population demographic, socioeconomics and household attributes (Mindell et al., 2012; Wei et al., 2013; Ghekiere et al., 2014). Overall, population (elasticity of 7.03) is positively correlated to bicycle crash frequency. However, there is heterogeneity for the population effect based on demographic and socioeconomic characteristics. For instances, increase in the proportion of elderly (age above 64) is correlated to the increase in bicycle crash frequency (elasticity of 0.70). This could be attributed to the degradation of cognitive performance and impaired mobility of elderly. Then, the crash likelihood might increase (Palamara and Broughton, 2013; Jiménez-Mejías et al., 2016). Also, increase in the proportion of male (elasticity of 6.39) is correlated to the increase in bicycle crash frequency. This might be because male commuters are generally more aggressive and have a higher tendency to commit convicted travel behavior (Guo et al., 2018b). Furthermore, proportion of white race (elasticity of 1.28), which constituted to over 75% of overall population in the Greater London, is positively correlated to bicycle crash. Again, there are heterogeneities for the effects of male and race on bicycle crash incidence. This implies the variations in safety perception and behaviors among male and people of the same race. It can be attributed to the differences in education level, cultural background and family influences, which are not captured in the prediction model, among the people in the same group. For example, people who have higher education are more risk averse and have lower tendency to violate traffic rules (Sami et al., 2013; Hung et al., 2011). These findings are indicative to the targeted safety education and promotion strategies that can enhance the safety perception of vulnerable road user groups (TfL, 2018). For the household attribute, results indicate that increase in medium household income by 100% is correlated to the increase in bicycle crash by 19%. Yet, current results only indicate the correlation between bicycle crash frequency and characteristics of residents. It is worth exploring the relationship between population demographic & socioeconomics, bicyclist behavior and potential crash risk, when comprehensive information on bicyclists' safety perception are available in the future survey.

6. CONCLUSIONS

This study examines the relationship between possible risk factors and bicycle crash frequency at the zonal level, using the population census, land use, traffic, bicycle use and crash data of the Greater London in 2012-2013. Random parameter negative binomial regression approach is adopted. Crash exposures are estimated based on the frequency and duration of usage of a public bicycle rental system in London.

Results indicate that model using the duration of bicycle use as the exposure measure is superior to that using the frequency of bicycle use or population. It can be expected as the duration of bicycle use is a better proxy to infer the potential interactions between bicycle and motor vehicles on the roads. Additionally, road density, bicycle facilities, land use, demographic, socioeconomics and household attributes are found correlated to bicycle crash incidence. It is indicative to the development of infrastructure, traffic management and enforcement strategies that can mitigate the hazards to bicyclist on roads. In particular, the

London Cycle Superhighway network, which has favorable effect on bicycle safety, could have been extended. Also, better traffic management and control measures can be implemented to mitigate the risk of bicyclists in the commercial areas, where the roadside pick-up & drop-off activities and interactions between bicycle and motor vehicles are frequent. Furthermore, separate bicycle crash prediction models are developed for different seasons. It is believed that the characteristics and travel behavior of bicyclists are different across different time periods. Casual bicyclists could ride more frequently for recreation purpose in the warm season. That is why proportion of green area is positively correlated to bicycle crash frequency in the warm season only. This is indicative to the effective education and promotion strategies that can enhance the safety perception and awareness of bicyclists, especially those of vulnerable groups. Yet, it is worth exploring the contributory factors to the safety perception, and therefore the behavior and revealed crash risk of bicyclists. Moreover, the bicycle exposure adopted in this study is limited to the usage data of a public bicycle rental system. There could be possible bias even that it constitutes 80% of overall bicycle trips in the study area. However, the system only records the origin and destination (i.e. docking station) of a bicycle trip. It is worth exploring to use bicycle travel distance as a proxy of bicycle crash exposure, when comprehensive and extensive bicycle count are available in future study.

ACKNOWLEDGEMENTS

The work described in this paper was supported by the grants from the Research Grants Council of Hong Kong (Project No. 25203717) and The Hong Kong Polytechnic University (1-ZE5V). This work was also supported by the National Natural Science Foundation of China (Grant No. 71701042) and the Key Project of National Natural Science Foundation of China (Grant No. 51638004). And this work also supported by the National Natural Science Foundation of China (71701046).

REFERENCES

Beck, L.F., Dellinger, A.M., O'Neil, M.E., 2007. Motor vehicle crash injury rates by mode of travel, United States: using exposure-based methods to quantify differences. Am. J. Epidemiol. 166 (2), 212–218.

Bhowmik, T., Rahman, M., Yasmin, S., Eluru, N., 2019a. Alternative Model Structures for Multivariate Crash Frequency Analysis: Comparing Simulation-based Multivariate Model with Copula-based Multivariate Model. Transportation Research Board (TRB) Annual Meeting, Washington D.C.

Bhowmik, T., Yasmin, S., Eluru, N., 2019b. Do We Need Multivariate Modeling Approaches to Model Crash Frequency by Crash Types? A Panel Mixed Approach to Modeling Crash Frequency by Crash Types. Anal. Methods Accid. Res. 24.

Bjørnskau, T., Ingebrigtsen, R., 2015. Alternative understandings of risk and exposure. Institute of Transport Economics Report 1449/2015. https://trid.trb.org/view/1395684

Blaizot, S., Papon, F., Haddak, M.M., Amoros, E., 2013. Injury incidence rates of cyclists compared to pedestrians, car occupants and powered two-wheeler riders, using a medical registry and mobility data, Rhone County, France. Accid. Anal. Prev. 58, 35–45.

Cai, Q., Lee, J., Eluru, N., Abdel-Aty, M., 2016. Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models. Accid. Anal. Prev. 93, 14-22.

Chen, J., Li, Z., Wang, W., 2016. Evaluating bicycle–vehicle conflicts and delays on urban streets with bike lane and on-street parking. Transportation Letters the International Journal of Transportation Research. 10 (1), 1-11.

Chen, P., 2015. Built environment factors in explaining the automobile-involved bicycle crash frequencies: a spatial statistic approach. Saf. Sci. 79, 336–343.

Chong, S., Tyebally, A., Chew, S., Lim, Y., Feng, X., Chin, S., Lee, L., 2017. Road traffic injuries among children and adolescents in Singapore -Who is at greatest risk? Accid. Anal. Prev. 100, 59–64.

Cottrill, C.D., Thakuriah, P.V., 2010. Evaluating pedestrian crashes in areas with high low-income or low-income or minority populations. Accid. Anal. Prev. 42, 1718–1728.

de Geus, B., Vandenbulcke, G., Int Panis, L., Thomas, I., Degraeuwe, B., Cumps, E., et al., 2012. A prospective cohort study on minor accidents involving commuter cyclists in Belgium. Accid. Anal. Prev. 45, 683–693.

El-Assi, W., Mahmoud, M., Habib, K., 2015. Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. Transportation. 44, 589-613.

Fournier, N., Christofa, E., Knodler, M.A., 2017. A sinusoidal model for seasonal bicycle demand estimation. Transportation Research Part D: Transport and Environment 50, 154-169.

Ghekiere, A., Van Cauwenberg, J., de Geus, B., Clarys, P., Cardon, G., Salmon, J., et al., 2014. Critical environmental factors for transportation cycling in children: a qualitative study using bike-along interviews. PLOS ONE. 9 (9), e106696.

Guo, Y., Osama, A., Sayed, T., 2018a. A cross-comparison of different techniques for modeling macro-level cyclist crashes. Accid Anal Prev. 113, 38-46.

Guo, Y., Li, Z., Wu, Y., Xu, C., 2018b. Evaluating factors affecting electric bike users' registration of license plate in China using Bayesian approach. Transportation Research Part F. 59, 212-221.

Guo, Y., Li, Z., Wu, Y., Xu, C., 2018c. Exploring unobserved heterogeneity in bicyclists' red-light running behaviors at different crossing facilities. Accid Anal Prev. 115, 118-127.

Guo, Y., Sayed, T., Zaki, M.H., 2019. Evaluating the safety impacts of powered two wheelers on a shared roadway in China using automated video analysis. Journal of Transportation Safety & Security, 11(4), 414-429.

Hamann, C., Peek-Asa, C., 2013. On-road bicycle facilities and bicycle crashes in Iowa, 2007–2010. Accid Anal Prev. 56, 103-109.

Hung, K., Huyen, L., 2011. Education influence in traffic safety: A case study in Vietnam. IATSS Research. 34, 87-93.

Ji, Y., Fan, Y., Ermagun, A., Cao, X., Wang, W., Das, K., 2016. Public bicycle as a feeder mode to rail transit in China: The role of gender, age, income, trip purpose, and bicycle theft experience. International Journal of Sustainable Transportation. 11, 308-317.

Jiménez-Mejías, E., Martínez-Ruiz, V., Amezcua-Prieto, C., Olmedo-Requena, R., de Dios Luna-del-Castillo, J., Lardelli-Claret, P., 2016. Pedestrian-and driver-related factors associated with the risk of causing collisions involving pedestrians in Spain. Accid. Anal. Prev. 92, 211–218.

Lajunen, T., Özkan, T., Porter, B.E., 2016. Bicycle safety. Transportation Research Part F. 41B, 179-181.

Lee, J., Abdel-Aty, M., Choi, K., Huang, H., 2015. Multi-level hot zone identification for pedestrian safety. Accid. Anal. Prev. 76, 64-73.

Li, H., Ding, H., Ren, G., Xu, C., 2018. Effects of the London Cycle Superhighways on the usage of the London Cycle Hire. Transportation Research Part A. 111, 304-315.

Li, H., Zhang, Y., Ding, H., Ren, G., 2019. Effects of dockless bike-sharing systems on the usage of the London Cycle Hire. Transportation Research Part A. 130, 398-411.

Lubbock, E., 1963. Parliamentary Debates (Hansard). House of Commons.

Mandic, S., Flaherty, C., Ergler, C., Kek, C.C., Pocock, T., Lawrie, D., Chillon, P., Bengoechea, E.G., 2018. Effects of cycle skills training on cycling-related knowledge, confidence and behaviour in adolescent girls. Journal of Transport & Health. 9, 253-263.

Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.

Mindell, J.S., Leslie, D., Wardlaw, M., 2012. Exposure-based 'like-for-like' assessment of road safety by travel mode using routine health data. PLOS ONE. 7 (12).

Miranda-Moreno, L., Strauss, J., Morency, P., 2011. Disaggregate exposure measures and injury frequency models of cyclist safety at signalized intersections. Transp. Res. Rec. J. Transp. Res. Board. 2236, 74–82.

Narayanamoorthy, S., Paleti, R., Bhat, C.R., 2013. On accommodating spatial dependence in bicycle and pedestrian injury counts by severity level. Transportation Research Part B. 55, 245-264.

Noland, R., Quddus, M., 2004. Analysis of pedestrian and bicycle casualties with regional panel data. Transp. Res. Rec.: J. Transp. Res. Board 28–33.

Nordback, K., Marshall, W., Janson, B., et al., 2013. Estimating Annual Average Daily

Bicyclists: Error and Accuracy. Transp. Res. Rec. J. Transp. Res. Board. 2339, 90-97.

Palamara, P., Broughton, M., 2013. An investigation of pedestrian crashes at traffic intersections in the Perth Central Business. J. Public Health. 93, 1456–1463.

Pei, X., Wong, S.C., Sze, N.N., 2012. The roles of exposure and speed in road safety analysis. Accid. Anal. Prev. 48, 464-471.

Poulos, R.G., Hatfield, J., Rissel, C., Flack, L.K., Murphy, S., Grzebieta, R., et al., 2015. An exposure based study of crash and injury rates in a cohort of transport and recreational cyclists in New South Wales, Australia. Accid. Anal. Prev. 78, 29–38.

Pulugurtha SS., Thakur V., 2015. Evaluating the effectiveness of on-street bicycle lane and assessing risk to bicyclists in Charlotte, North Carolina. Accid Anal Prev. 76, 34-41.

Reynolds, C., Harris, M., Teschke, K., Cripton, P., Winters, M., 2009. The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. Environ. Health. 8 (1), 47.

Rodgers, G.B., 1995. Bicyclist deaths and fatality risk patterns. Accid. Anal. Prev. 27 (2), 215-223.

Sami, A., Najafi, A., Yamini, N., Moafian, G., et al., 2013. Educational level and age as contributing factors to road traffic accidents. Chinese Journal of Traumatology. 16, 281-285.

Siddiqui, C., Abdel-Aty, M., Choi, K., 2012. Macroscopic spatial analysis of pedestrian and bicycle crashes. Accid Anal Prev. 45, 382-391.

Strauss, J., Miranda-Moreno, L.F., 2013a. Spatial modeling of bicycle activity at signalized intersections. Journal of Transport and Land Use. 6(2), 47-58.

Strauss, J., Miranda-Moreno, L.F., Morency, P., 2013b. Cyclist activity and injury risk analysis at signalized intersections: A Bayesian modelling approach. Accid Anal Prev. 59, 9-17.

Sze, N.N., Tsui, K.L., Wong, S.C., So, F.L., 2011. Bicycle-related crashes in Hong Kong: Is it possible to reduce mortality and severe injury in the metropolitan area? Hong Kong Journal of Emerg Med. 18, 136-143.

Sze, N.N., Su, J., Bai, L., 2019. Exposure to pedestrian crash based on household survey data: Effect of trip purpose. Accid Anal Prev. 128, 17-24.

Tin Tin, S., Woodward, A., Ameratunga, S., 2010. Injuries to pedal cyclists on New Zealand roads, 1988–2007. BMC Public Health. 10, 655.

Transport for London, TfL, 2013. Travel in London, report. http://content.tfl.gov.uk/travel-in-london-report-6.pdf

Transport for London, TfL, 2014. Travel in London, report. http://content.tfl.gov.uk/travel-in-london-report-7.pdf

Transport for London, TfL, 2018. Travel in London, report. http://content.tfl.gov.uk/travel-in-london-report-11.pdf

Turner, S., Francis, T., Roozenburg, A., Land Transport, N., 2006. Predicting accident rates for cyclists and pedestrians. Land Transport New Zealand Research Report 289.

Vandenbulcke, G., Thomas, I., Panis, L., 2014. Predicting cycling accident risk in Brussels: a spatial case–control approach. Accid. Anal. Prev. 62, 341–357.

Vanparijs, J., Panis, L.I., Meeusen, R., de Geus, B., 2015. Exposure measurement in bicycle safety analysis: A review of the literature. Accid. Anal. Prev. 84, 9-19.

Wang, C., Xu C., Xia J., Qian Z., 2016. Modelling Faults among E-bike related Fatal Crashes in China. Traffic Inj. Prev. 18, 175-181.

Wang, J., Huang, H., Zeng, Q., 2017. The effect of zonal factors in estimating crash risks by transportation modes: motor vehicle, bicycle and pedestrian. Accid. Anal. Prev. 98, 223-231.

Wen, H., Zhang, X. Zeng, Q., Sze, N.N., 2019. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence. Accid. Anal. Prev. 132, 105249.

Wei, F., Lovegrove, G., 2013. An empirical tool to evaluate the safety of cyclists: Community based macro-level collision prediction models using negative binomial regression. Accid Anal Prev. 61, 129-137.

Wong, S.C., Sze, N.N., Li, Y.C., 2007. Contributory factors to traffic crashes at signalized intersections intersections in Hong Kong. Accid. Anal. Prev. 39, 1107–1113.

Xing, F., Huang, H., Zhan, Z.Y., Zhai, X., Ou, C., Sze, N.N., Hon, K.K., 2019. Hourly associations between weather factors and traffic crashes: non-linear and lag effects. Anal.

Methods Accid. Res., in press.
Yao, S. and Loo, B.P.Y., 2016. Safety in numbers for cyclists beyond national-level and city level data: a study on the non-linearity of risk within the city of Hong Kong. Inj. Prev. 22, 379-385.
Yasmin, S., Bhowmik, T., Rahman, M., & Eluru, N., 2018. Enhancing Non-Motorized Safety by Simulating Non-Motorized Exposure Using a Transportation Planning Approach. 97th annual meeting in TRB, Washington D.C.
Yasmin, S., & Eluru, N., 2016. Latent segmentation based count models: analysis of bicycle safety in Montreal and Toronto. Accid Anal Prev. 95, 157-171.
Zhai, X., Huang, H., Sze, N.N., Song, Z., Hon, K.K., 2019. Diagnostic Analysis of the Effects of Weather Condition on Pedestrian Crash Severity. Accid Anal Prev. 122, 318-324.