Accident Analysis and Prevention 154 (2021) 106063

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Cycling injury risk in Britain: A case-crossover study of infrastructural and route environment correlates

Georgios Kapousizis [a], Anna Goodman [b], Rachel Aldred [c,*]

[a] University of Westminster, United Kingdom
[b] London School of Hygiene and Tropical Medicine, United Kingdom
[c] University of Westminster, United Kingdom

ARTICLE INFO

Keywords:
Cycling
Injury
Route environment
Case-crossover
Collision
Infrastructure

ABSTRACT

This paper examines infrastructural and route environment correlates of cycling injury risk in Britain. We used a case-crossover design, randomly selecting control sites from modelled cyclist routes, comparing these with sites where cyclists were injured. We then used conditional logistic regression for matched case–control groups modelling to compare characteristics of control and injury sites.

Intersections were strongly associated with injury risk. High streets were associated with an elevated injury risk in final adjusted models, as was road type being primary, and a more downhill gradient. Lower speed limits and lower motor traffic connectivity were initially associated with lower injury risk, but these effects were no longer statistically significant in adjusted models. Increased road width was associated with increased injury risk in all models.

Increased injury risk was associated in all models with presence of bus lane (somewhat mitigated at stops), guardrail, and fuel station or parking lot. Presence of parked cars in street view data raised injury risk in fully adjusted models, as did congestion (measured by low morning peak speeds), while higher volumes of people cycling along the street reduced it.

In fully adjusted models, a statistically significant increase in risk was associated with presence of an on-road painted cycle lane. Most cycle lanes or tracks at control and injury sites were very poor, with narrow lanes, shared footways, and lack of protection at junctions. Given findings from other studies showing protective effects of cycle infrastructure, Britain must create higher quality cycle provision, avoiding narrow on-road painted lanes.

1. Introduction

Countries and cities with higher levels of cycling tend to have lower per-cyclist injury risks (Buehler and Pucher, 2017). As low-cycling countries and cities (such as the UK) seek to increase levels of cycling, they also aim to improve cycling safety, such that an increase in riders is not accompanied with a similar increase in injuries. However in Britain between 2001 and 2011, injury risk per cycle commuter increased both in absolute terms and in relation to other modes, where risk per commuter declined (Aldred et al., 2019). Hence, identifying infrastructural and route environment changes to reduce cycling injury risk is crucial.

Our study uses a case-crossover method to investigate this question. Specifically, we seek to examine the extent to which a wide range of variables (from actual speeds and speed limits, to bus lanes and cycle lanes, to car parks and off-street car parking) are associated with elevated or reduced risk per cycle journey. Risk in relation to the amount of cycling is a critical metric for a country seeking to increase levels of cycling, while reducing the risk to each rider.

2. Literature review

Much analysis of the infrastructural causes or correlates of cycling injuries uses an outcome variable as being either injury numbers or injury severity (e.g. Chen, 2015). However, predicting crash frequency without including a measure of bicycle volume or distance travelled means that it is not possible to separate the risk that a (type of) location poses to each individual cyclist from the number of cyclists using that (type of) location. Hence, while analysis can identify characteristics of sites with high numbers of injuries or where injuries are relatively

* Corresponding author.
E-mail address: r.aldred@westminster.ac.uk (R. Aldred).

https://doi.org/10.1016/j.aap.2021.106063
Received 6 July 2020; Received in revised form 19 January 2021; Accepted 23 February 2021
Available online 16 March 2021

severe, it may fail to identify route characteristics that keep cyclists safer but simultaneously attract more cyclists (or conversely, route characteristics that raise injury risk but simultaneously deter cycling).

One reason for this methodological limitation is a lack of cycling flow data, on which exposure calculations could be based (Dozza, 2017); another is the traditionally poor spatial data on characteristics of the street environment that might be associated with injury risk. Because of these limitations, analysis that does seek to control for exposure has frequently focused on only a small number of sites (e.g. Lusk et al., 2011) as this facilitates collection of count and infrastructure data that may not be available across a wider network. There are relatively few global analyses taking in a range of sites and a range of infrastructure types, with some (e.g. Vandenbulcke et al., 2009) using area-level data, which is limited by an inability to link risk directly to route segment characteristics.

The following sub-sections review evidence related to route environment factors that may affect cycling injury risk. As per above, much of the evidence relates to injury numbers or injury severity; however, some work does cover risk in relation to exposure and we focus on these results in the summary below.

2.1. Intersections and road hierarchy

A disproportionate amount of cycling injuries take place at or near intersections (DfT, 2017), where conflicting movements occur. In a study incorporating GPS-derived measures of cyclist flows, Strauss et al. (2015) found that signalized intersections, which are often located at the intersection of major arterials, witness 4 times more injuries and 2.5 times greater risk than non-signalized intersections. Similarly, Strauss et al. (2015) found arterial roads have higher risk than do minor roads, a finding replicated by other studies (e.g. Williams et al., 2018; Aldred et al., 2018).

2.2. Motor traffic volumes and speeds

Higher motor traffic volumes have been found to be associated with higher injury risk, with Aldred et al. (2018) finding that this has an impact independent of road class (arterial roads would generally be expected to carry more motor traffic thasn residential roads). Speed is established as a risk factor for injury severity (Chen et al., 2010).

2.3. Speed limits

Some papers have examined the impact of speed limits on cycling injury risk, as opposed to actual speeds. In London, Aldred et al. (2018) found a reduction in cycling injury odds of 21 % for 20 mph compared to 30mph streets. Kaplan et al. (2014) found similar results for Denmark.

2.4. Topography

Teschke et al. (2012) found downhill route gradients were associated with elevated injury risk, while Vandenbulcke et al's (2009) area-level study found that hilly areas had higher risks.

2.5. Land use

Studies in the USA (Cho et al., 2009), China (Ma et al., 2010), and New Zealand (Williams et al., 2018) have found relationships between land use and cyclist injuries. Some have highlighted mixed land use and/or high street locations as a risk; while Chen and Shen (2019) found that mixed land use areas have less severe injuries.

2.6. Guard railing

In Britain many cities and towns have installed ‘guard railings’ between footways and roads, particularly at ‘desire lines’ without controlled crossings, and close to junctions and crossings. A 2017 report by Transport for London found that removing guard railing reduced collisions for pedestrians and all users, however.

2.7. Cycle infrastructure

Studies that control for cycling volume (often higher on cycle lanes and tracks) generally find that cycle infrastructure plays a protective role, although with some conflicting findings regarding infrastructure type, and differences by context. Strauss et al’s (2015) Canadian study found that while there were more cyclist injuries where there were cycle tracks, this was due to higher cycling volumes, and hence the risk per cyclist was lower than on streets without cycle tracks. Again in Canada, Teschke et al. (2012) found a nine-fold reduction in cycling injury odds (albeit with large confidence intervals) for cycle tracks compared to major roads with parked cars, although they did not find a similar reduction for painted cycle lanes; nor for off-road routes. However, Williams et al. (2017) found cycle lanes (on-road, painted) in New Zealand were associated with reduced injury risk.

While Li et al. (2017) London study found no difference between cyclist injury risk on London Cycle Superhighways and other roads, their results showed ‘that it is much safer to cycle on CS3’ [Cycle Superhighway 3, which was then the only Superhighway in the study largely consisting of separate cycle tracks] than on roads with painted or no cycle infrastructure. Adams and Aldred (2020) found similar results, with separated cycle infrastructure in London associated with a 40–65 % reduction in injury odds, whereas painted lanes increased risk. By contrast, Jensen (2007) found that introduction of cycle tracks in Copenhagen during 1978–2003 was associated with a 10 % increase in cycling injury risk.

2.8. Summary

The discussion above highlights some key findings and areas of debate in the literature. As mentioned above, more evidence is still needed, especially covering a global network as opposed to (for instance) several intersection sites. Methods to control for exposure are needed to separate out the impacts of increased risk and increased usage. Where exposure data exists case-control methods can be used (e.g. Aldred et al., 2018, Williams et al., 2017, Vandenbulcke et al., 2014); however, at a national level this is rarely available.

3. Methods

3.1. Approach

This paper examines correlates of cycling injury risk in Britain, using a case-crossover method. This method avoids the problem of needing network-wide exposure data, as individual injured cyclists effectively act as their own controls. Ethical approval for the study was given by the University of Westminster. Like Teschke et al. (2012), the study uses a case-crossover method which selects control points from individual cyclist routes, then building a model to compare characteristics of control and injury sites. The method has the further strength of controlling for differences between individuals, as these individuals also act as their own controls. However, unlike Teschke et al. (2012) we do not have actual routes. Instead, we have used the Cyclestreets fastest-route journey planner to model cyclist routes prior to injury. Comparison with observed cyclists routes (see Appendix A) and other work (e.g. Meade, 2018) has suggested that this predicts sufficiently well the types of routes that cyclists tend to follow (directness being a major factor, but not the only one).

3.2. Data sources

We obtained home postcode data from the Department for Transport covering Great Britain, for all cyclists injured during 2017[1]. To identify home locations from these postcodes, we used the data from the National Statistics Postcode Lookup Centroids (NSPLC) data, provided by the Office for National Statistics. While not as useful as journey start location, for many trips the start location is a person's home, and this can be accurately predicted based on trip timing given that that >95 % of cycle trips during the morning peak start from home. We used the home postcode data alongside publically available Stats19 police road injury data, which includes data about a range of variables from injury location to involvement of other vehicles in collisions, and data on casualty gender and age group.

3.3. Generation of routes and control points

In Great Britain between 5am and 9:59am, Monday to Friday, 4303 cyclists were injured during 2017. However, only 3507 (81.5 %) had full home postcode data that we could use for modelling routes. We used postcode data lookups (to postcode area centroids) to identify home locations, and then used the Cyclestreets API (fastest-route option) to model routes from these home locations to the injury points.

We then excluded any points associated with routes longer than 25 km (137 routes, or 3.9 % of the total). With such long distances, it is likely that the person's cycle journey did not start from their home address; for instance, perhaps they were making a mixed-mode trip in which their cycle stage started at a station. While an exclusion criterion of 25 km may sound high, the study by Teschke et al. (2012) found that 4.9 % of injured cyclists had travelled over 20 km at the time they were injured. We also excluded 29 points where injury occurred <100 m from home, as we judged this did not give sufficient scope for the control and injury point to differ in their characteristics, leading to over-matching. We then generated one control point randomly from each of the 3341 remaining routes, using the ArcGIS Random Points tool (ESRI, 2020a).

3.4. Identifying route network segments and excluding off-highway control points

As we had initially routed cyclists using Cyclestreets, they were matched to the OpenStreetMap network, which Cyclestreets uses. The OSM network includes a variety of 'ways', including those open to motor traffic alongside cycleways and pedestrian paths. Hence, both injury and control points might be matched to non-highway points, depending on what happened to be the nearest 'way'.

However, injury point locations cannot be interpreted as telling us reliably on which part of a carriageway an incident took place (e.g. a police officer may use GPS to record an injury location while standing on a cycle track or footway, but the injury actually happened in the adjacent carriageway). We also do not have comparable data for control points; so even if an injury 'really' took place on a footway, Cyclestreets would never send a cyclist along that footway but rather on the adjacent carriageway. In a separate issue, we needed to exclude control points that were completely off-highway (e.g. on a canal or on a bridleway away from roads) as locations that are not adjacent to the carriageway would not be included in police injury data. Hence our results would be biased if such points were included in analysis comparing characteristics of control and injury points.

[1] While we did also have data from Northern Ireland, this represented only ~1% of all cycle injuries, and much route environment data only covered GB. Hence we decided to only cover GB in this analysis.

Our first step was thus to reclassify 552 points that we had matched to non-highway route types (cycleway, footway, bridleway, pedestrian, path, and step) rather than an adjacent highway. We first reclassified 410 (171 control and 239 injury) of these points to an adjacent highway that lay within 10 m of the point. This left 142 control points matched to a non-highway route type and located 10 m or more from a highway. These points were likely to lie away from a highway, but we also manually checked them as GIS data represents carriageways as a single line, so any distance criterion cannot accurately identify whether a point is on infrastructure adjacent to a carriageway, or not.

Our manual inspection showed that 45 of those remaining 142 points were in fact located adjacent to a carriageway, with the person routed along a parallel cycleway or other non-highway infrastructure lying 10 m or more from the centre-line of the highway. These points, as with the 410 points described above, were assigned to that highway. This left 97 points which seemed truly to be away from highways, for instance, on a river towpath. For these points, we created completely new control points selected from those segments of their routes that lay within 10 m of a highway.

3.5. Route environment data

After following the process above to avoid matching to off-highway control points, our analysis is based on analysis of 3341 injury and 3341 control points, all lying within or adjacent to a public highway. The following discussion explains how we matched these points to different characteristics of the route environment. We sourced route environment data in a range of ways. This included the use of datasets provided by partners (e.g. Basemap) or online (e.g. OpenStreetMap) and use of Google Street View (manual lookups). For details of data sources and how we derived the resultant variables see Appendix B.

We assigned each point the following route environment characteristics, grouped *a priori* into four different categories:

1 Area type: urban/rural status, high street status, average small area deprivation.
2 Road type: road class, road width, road gradient, speed limit, motor connectivity ranking.
3 Nearby street infrastructure: Cycle infrastructure, guard railing, bus lane, bus stop, metro/rail/tram stop, fuel station/parking lot, intersection status.
4 Travel behaviour: average AM peak speed, parked cars, cycle commuter flow.

Note that intersections include intersections with other highways, including e.g. entrances to car parks, but do not include private driveways.

3.6. Statistical analysis

We first used univariable descriptive statistics to highlight characteristics of injury and control points. To separate the impacts of the different aspects of the route environment, we used regression modelling to predict whether a point was a control or an injury site, entering our route environment factors as predictor variables. We used conditional fixed effects logistic regression, matching each injury point to its sampled control point, and we present our results as odds ratios.

We fitted the adjusted regression models using a hierarchal modelling structure, guided by the classification of route environment characteristics into four categories described above. Specifically, we started with the 'area type' variables which we conceptualised as most distal to the outcome. We then added the 'road type' and 'nearby street infrastructure' variables, which we hypothesised might mediate to some extent the effect of area type. Finally, we additionally adjusted for 'travel behaviour', which we conceptualised might in turn mediate some of the effects of road type and street infrastructure.

We first ran the analysis for all points. We then conducted sensitivity analyses stratifying between KSI (Killed and Seriously Injured) casualties and slight injuries, and present results for tests for interaction between each predictor and whether the injury was a KSI versus a slight injury.

Note that because our study is focusing on injuries occurring during the morning commute, we expect control points to be closer to a person's home and further from a person's work than their injury point. In addition, on average the places where people work will be less residential and more commercial than the places where people live. These two facts in combination meant that we expected that injury points would generally have a higher workplace density than control points, simply as an artefact of our methodology. Workplace density is obtained form the Classification of Workplace Zones, provided by the Consumer Data Research Centre, and measures the number of workers located within each small zone during the working day, based on Census 2011 data. This trend was indeed observed: workplace density was higher in the injury point for 1207 participants (36.1 %), was higher in the control point for 728 participants (21.8 %), and was similar (within 0.05) for 1406 participants (42.1 %). This pattern has the potential to create confounding if some aspects of the road environment vary systematically with workplace density. To attempt to reduce any such confounding, we therefore included workplace density in all adjusted models as a covariate.

For cycling volume data we used estimates produced by the Propensity to Cycle Tool (PCT) route network. This PCT route network was created using the Cyclestreets algorithm, as were our modelled cyclist routes. Injury points can happen anywhere that cyclists travel, potentially including entirely away from this PCT route network. Such 'off the PCT network' points may often either have no PCT route within 20 m or else only a minor one containing a very small number of cyclists. By contrast, control points are selected after using the Cyclestreets algorithm to determine a route between the person's home and their injury point. Our method therefore means control points are less likely to be 'off the PCT network', and therefore less likely to get a zero or very low cycle volume value. We therefore expected to see a disproportionate number of injury points associated with very low cycle volumes simply as an artefact of this methodology. Visual inspection indicated that this was indeed the case, with the effect limited to routes with a volume $<=$ 5. For this reason, we decided that when modelling cycle volume as a continuous variable we would simultaneously enter a binary dummy variable identifying whether the route contained 0–5 versus 6+ cyclists. In the appendix we also offer an alternative presentation in which cycle flow is entered as a categorical variable (which is more transparent, but less well powered).

We examined crude associations to guide how continuous variable should be entered into our model. Motor connectivity ranking was highly correlated with road class and other road type variables, and we therefore decided to enter it as a categorical variable. Otherwise we preferred where possible to enter continuous variables as linear terms, to increase power and to avoid the complications of interpretation that come with using quadratic terms. To limit the effect of outliers, we decided to cap road width at 15 m (252 higher values, or 3.8 %, rounded down to 15), to cap average peak speed at 50 miles/hr (370 higher values, or 5.5 %, rounded down to 50) and to cap the number of cycle commuters at 1000 (88 higher values, or 1.3 %, rounded down to 1000). After doing this, all continuous variables showed an approximately linear relationship in visual inspection, and there was no evidence of non-nonlinearity as judged by the inclusion of a quadratic term (all $p >$ 0.05 in adjusted analyses).

The proportion of variables with missing data ranged from 0 to 6.4 %. We imputed this data using multiple imputation (25 imputations) under an assumption of Missing at Random. We confirmed in sensitivity analyses that the results were similar when using a complete case analysis on the 2589 participants (77.5 %) with complete data for both injury and control points.

Table 1
Characteristics of individuals and of their crash.

Characteristic	Level	N (%)
Full sample	–	3341 (100 %)
Country	England	3159 (94.6 %)
	Scotland	131 (3.9 %)
	Wales	51 (1.5 %)
Sex	Male	2579 (77.2 %)
	Female	762 (22.8 %)
Age	0–15	293 (8.9 %)
	16–24	415 (12.5 %)
	25–39	1276 (38.5 %)
	40–59	1139 (34.4 %)
	60–74	155 (4.7 %)
	75+	34 (1.0 %)
Small-area deprivation	Fifth 1 (richest)	546 (17.3 %)
	Fifth 2	569 (18.0 %)
	Fifth 3	642 (20.3 %)
	Fifth 4	778 (24.6 %)
	Fifth 5 (poorest)	623 (19.7 %)
Injury severity	Fatal	14 (0.4 %)
	Serious	578 (17.3 %)
	Slight	2749 (82.3 %)
Striking vehicle	No other vehicle	188 (5.6 %)
	Cycle	20 (0.6 %)
	HGV	70 (2.1 %)
	Bus	38 (1.1 %)
	Other motor vehicle (mostly cars)	3025 (90.5 %)
Light conditions	Light	2933 (87.8 %)
	Dark	408 (12.2 %)
Weather conditions	Fine, no high winds	2708 (85.4 %)
	Other	464 (14.6 %)
Road surface conditions	Dry	2401 (74.3 %)
	Other	832 (25.7 %)

Numbers add to less than 3341 for some variables due to missing data: in these cases, the % is calculated relative to those with non-missing data.

4. Results

4.1. Sample characteristics

The characteristics of the 3341 individuals in our sample are shown in Table 1. As this shows, the large majority were from England. 77 % of individuals were male, and 73 % were aged 25–59. There were not large differences by small-area deprivation, but people in the richest two-fifths were somewhat underrepresented. With regard to the injury, 82 % were slight, 17 % serious and 0.4 % fatal. In 6% of cases the injury occurred with no other vehicle involved as the striking vehicle. Otherwise, a very small proportion involved collisions with cyclists (0.6 %), HGVs (2%) or buses (1%). The large majority, 91 %, involved other motor vehicles, mostly cars. Most injuries occurred when it was light (as can be expected given we only selected injuries during the morning commute), and in fine weather with dry road conditions.

4.2. The effects of area, road, street infrastructure and travel behaviour

4.2.1. Effects of area-type variables

Being an urban area and being on a high street were both significantly associated with an increased odds of injury in univariable analyses, but there was no association with area deprivation. The impact of being in an urban area attenuated after mutual adjustment for whether a

Table 2
Predictors of injury, all points.

Category	Predictor	Level	N points	% of which injury points	Univariable	Adjusted 1	Adjusted 2	Adjusted 3
Area Type	**Urban**	**Rural**	490	45%	1***	1***	1	1
		Urban	6192	50%	2.02 (1.43, 2.85)	1.88 (1.33, 2.66)	1.37 (0.90, 2.08)	1.40 (0.90, 2.17)
	High Street	**No**	6014	48%	1***	1***	1***	1**
		Yes	668	67%	2.52 (2.08, 3.06)	2.15 (1.77, 2.62)	1.75 (1.40, 2.20)	1.48 (1.17, 1.86)
	Average deprivation	**Change per standard deviation increase**	–	–	1.09 (1.01, 1.16)*	1.09 (1.01, 1.17)*	1.04 (0.95, 1.13)	1.03 (0.95, 1.13)
Road type	**Road class**	**Primary**	2511	59%	1***		1***	1***
		Secondary	744	49 %	0.51 (0.41, 0.63)		0.68 (0.53, 0.87)	0.70 (0.54, 0.90)
		Tertiary	1210	45%	0.41 (0.34, 0.49)		0.59 (0.47, 0.73)	0.59 (0.47, 0.73)
		Residential or other	2216	43%	0.40 (0.34, 0.46)		0.64 (0.51, 0.81)	0.50 (0.39, 0.65)
	Road width	**Change per 1 m increase**	–	–	1.23 (1.20, 1.26)***		1.12 (1.09, 1.16)***	1.11 (1.08, 1.15)***
	Gradient	**Change per 1% increase in incline (downhill = negative)**	–	–	0.96 (0.94, 0.99)**		0.97 (0.94, 1.00)*	0.96 (0.93, 1.00)*
	Speed limit	**20 mph or less**	1257	47%	1***		1	1
		30mph	4582	51%	1.39 (1.16, 1.67)		0.92 (0.73, 1.17)	0.93 (0.73, 1.19)
		40mph	424	49 %	1.21 (0.92, 1.61)		0.87 (0.60, 1.26)	1.10 (0.74, 1.62)
		over 40mph	382	45%	0.94 (0.68, 1.30)		0.99 (0.65, 1.50)	1.38 (0.88, 2.16)
	Connectivity rank	**0—24%**	327	40%	1***		1	1
		25—49%	620	43%	1.14 (0.85, 1.51)		1.09 (0.78, 1.52)	1.20 (0.85, 1.68)
		50—74%	1281	46%	1.33 (1.03, 1.72)		1.19 (0.88, 1.62)	1.43 (1.04, 1.96)
		75—100%	4170	53%	1.93 (1.51, 2.46)		1.03 (0.74, 1.42)	1.39 (1.00, 1.94)
Nearby street infra-structure	**Bicycle infrastructure**	**None**	5209	48%	1***		1***	1***
		Track (no lane)	571	53%	1.32 (1.09, 1.59)		1.13 (0.90, 1.41)	1.17 (0.93, 1.48)
		Lane (no track)	627	60%	1.84 (1.51, 2.23)		1.34 (1.06, 1.69)	1.39 (1.10, 1.76)
		Track and Lane	66	88%	9.55 (4.34, 21.0)		5.99 (2.55, 14.0)	6.35 (2.61, 15.4)
		Other, e.g. sign	131	54%	1.34 (0.94, 1.91)		1.14 (0.75, 1.73)	1.22 (0.80, 1.88)
	Guardrail	**No**	5704	47%	1***		1***	1***
		Yes	900	66%	2.33 (1.99, 2.73)		1.57 (1.31, 1.89)	1.48 (1.23, 1.79)
	Bus lane	**No**	6250	49 %	1***		1**	1***
		Yes	354	69%	2.56 (1.99, 3.29)		1.58 (1.17, 2.14)	1.60 (1.18, 2.17)
	Bus stop	**No**	5987	51%	1*		1*	1*
		Yes	695	45%	0.81 (0.69, 0.95)		0.82 (0.68, 0.99)	0.82 (0.68, 1.00)
	Metro/rail/ tram stop	**No**	6640	50%	1*		1	1
		Yes	42	67%	2.00 (1.05, 3.80)		1.30 (0.60, 2.84)	1.17 (0.53, 2.57)
	Fuel station or parking lot	**No**	6289	49 %	1***		1**	1**
		Yes	393	63%	1.82 (1.46, 2.27)		1.50 (1.16, 1.95)	1.47 (1.13, 1.92)
	Intersection	**No**	2400	29%	1***		1***	1***
		Yes	4282	62%	4.42 (3.90, 5.00)		3.59 (3.14, 4.10)	3.43 (2.99, 3.93)
Travel behaviour	**2-way average morning peak speed**	**Change per 10 mph increase**	–	–	0.71 (0.67, 0.75)***			0.76 (0.70, 0.83)***
	Parked cars	**No**	2832	48%	1*			1***
		Yes	3772	51%	1.15 (1.03, 1.28)			1.35 (1.17, 1.55)
	No. cycle commuters on segment	**Change per 100 cyclists increase**	–	–	1.00 (0.96, 1.04)			0.95 (0.90, 0.99)*

†p<0.1, *p < 0.05, **p < 0.01, ***p < 0.001 in tests for heterogeneity. Numbers in the 'N' column add to less than 6682 points for some variables due to missing data. In all other columns all 6682 points are used, using multiple imputation. All adjusted models additionally adjust for workplace density, as linear and quadratic terms, and when examining number of commuters on the segment we additionally included a dummy variable '0–5 cycle commuters versus 6+'.

high street and became no longer significant (adjusted model 1), and then further attenuated upon additional adjustment. This suggests the univariable urban effect reflected the types of roads found in urban areas plus the higher concentration of high streets. The impact of being on a high street was also somewhat attenuated after adjusting for road type, street infrastructure and travel behaviour, suggesting that these characteristics explain some of the original effect. Nevertheless, there was still a significant independent effect (OR 1.48, 95 % CI 1.17–1.86) in the final adjusted model (adjusted model 3), suggesting some risk posed by aspects of the high street not captured in the other variables.

4.2.2. Effects of road type variables

All five variables were significantly associated with the odds of injury in univariable analyses. After mutual adjustment plus adjusting for area type and nearby street infrastructure (adjusted model 2), injury was independently predicted by primary road type (with less variation among the remaining road types); greater road width; and a lower gradient value. Note that the gradient value includes negative values for downhill travel, i.e. there was a higher odds of injury for downhill travel than flat travel, and for flat travel than uphill travel. There was no longer evidence in adjusted analysis of an independent effect of speed limit (whereas in univariable analyses 30mph streets had a higher injury odds than 20 mph streets) or motor connectivity (whereas in univariable analysis there was an association between higher connectivity and higher odds of injury).

4.2.3. Effects of street infrastructure variables

Five of the six variables were significantly associated with odds of injury in univariable analyses, the exception being having a nearby bus stop. After mutual adjustment, plus adjusting for area type and road type (adjusted model 2), injury was independently predicted by the presence of a bicycle lane or a bicycle track plus a lane. There was also a trend towards higher odds of injury in the presence of a bicycle track with no lane, although this did not reach statistical significance in adjusted models. These associations with bicycle infrastructure type changed little after adjusting for travel behaviour, indicating that the association cannot readily be explained by e.g. a greater volume of cyclists on those streets. The highest odds was observed in the presence of the combination of both a track and a lane but this may serve to some extent as a proxy for intersection status.

Increased odds of injury was also independently predicted by the presence of a guardrail, a bus lane, a fuel station or a parking lot; or an intersection. The effect of being near an intersection was particularly strong with an adjusted odds ratio of 3.43 (95 % CI 2.99, 3.93). Again, none of these associations changed much after adjusting for the travel behaviour variables. Interestingly, after adjusting for the greater risk conferred by the presence of a bus lane, there was weak evidence of a protective effect of being near a bus stop ($p = 0.05$), which would partly counteract the raised risk of bus lanes at those locations. There was no evidence of an effect of being near a metro/rail/tram stop in adjusted analyses, although this may reflect lack of power given that only 42 injury or control points in our sample were near a metro/rail/tram stop.

4.2.4. Effects of travel behaviour variables

Higher average speed was associated with a lower odds of injury in both univariable and adjusted analyses. Parked cars were associated with a somewhat raised odds of injury in univariable analyses, and this effect strengthened and became more significant in adjusted models. This strengthening upon adjustment seemed to reflect adjusting for road type (parked cars are more common on residential streets). In univariable analyses there was no association between odds of injury and the volume of cycling, but after adjustment for other factors a higher volume of cyclists was associated with a lower odds of injury. In particular, it seemed the protective effect of higher cycle volume was initially masked the fact that higher cycle volume is positively correlated with greater road width. After adjusting for road width, the protective effect of cycle volume became significant.

4.3. Examination of differential effects between slight injuries versus KSI

We conducted stratified analyses comparing the 2749 individuals with a slight injury to the 592 individuals who were killed or seriously injured (KSI). The results are tabulated in the Appendix. In general, the point estimates for effect were similar between the two injury types, although less often statistically significant for KSI because of the much smaller sample size. There was never evidence of an interaction between any of the 17 predictor variables shown in Table 2 and KSI status (all $p \geq 0.07$). It therefore appears that our findings presented in relation to Table 2 apply both to slight injuries and to KSI.

5. Discussion

5.1. Summary of findings

Intersections were strongly associated with risk of injury. High street status was associated with an elevated injury risk in final adjusted models, while urban area status was not, an initial effect becoming attenuated when adjusting for other variables. In adjusted models, injury risk was also independently predicted by road type being primary, and by a more downhill gradient. Lower speed limits and lower motor traffic connectivity (to some extent, a proxy for motor traffic volume) were initially associated with lower injury risk, but these effects were no longer statistically significant when adjusting for other variables. Increased road width was associated with increased injury risk in all models.

Increased injury risk was independently predicted by presence of a bus lane (partly counteracted by a smaller reduction in risk associated with bus stop proximity), a guardrail, a fuel station or parking lot. None of these associations changed much in fully adjusted models. In fully adjusted models, a statistically significant increase in risk was associated with presence of an on-road cycle lane, or presence of both a cycle lane and an off-road cycle track. An off-road track alone was associated with increased risk in univariable modelling, but this was not statistically significant in adjusted models. The presence of parked cars in street view data raised injury risk, as did congestion (measured by low morning peak speeds), while higher volumes of people cycling along the street reduced it.

Findings suggest that injury risk is increased by width and classification of road, and by factors generating potentially conflicting movements by other road users – i.e. intersections, shops, fuel stations and parking lots, and parked cars; although the presence of other cyclists reduced risk. Bus lanes, a principal form of provision for cycling on busy roads, are shown to increase injury risk, although this increased risk is somewhat mitigated close to bus stops. Perhaps surprisingly, on-road cycle lanes are associated with a increase in risk similar to presence of a bus lane combined with a bus stop, and although off-road cycle tracks were not associated with a statistically significant increase in risk in the adjusted models, they did not appear to be protective.

Fig. 1. Typical “cycle lane” examples.

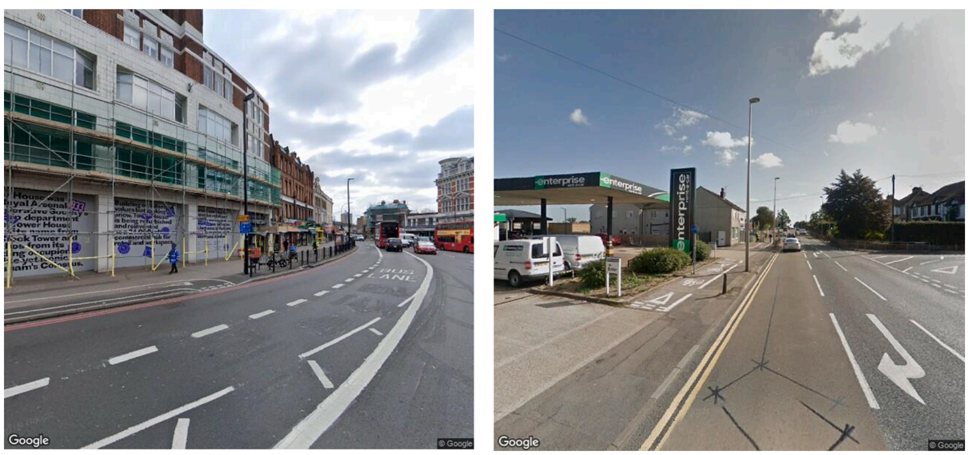

Fig. 2. Typical "cycle track" examples (on the left, the track is on the footway to the left of the picture).

5.2. Limitations

This study is limited in a variety of ways. We were only able to include weekday morning peak journeys, due to not having journey start location and hence having to use home postcode as a proxy. Data was incomplete so we had to exclude those injured cyclists for whom home postcode was not known. Our use of a modelling algorithm to route the cyclists could lead to bias, for instance, if cyclists in practice make more use of residential roads than is suggested by the algorithm. However, use of a relatively direct route (the Cyclestreets ‘fast route’ algorithm prioritises directness, but does avoid the very busiest roads where possible) is, we believe, likely to represent well enough cyclist routes, especially at commuting times. We were limited in the route environment data sources available to us, and use of current street view images may introduce bias, if for instance infrastructure has been built post-2017 partly in response to perceived dangerous environments. Our data predominantly relates to slight injuries, these being the large majority of injuries recorded by the police. Most slight injuries, however, are not reported to the police so Stats19 only contains a subset of such injuries. Even most cycling injuries requiring hospitalisation may not be found in Stats19 (Jeffrey et al., 2009) which predominately contains injuries involving motor vehicles.

5.3. Strengths

The study is able to use national data and to control for cyclist volume, and for individual characteristics, through the case-crossover approach used. This is unusual and represents an innovative use of secondary data and of modelling, allowing the research to be conducted without potentially intrusive and time-consuming primary data collection.

5.4. Meanings of our findings

Unsurprisingly, our findings confirmed that main roads and wider roads are riskier for people cycling. Adjusting for these factors meant that the impact of speed limits became statistically insignificant. This suggests that perhaps in practice road design is more important in injury risk than formal measures to reduce speeds alone. Our modelling of actual motor traffic speeds in the morning peak suggested that congestion may also increase injury risk, with roads with very low motor traffic speeds seeing higher risks. The finding for guard railing – to our knowledge our study is the first to examine this in relation to cyclist injury risk– suggests that this (anti)pedestrian infrastructure may help to create a perception among drivers that they will not encounter conflict with non-motorised users (TfL, 2017), to cyclists' as well as pedestrians' detriment, as well as offering the potential for cyclists to be physically crushed against the railing.

The negative impact of environments with conflicting motor traffic movements appears clear in most cases, particularly related to kerbside activity. This may account for the somewhat protective effect of bus stops, sometimes counteracted by the larger negative impact of bus lanes. Car parking is likely to be restricted at bus stops, to allow pedestrian access and egress, providing greater visibility for people cycling. Conversely, presence of parked cars increases risk and often in practice will reduce the protective impact of residential roads. As in other studies, we found a safety in numbers impact from other cyclists being present on the road segment; there did not appear to be a negative impact from conflicting movements in relation to other cyclists.

Our findings in relation to cycle infrastructure are contrary to other literature, which generally finds a protective impact. Instead, we find cycle lanes (and those few locations with combined lanes and tracks) are associated with elevated injury risk. Cycle tracks initially showed the same pattern but this attenuated and became statistically insignificant on adjustment for other factors. Still, one might expect a protective effect from this infrastructure. Why have we not found this? Assuming that our algorithm has not introduced bias (e.g. cyclists in practice are more likely to use roads with cycle infrastructure than predicted by the Cyclestreets direct routing), we believe the explanation likely lies in the quality of the cycle infrastructure encountered by the cyclists.

Cycle lanes were frequently narrow, non-mandatory, or disappeared at road narrowings or on encountering allocated car parking (Figs. 1 and 2). The infrastructure characterised here as 'tracks' frequently meant shared footways rather than true tracks, and like lanes, these had a tendency to disappear at junctions or at road narrowings. We do not of course know whether cyclists were actually riding in a given track or lane at any point, and in many cases they may not have been doing so (for instance, where a piece of infrastructure is a bumpy, shared-use footway that frequently gives way to side roads). When we examined Stats19 data on casualty locations for our sample of injured cyclists, we found that only 7% were recorded as being in a cycle lane or track, although Table 2 shows that around a quarter of our injury sites contained some kind of cycle infrastructure[2]. In any case, unlike in some other studies (Teschke et al., 2012, Williams 2017) British cycle infrastructure does not appear protective, and its presence may (especially in the case of lanes) even raise risk.

Given (i) the evidence from international studies (e.g. Teschke et al., 2012, and in London Li et al., 2017; Adams and Aldred, 2020) that high-quality separated cycle infrastructure should reduce, not increase risk, (ii) the evidence from other international studies that such infrastructure acts to encourage cycle uptake, we would suggest that in the British context, higher standards for cycle infrastructure are urgently needed (at the time of publication, new guidance has now been produced for England which draws on the higher standards developed in London), and poor quality paint-based infrastructure lacking physical separation from motor traffic should not be implemented.

Author statement

RA led the project, GK led the data processing and spatial analysis, AG led the statistical analysis. All contributed to the paper.

Funder

This project was funded by the Road Safety Trust, UK.

[2] Note that police recording of casualty location may not necessarily reflect where the person was cycling; for instance, if they were cycling in a narrow advisory lane this may not be where they are found post-collision.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

We would like to thank our stakeholder group for input into the project, and Irena Itova for assistance with running the PCT model for Scotland, helping with the analysis of Points of Interest data, and contributing to the analysis of GSV (Google Street View) data. Keny Toto and Eno Omorogieva also assisted with GSV data coding.

Appendix A. Selection of routing method

Our approach to modelling routes was informed by analysis of 230 tracks from early-morning commuters provided to us by Beeline, a new company which produces a navigation app like a compass for cyclists (Beeline, 2019). 107 from those points were for the London area and 123 for the rest of the country. We excluded training, leisure rides and wrong data (straight lines): in London, 7 tracks were recognized as training, being routes that consisted of circles around parks rather than routes from one place to another, 40 as leisure trips and 14 were wrong data (straight lines). Outside of the London area, there were 36 leisure rides, 16 training tracks and 19 tracks that were with straight lines. This left 98 valid tracks (46 for London and 52 out of London) that were used to decide which algorithm we should use for routing.

From these routes, we used the start and the end points to create new routes through two alternative methods in order to find which corresponded better to the actually observed routes. These alternatives were ArcGIS and Cyclestreets API.

1 To model routes in ArcGIS, the tool Network Analyst was used. This is based on Dijkstra's algorithm which generates solutions for a shortest-path problem on an undirected, nonnegative, weighted graph, in our case the road network (ESRI, 2020b). We adapted the parameters such that road types such as motorway and footway were excluded.
2 Cyclestreets.net is a journey planner for cyclists, which uses a related algorithm but incorporating other variables (for instance, likely cycling speed on different types of segment). Cyclestreets offers different options (Fast, Balanced and Quiet) which trade off directness against route comfort. Initial investigations (and evidence from Meade and Stewart, 2018) suggested that only the Fast option was likely to well represent commuter cycling behaviour, with the other options creating relatively long detours due to the paucity of cycling infrastructure in much of Britain.

Once the routes were created into ArcGIS, the random point tool in ArcGIS was used to create 20 random points for each route per source (20 points*98 tracks*3 sources of tracks, i.e. Beeline, Cyclestreets and Dijkstra), resulting in 5880 points in total. Every point corresponded to a road segment of each track. The approach aimed to draw out information from the road network segment for each one of these 5880 points, by using the spatial join tool in ArcGIS, since the aim was not to model exactly where people went, but rather represent well the types of the routes they chose.

Comparing the subsequent road types across the three route types (actual, Dijkstra, and Cyclestreets), we found that Cyclestreets provided the closer comparison to the actual routes followed. For instance, 27 % of the actual route points were located on residential or unclassified streets, compared to 30 % for the Cyclestreets algorithm but only 20 % for the Dijkstra algorithm. This informed our decision to use CycleStreets to route start and end points in the present paper.

Appendix B. Route environment data sources

Table B1

Table B1
route environment data sources.

Sequence number	Variables and contributing factors	Value	Type of variable	Operationalisation of variables	Dataset name, owner, and date	Data location
Area type						
1	Urban	1 Rural 2 Urban	Polygon	We matched the Rural Urban Classification with the boundaries for England, Wales and Scotland using the Lower Layer Super Output Areas code. Then, we identified where the injury and control points are located within the boundaries of LSOA.	Rural Urban Classification, GOV.UK, Department for Environment, Food & Rural Affairs, January 2020 Urban Rural Classification, Scottish Government, Geographic Information Science & Analysis Team, January 2020	https://data.gov.uk/dataset/b1165cea-2655−4cf7-bf22-dfbd3cdeb242/rural-urban-classification-2011-of-lower-layer-super-output-areas-in-england-and-wales https://statistics.gov.scot/data/urban-rural-classification
2	High Street	0 Not on or close to a high street 1 On or close to a high street	Point	We used the POIs catalogue but only some of the categories. These were Retail, Eating and drinking, Education and health, Sport and entertainment, Attractions, Commercial services. Once we selected the classification, we matched them with the corresponding data from the whole POIs dataset and we created polygon clustering based on the point data using ArcGIS. Then we selected the road network from OSM within the polygon cluster. At the final step, we selected all the injury and control points that are located within 25 m of the selected road network.	Points of Interest, Ordnance Survey, November 2018	https://www.ordnancesurvey.co.uk/documents/product-support/support/points-of-interest-classification-scheme.pdf https://digimap.edina.ac.uk/
3	Average deprivation	Change per standard deviation increase	Polygon	We located injury and control points inside each zone and looked up the deprivation levels per household for Lower Super Output Areas and Data Zones.	Classification of household deprivation (Great Britain) 2011 - Lower Super Output Areas and Data Zones, UK Data service, dataset used December 2019	https://www.statistics.digitalresources.jisc.ac.uk/
4	Workplace density		Polygon	The workplace population with the boundaries has been matched. Then we located injury and control points inside each zone and looked up the workplace density.	Classification of Workplace Zones, Consumer Data Research Centre, dataset used Janury 2020	https://data.cdrc.ac.uk/
Road Type						
5	Road class (hierarchy)	Primary (A road, motorway) Secondary (B road) Tertiary (C class road) Residential or other unclassified road	Line	We mapped injury and control points to the nearest OSM road segment. As vector datasets represent roads as lines, and injuries are more frequent on major than minor roads, matching off-network points by distance tends to disproportionately allocate the points to minor roads, at intersection locations (Aldred et al., 2018). We similarly found that when comparing our initial distance-based matching of injury points to route segments, only 31.6% were matched to major roads, compared to an allocation of 43.3% by the police for the same set of points. While police data is not always completely accurate, this disparity suggests that at or close to intersections, our matching was biased towards minor roads. Hence, we carried out the following process. Points lying within 10 m of an	Great Britain (England, Scotland and Wales) datasets, Open Street Map, dataset used March 2019	https://www.geofabrik.de/data/download.html

(continued on next page)

Table B1 (*continued*)

Sequence number	Variables and contributing factors	Value	Type of variable	Operationalisation of variables	Dataset name, owner, and date	Data location
				intersection (267 locations) were reclassified to a major road, where they had initially been assigned to a minor road. In total, this then gave 1472 injury points located at a major road, a number that represents 43.7% of injury points, close to the proportion recorded by the police (43.3%).		
6	Road width	Change per 1 m increase	Line	We used the OS Mastermap road network. Then the nearest roads on a range (buffer zone) of 20 m. of the injury and control points were selected. We used the average road width classification from the dataset.	Highways Network Road, Ordnance Survey, November 2019 dataset used	https://www.basemap.co.uk/
7	Gradient	Change per 1% increase in incline	API	The elevations and the distances from the Cyclestreets API have been used. We used road segments up to 250 m before the injury and control points with the same slope in order to calculate the gradient.	Cyclestreers API, Cyclestreets, journey planner system, APi used March 2020	https://www.cyclestreets.net/api/
8	Speed limit	1 20 mph or less 2 30 mph 3 40 mph 4 over 40 mph	Line	We selected the nearest road on a range (buffer zone) of 20 m. of the injury and control points.	Basemap (the creator of the dataset) directly provided speed limit data from 2017 to us; speed limit data is also now available via Ordnance Survey Public Sector Mapping Agreement: https://www.ordnancesurvey.co.uk/business-government/products/mastermap-highways-speed-data	Basemap (the creator of the dataset) directly provided speed limit data from 2017 to us; speed limit data is also now available via Ordnance Survey Public S https://www.ordnancesurvey.co.uk/business-government/products/mastermap-highways-speed-data ector Mapping Agreement:
9	Connectivity rank	0–24% 25–49% 50–74% 75–100%	Line	The SpaceSyntax dataset has been used. It is a linear dataset and we used the 10 km Choice Rank classification. The nearest road segment on a range of 20 m of the injury and control points has been used.	Space Syntax OpenMapping, Spacesyntax, dataset used January 2020	https://spacesyntax.com/openmapping/
Nearby street infrastructure						
10	Bike infrastructure	0 No bicycle infrastructure 1 Track (no lane) 2 Lane (no track) 3 Track and Lane 4 Other, e.g. sign	GSV	Lookups to see whether any bicycle infrastructure was present at any of the four streetview images that were downloaded for each point (where available). Then coding of the bicycle infrastructure type, separating lanes (on-road, paint-based) from tracks (off-road, separated from motor vehicles in some way).	Google Street View images, Google API, API used November 2019-March 2020	https://rrwen.github.io/google_streetview/https://developers.google.com/maps/documentation/streetview/intro
11	Bus lane	0 Not bus lane 1 Yes, bus lane	GSV	GSV lookups to see whether a bus lane was visible in any of the four lookup images.	Google Street View images, Google API, API used November 2019-March 2020	https://rrwen.github.io/google_streetview/ https://developers.google.com/maps/documentation/streetview/intro
12	Bus stops	0 No, bus stops in a range of 20 m 1 Yes, bus stops in a range of 20 m	Point	We used data from NAPTAN. We created the point based on the coordinates and then used a 20 m range (buffer zone) from injury and control points in order to select all the relative points (Bus stops)	National Public Transport Access Nodes, Department for Transport, dataset used December 2019	https://data.gov.uk/dataset/ff93ffc1–6656-47d8–9155-85ea0b8f2251/national-public-transport-access-nodes-naptan
13	Metro/rail/tram stops	0 No, bus stops in a range of 20 m 1 Yes, bus stops in a range of 20 m	Point	We used data from NAPTAN. We created the point based on the coordinates and then used a 20 m range (buffer zone) from injury and control points in order to select all the relative points (Metro/rail/tram stops)	National Public Transport Access Nodes, Department for Transport, dataset used December 2019	https://data.gov.uk/dataset/ff93ffc1–6656-47d8–9155-85ea0b8f2251/national-public-transport-access-nodes-naptan
14	Fuel station or parking lot	0 Without Fuel station or parking lot on a range of 20 m	Point, polygon	Data from OSM was used. Then we selected all the points that are related to the fuel station or parking lot in a range (buffer zone)	Great Britain (England, Scotland and Wales) datasets, Open Street Map, dataset used January 2020	https://www.geofabrik.de/data/download.html

(*continued on next page*)

Table B1 (*continued*)

Sequence number	Variables and contributing factors	Value	Type of variable	Operationalisation of variables	Dataset name, owner, and date	Data location
		1 Within Fuel station or parking lot on a range of 20 m		of 20 m from injury and control points.		
15	Intersection	0 Without an intersection in 20 m range 1 Within an intersection in 20 m range	Line, point	The OSM road network was used. We identify the intersections using ArcGIS. Then we used a range (buffer zone) of 20 m of the injury and control points that lay near to an intersection.	Great Britain (England, Scotland and Wales) datasets, Open Street Map, dataset used March 2019	https://www.geofabrik.de/data/download.html
Travel behaviour						
18	2-way average morning peak speed	Change per 10 mph increase	Line	We used the average speed based on 2017 from basemap. We matched the speed data with the Master map network based on the TOID number. Then the nearest road on a range (buffer zone) of 20 m. of the injury and control points were selected	TOIDs (based on 2017) which have the average speed for the morning peak, Basemap (the creator of the dataset) directly provided the average speed data for the morning peak from 2017 to us.	https://www.basemap.co.uk/
19	Parked cars	0 Not on or close to cars parked 1 On or close to cars parked	GSV	GSV lookups to see whether parked cars were visible in any of the four lookup images.		https://www.geofabrik.de/data/download.html https://rrwen.github.io/google_streetview/https://developers.google.com/maps/documentation/streetview/intro
20	Cycle commuters on segment	Change per 100 cyclists increase	Line	We used the PCT tool which uses Census origin-destination data to allocate commuter cyclists across the route network within England and Wales. The nearest road segment on a range (buffer zone) of 20 m of the injury and control points has been used. As the PCT does not cover Scotland, we used the stplanr package in R (developed for the PCT) to create cycling volume using data from Census 2011	Cycle commuters, Propensity to Cycle Tool (PCT), dataset used December 2019 Census Scotland 2011, National Records of Scotland, dataset used January 2020	https://www.pct.bike/ https://github.com/ropensci/stplanr https://www.scotlandscensus.gov.uk/

Appendix C. results split by KSI status

Table C1

Table C1
Results split by KSI status.

Category	Predictor	Level	Slight injuries (N = 5498 points)			KSI (N = 1184 points)			P for Interaction with KSI status[a]
			N points	% injury	Adjusted	N points	% injury	Adjusted	
Area type	**Urban**	**Rural**	351	44%	1	139	48%	1	p = 0.48
		Urban	5147	50%	1.50 (0.87, 2.61)	1045	50%	1.30 (0.60, 2.80)	
	High Street	**No**	4921	48%	1**	1093	48%	1	p = 0.81
		Yes	577	66%	1.45 (1.13, 1.87)	91	69%	1.73 (0.92, 3.25)	
	Average deprivation	**Change per standard deviation increase**	–	–	1.01 (0.92, 1.12)	–	–	1.13 (0.90, 1.42)	p = 0.47
Road type	**Road class**	**Primary**	2094	59%	1***	417	60%	1	p = 0.51
		Secondary	584	48%	0.70 (0.53, 0.93)	160	51%	0.63 (0.33, 1.23)	
		Tertiary	982	45%	0.58 (0.45, 0.75)	228	46%	0.58 (0.33, 1.02)	
		Residential or other	1837	44%	0.48 (0.37, 0.63)	379	41%	0.57 (0.31, 1.05)	
	Road width	**Change per 1 m increase**	–	–	1.11 (1.07, 1.15)***	–	–	1.14 (1.05, 1.24)**	p = 0.17
	Gradient	**Change per 1% increase in incline**	–	–	0.97 (0.94, 1.01)	–	–	0.93 (0.86, 1.00)*	p = 0.24

(continued on next page)

Table C1 (*continued*)

Category	Predictor	Level	Slight injuries (N = 5498 points)			KSI (N = 1184 points)			P for Interaction with KSI status[a]
			N points	% injury	Adjusted	N points	% injury	Adjusted	
	Speed limit	**20 mph or less**	1061	48%	1	196	44%	1	p = 0.13
		30mph	3789	51%	0.88 (0.67, 1.15)	793	52%	1.15 (0.62, 2.13)	
		40mph	329	49 %	1.07 (0.69, 1.66)	95	47%	1.13 (0.46, 2.78)	
		over 40mph	287	43%	1.10 (0.65, 1.84)	95	51%	2.81 (1.05, 7.54)	
	Connectivity rank	**0−24%**	263	41%	1	64	36%	1	p = 0.15
		25−49%	501	44%	1.10 (0.75, 1.60)	119	42%	1.67 (0.74, 3.77)	
		50−74%	1064	46%	1.30 (0.92, 1.84)	217	45%	1.97 (0.92, 4.21)	
		75−100%	3425	53%	1.23 (0.85, 1.78)	745	54%	2.18 (0.99, 4.80)	
	Bicycle infrastructure	**None**	4266	48%	1***	943	48%	1	p = 0.74
Nearby street infrastructure		**Track (no lane)**	477	52%	1.14 (0.88, 1.48)	94	58%	1.37 (0.74, 2.54)	
		Lane (no track)	533	60%	1.46 (1.13, 1.90)	94	55%	1.04 (0.57, 1.90)	
		Track and Lane	57	86%	5.44 (2.19, 13.5)	9	100 %	[cannot converge]	
		Other, e.g. sign	104	53%	1.08 (0.67, 1.75)	27	59%	2.01 (0.75, 5.43)	
	Guardrail	**No**	4663	48%	1**	1041	47%	1*	p = 0.22
		Yes	774	65 %	1.42 (1.16, 1.75)	126	74 %	1.87 (1.09, 3.21)	
	Bus lane	**No**	5137	49 %	1**	1113	49 %	1	p = 0.54
		Yes	300	68%	1.72 (1.23, 2.41)	54	70%	1.24 (0.57, 2.68)	
	Bus stop	**No**	4917	51%	1	1070	50%	1	p = 0.93
		Yes	581	45%	0.83 (0.67, 1.02)	114	49 %	0.85 (0.50, 1.43)	
	Metro/rail/ tram stop	**No**	5465	50%	1	1175	50%	1	p = 0.47
		Yes	33	70%	1.39 (0.56, 3.42)	9	56%	0.77 (0.11, 5.20)	
	Fuel station or parking lot	**No**	5168	49 %	1**	1121	49 %	1	p = 0.25
		Yes	330	63%	1.58 (1.18, 2.13)	63	60%	1.09 (0.58, 2.06)	
	Intersection	**No**	1963	29%	1***	437	27 %	1***	p = 0.83
		Yes	3535	62%	3.41 (2.93, 3.97)	747	63%	3.75 (2.69, 5.23)	
Travel behaviour	**2-way average morning peak speed**	**Change per 10 mph increase**	–	–	0.75 (0.68, 0.82)***	–	–	0.83 (0.69, 1.00)*	p = 0.07
	Parked cars	**No**	2303	48%	1***	529	48%	1	p = 0.20
		Yes	3134	51%	1.37 (1.17, 1.60)	638	51%	1.15 (0.80, 1.66)	
	No. cycle commuters on segment	**Change per 100 cyclists increase**			0.94 (0.90, 1.00)*			1.00 (0.88, 1.13)	p = 0.30

[†]p<0.1, *p < 0.05, **p < 0.01, ***p < 0.001 in tests for heterogeneity. All models additionally adjust for workplace density, as linear and quadratic terms, and a dummy variable ‘0–5 cycle commuters versus 6+’.

[a] In tests for interaction, we fitted the conditional logistic regression model for all 6682 points plus an interaction term for KSI status and one other predictor. We did this separately for each of the 18 predictors, to generate the 18 terms for interaction shown.

References

Adams, T., Aldred, R., 2020. Cycling Injury Risk in London: Impacts of Road Characteristics and Infrastructure. Transport Findings.

Aldred, R., Goodman, A., Gulliver, J., Woodcock, J., 2018. Cycling injury risk in London: a case-control study exploring the impact of cycle volumes, motor vehicle volumes, and road characteristics including speed limits. Accid. Anal. Prev. 117 (November 2017), 75–84. https://doi.org/10.1016/j.aap.2018.03.003.

Aldred, R., Goel, R., Woodcock, J., Goodman, A., 2019. Contextualising safety in numbers: a longitudinal investigation into change in cycling safety in Britain, 1991–2001 and 2001–2011. Inj. Prev. 25 (3), 236–241.

Beeline, 2019. Beeline, 11th June. https://beeline.co/.

Buehler, R., Pucher, J., 2017. Trends in walking and cycling safety: recent evidence from high-income countries, with a focus on the United States and Germany. Am. J. Public Health 107 (February 2), 281–287.

Chen, P., 2015. Built environment factors in explaining the automobile-involved bicycle crash frequencies: a spatial statistic approach. Saf. Sci. 79, 336–343. https://doi.org/10.1016/j.ssci.2015.06.016.

Chen, P., Shen, Q., 2019. Identifying high-risk built environments for severe bicycling injuries. J. Saf. Res. 68, 1–7. https://doi.org/10.1016/j.jsr.2018.11.002.

Chen, Y., Yang, J., Otte, D., 2010. Load and Impact Conditions for Head Injuries in Car-to-Pedestrian and Car-to-Cyclist Accidents – A Comparison of Real Accidents and Simulations.

Cho, G., Rodríguez, D.A., Khattak, A.J., 2009. The role of the built environment in explaining relationships between perceived and actual pedestrian and bicyclist safety. Accid. Anal. Prev. 41 (4), 692–702. https://doi.org/10.1016/j.aap.2009.03.008.

CycleStreets.net. 2021 CycleStreets.net. n.d. https://www.cyclestreets.net/.

Department for Transport, 2017. Pedal Cycle Factsheet. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/686969/pedal-cycle-factsheet-2017.pdf.

Department for Transport. data.gov.uk. 20th January 2019. https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data.

Dozza, M., 2017. Crash risk: how cycling flow can help explain crash data. Accid. Anal. Prev. 105, 21–29. https://doi.org/10.1016/j.aap.2016.04.033.

ESRI, 2020a. Random Points Tool, 25th November 2020. https://pro.arcgis.com/en/pro-app/tool-reference/data-management/create-random-points.htm.

ESRI, 2020b. Network Analyst Extension, 23rd November 2020. https://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/algorithms-used-by-network-analyst.htm.

Jeffrey, S., Stone, D.H., Blamey, A., et al., 2009. An evaluation of police reporting of road casualties. Inj. Prev. 2009 (15), 13–18.

Jensen, S.U., 2007. Bicycle Tracks and Lanes: a Before-After Study. available at. https://www.researchgate.net/publication/237524182_Bicycle_Tracks_and_Lanes_a_Before-After_Study/link/5a548377458515e7b732688e/download.

Kaplan, S., Vavatsoulas, K., Prato, C.G., 2014. Aggravating and mitigating factors associated with cyclist injury severity in Denmark. J. Safety Res. 50, 75–82. https://doi.org/10.1016/j.jsr.2014.03.012.

Li, H., Graham, D.J., Liu, P., 2017. Safety effects of the London cycle superhighways on cycle collisions. Accid. Anal. Prev. 99 (A), 90–101.

Lusk, A.C., Furth, P.G., Morency, P., Miranda-Moreno, L.F., Willett, W.C., Dennerlein, J. T., 2011. Risk of injury for bicycling on cycle tracks versus in the street. Inj. Prev. 17 (2), 131–135. https://doi.org/10.1136/ip.2010.028696.

Ma, M., Yan, X., Abdel-Aty, M., Huang, H., Wang, X., 2010. Safety analysis of urban arterials under mixed-traffic patterns in Beijing. Transp. Res. Rec.: J. Transp. Res. Board 2193, 105–115. https://doi.org/10.3141/2193-13.

Meade, S., Stewart, K., 2018. The visualisation of SiN in Edinburgh. Scottish Transp. Appl. Res. 2020, 1–15.

Office for National Statistics. 6th November 2019. https://geoportal.statistics.gov.uk/datasets/national-statistics-postcode-lookup-latest-centroids.

Strauss, J., Miranda-Moreno, L.F., Morency, P., 2015. Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. Accid. Anal. Prev. 83, 132–142. https://doi.org/10.1016/j.aap.2015.07.014.

Teschke, K., Harris, M.A., Reynolds, C.C.O., Winters, M., Babul, S., Chipman, M., et al., 2012. Route infrastructure and the risk of injuries to bicyclists: a case-crossover study. Am. J. Public Health 102 (12), 2336–2343. https://doi.org/10.2105/AJPH.2012.300762.

Transport for London, 2017. Collisions Before and After the Removal of Pedestrian Railings at 70 Junctions and Crossings on the Transport for London Road Network. Report no. SB257. TfL, London.

Vandenbulcke, G., Thomas, I., de Geus, B., Degraeuwe, B., Torfs, R., Meeusen, R., Int Panis, L., 2009. Mapping bicycle use and the risk of accidents for commuters who cycle to work in Belgium. Transp. Policy 16 (2), 77–87. https://doi.org/10.1016/j.tranpol.2009.03.004.

Williams, T., Page, S., Doscher, C., 2018. Spatial characteristics of bicycle–motor vehicle crashes in Christchurch, New Zealand: a case-control approach. J. Transp. Land Use 11 (1), 849–864. https://doi.org/10.5198/jtlu.2018.1147.