The impact of Low Traffic Neighbourhoods in London on the characteristics of people walking, cycling and wheeling: statistical analysis plan

Dr Anna Goodman,[a] Dr Jamie Furlong,[b] Professor Ben Armstrong,[c] Professor Phil Edwards,[a] Dr Ersilia Verlinghieri,[b] Dr Jack Benton,[d] Dr Harrie Larrington-Spencer,[b] Dr David Fevyer,[b] Dr Anthony Laverty,[e] Professor Rachel Aldred[b]

[a] Department of Population Health, London School of Hygiene and Tropical Medicine

[b] Active Travel Academy, School of Architecture and Cities, University of Westminster

[c] Department of Public Health, Environments and Society, London School of Hygiene and Tropical Medicine

[d] Manchester Centre for Health Psychology, Division of Psychology & Mental Health, University of Manchester

[e] School of Public Health, Imperial College London

Study aim

To examine whether the introduction of new Low Traffic Neighbourhoods (LTN) is associated with a change in the characteristics of people walking, cycling and wheeling within the area. This includes examining impacts on: the age and gender profile of pedestrians and cyclists; the proportion of pedestrians using mobility aids; independent mobility among children; and the use of helmets, high viz clothes and sports clothes by cyclists.

Proposed methods

LTN areas

This study prospectively identified 8 proposed LTNs in London and matched them to Control areas. Key variables used in matching each control area to a proposed LTN included the geographic size of the area; the demographic and socio-economic characteristics of its residents; and the similarity of travel patterns within the area. In addition, we required the control area to be suitable to become an LTN in principle but with no LTN planned there in practice; and not to be located immediately adjacent to its matched LTN.[1]

We will include in this analysis the four LTNs that have been implemented and remain operational as of July 2024, as shown in Table 1.[2] Note that the scheme in Camden Square is considerably smaller than the other schemes assessed (0.08km^2, vs 0.27-0.61km^2 for the other 3 schemes: see Table 1). As such, it only just meets our scope/extent criteria for what constitutes an LTN.[3] The limited size of the scheme might influence its observed impacts, and we therefore propose to present sensitivity analyses excluding Camden Square.

---

[1] Further details in https://westminsterresearch.westminster.ac.uk/item/w1q51/statistical-analysis-plan-low-traffic-neighbourhoods-in-london-interrupted-time-series-analysis-of-sensor-count-data

[2] Streatham Wells was implemented in October 2023 but then suspended in March 2024, before we could collect any follow-up data for this research question.

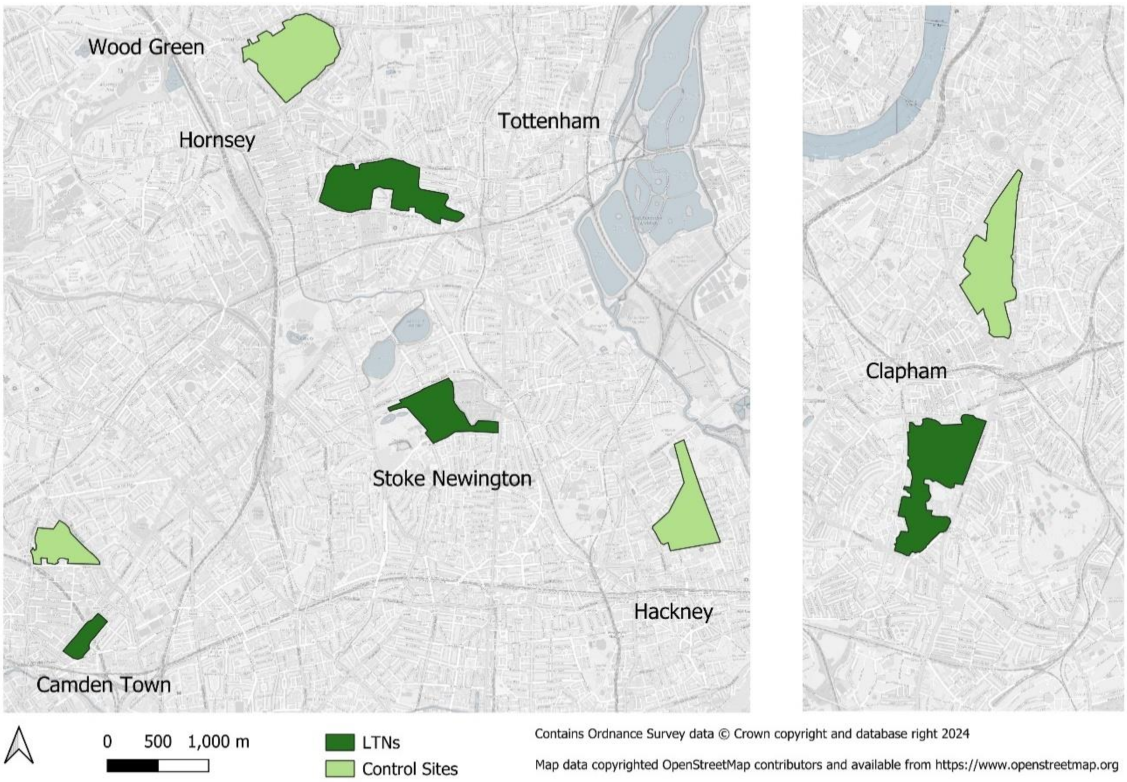
[3] Further details on our definition of LTNs can be found here: https://westminsterresearch.westminster.ac.uk/item/wqx4q/road-traffic-injuries-and-ltns-statistical-analysis-plan

Table 1: Overview of study LTNs and the data collected from them

Local Authority / short name	Scheme name	LTN area in km²	Date implemented	No. observation points, LTN / Control	Dates of 'before' data collection	Dates of 'after' data collection
Hackney	Stoke Newington	0.27	20/09/2021	2 / 2	08/09/2021 to 11/09/2021	07/09/2022 to 10/09/2022 06/09/2023 to 09/09/2023 04/09/2024 to 07/09/2024
Camden	Camden Square	0.08	16/12/2021	1 / 2 †	30/06/2021 to 03/07/2021	08/06/2022 to 11/06/2022 07/06/2023 to 10/06/2023 05/06/2024 to 08/06/2024
Haringey	St Anns	0.47	22/08/2022	2 / 2	08/06/2022 to 11/06/2022	07/06/2023 to 10/06/2023 05/06/2024 to 08/06/2024
Lambeth	Brixton Hill	0.61	04/09/2023	2 / 2	07/06/2023 to 10/06/2023	05/06/2024 to 08/06/2024

This table is limited to LTNs which are operational as of July 2024, in time for inclusion in this planned analysis. † Note that the Camden LTN scheme was smaller than we had originally anticipated, meaning that one of our planned observation points was outside the LTN area and was not used in this analysis. In addition, one of the Camden control sites had poor visibility of the left-hand pavement because of overhanging tree branches, and so only we only coded people in the roadway or on the right-hand pavement.

Figure 1: Our study's LTN schemes and their matched control areas



The three LTN schemes on the right-hand map are in north London. They are as follows: bottom left = Camden Square LTN in Camden and its matched control; top of map = St Anns LTN in Haringey and its matched control; bottom right quadrant of map = Stoke Newington LTN in Hackney and its matched control. The one LTN scheme on the left-hand map are in south London. It shows Brixton Hill LTN Lambeth and its matched control.

Selecting measurement locations inside LTNs and Control areas

We selected two observation points inside each LTN and control area. These points were used in our study to collect two data sources: a) continuous 24/7 counts of pedestrians, cyclists and motor vehicles using VivaCity sensors (data which is not used in the present analysis) and b) the video data that will be used for the present analysis

As we have previously described,[4] the method for selecting these observation points involved identifying road segments that were travel desire lines within each LTN and control area. We chose to focus on travel desire lines because we wanted to study streets where any impact of an LTN would be measurable, as opposed to already quiet streets that might be less likely to see any change and where we would have less statistical power to detect any change. We then purposively identified two observation points in each area that a) were >200m crow-fly distance from each other, and b) covered different desire lines (e.g., North-South and East-West, or two different East-West lines). These criteria were used to reduce double counting. In addition, the observation point had to be a lamppost suitable for installing VivaCity sensors and video cameras, e.g. with a clear view of the street.

Advice from local authority stakeholders helped us identify road segments in control areas that the stakeholders felt had comparable walking, cycling, and motor traffic flows to our selected road segments in the intervention area (generally there was no existing data on such flows for these streets, so this had to be based on their local knowledge). Where intervention LTN areas had key destinations likely to affect travel and traffic flows (e.g., a park or a school), we either identified an observation point at a similar location in a control area (e.g. both LTN and control observation point near to a school) or, if this was not possible, chose an observation point away from that destination.

[4] https://westminsterresearch.westminster.ac.uk/item/w1q51/statistical-analysis-plan-low-traffic-neighbourhoods-in-london-interrupted-time-series-analysis-of-sensor-count-data

Collecting video data and coding the characteristics of people walking, cycling and wheeling

At each observation point, we commissioned CTS Traffic and Transportation Ltd to erect temporary cameras and collect four days of video data (see Table 1 for dates). Video data was collected by CTS Traffic and Transportation Ltd from 7am to 7pm on Wednesday to Saturday, thereby covering a mixture of week and weekend days. In compliance with GDPR requirements, the video footage collected was of a low/pixellated quality such that numberplates could not be read, and facial features were blurred.

We limited seasonal impacts by collecting data at very similar times of year each year, and by seeking to avoid collecting data on days with rain, high wind or extreme temperatures. In addition, by always using the same four days for the LTN and control group, we ensured any variation in environmental conditions would be balanced between the observation points.

Contractors for CTS Traffic and Transportation Ltd then watched and manually coded the video footage to record the characteristics of people walking and cycling past the observation point. For each observed person, the characteristics recorded were as follows:

- Date and time stamp
- Estimated age group: 0 to 3 years, 4 to 10 years, 11 to 16 years, 17 years plus.[5]
- Estimated gender: Male, Female, Unsure
- Mode of travel: Walk, Jog, Scoot/skate/rollerblade, Wheelchair, Mobility scooter, Other pedestrian mobility aid, Normal bike, Cargo bike, Long bike, Bike trailer, Mobility bike, Other modified bike, e-scooter
- [Cyclists and e-scooter users only]: Are they a delivery rider?
- Number of children being pushed or carried by the person walking or cycling
- [Children only]: Are they alone, travelling only with other children, travelling with adult(s) but not holding hands with an adult, travelling with adult(s) and holding hands with an adult.
- [Cyclists and e-scooter users only]: Are they wearing a helmet? High-viz clothing? Cycle sportswear?

Appendix 1 provides a copy of the instructions given to video raters, as supplemented by verbal clarification of any points that were unclear. Appendix 2 provides an example extract of the spreadsheet that they completed.

After the above data being collected, a member of the research team identified and coded school groups (e.g. teachers and pupils on an outing together). This was done by manually reviewing the video data whenever, on a weekday between 9am and 3:15pm, 3 or more children aged 4-16 were recorded as passing within the space of a single minute. We did this because sites can show considerable day-to-day variation as to whether a school group is present, and how large the groups is. This has the potential to introduce measurement error / 'noise' when analysing outcomes involving children's travel.

---

[5] We initially sought to distinguish between adults aged 17 to 64 versus aged 65 and over, but comparisons between video coding and roadside enumerator coding indicated that this distinction could not be made with confidence. See Appendix 3 for more details. We also considered whether it would be possible to code ethnicity but decided not to pursue this before the pilot stage. Current government standards for collecting ethnicity data note that using self-reported ethnicity is the best option, and that assigning ethnicity based on visual appearance "will generally be of lower quality than when someone reports their own ethnicity - it might not necessarily reflect the ethnicity the person themselves would respond with." (https://www.gov.uk/government/consultations/standards-for-ethnicity-data/standards-for-ethnicity-data) . e expected this limitation to be particularly acute in the case of trying to assign ethnicity using blurred video data.

Assessing accuracy

Manual classification of pedestrian and cyclist characteristics from videos is expected to yield some measurement error. The nature and extent of any errors is expected to be the same in both our before and after data, and in our intervention and control sites (note that the people classifying the videos were not aware of our study aims or hypotheses, nor were they told which sites were intervention versus control sites). As such, these errors are not expected to be systematically biased. Instead, they are expected to be a form of random measurement error, that can be expected to reduce the size of any observed effects.

We took the following steps to try to minimise this measurement error:

1. Providing clear definitions/instructions, as shown in Appendix 1.
2. Our subcontractors employ raters with extensive experience in the manual classification of videos. Where possible, the same small pool of raters was used to code all data from a given wave of data collection. These subcontractors randomly spot checked 10% of results, investigating and correcting any anomalies found before supplying data to us.
3. The categories ‘Mobility bike’ and ‘Other modified/specialist bike’ were double checked by one of the research team for all baseline and follow-up data collection, to ensure comparability of these unusual and variable bicycle types.

In addition, to quantify the likely magnitude of any remaining measurement error we performed the following two analyses using data collected in 2021. These are summarised below and described in more detail in Appendix 3.

<u>Spot-checking for clear errors</u>

We compared data coded by the subcontractor with data independently coded by a member of the research team for a random sample of 657 active travel users in total (N=306 from across 4 pilot sites in 2021, and by N=351 across 28 main sites in 2021).

For each characteristic for each active travel user, we then determined whether the coding given by the subcontractor:

a) was judged to be correct by the research team. This included a small number of cases where upon review we found a definite error in the research team coding.
b) disagreed with the coding of the research team but the footage was ambiguous, and we were unclear who was correct.
c) did not agree with the research team, and we judged the subcontractor to have made an error.

A high proportion (89-100%, depending on the characteristic in question) of all subcontractor classifications were judged correct by our research team for most characteristics. In addition, where there was disagreement, it was usually judged that this was reasonable in the face of ambiguous footage, with only 0-4% of classifications judged to be definite subcontractor errors.

<u>Calculating subcontractor agreement with manual roadside classification for age and gender</u>

At the 4 pilot sites, enumerators coded the age and gender of cyclists and/or pedestrians in person at the roadside, simultaneous to the video recordings being made and subsequently independently coded. We matched these manual classifications to those made by the subcontractors on the video data and calculated the level of agreement.

We generally found good agreement in individual rating (Kappas >0.65) and also little evidence of systematic difference in the aggregate distributions. The exceptions to this related to coding gender for cyclists. First, we found poor agreement for coding gender of child cyclists aged 4-10 (Kappa=0.37). For older cyclists, the individual level agreement was acceptable (Kappa=0.66) but there was a systematic bias towards the video rater classifying fewer cyclists as female than the roadside enumerator (26% vs 35% for those age 11+).[6] We therefore intend to focus on adults in our analyses of the gender of cyclists. We believe our data may be sensitive to changes in the gender ratio over time, given the acceptable inter-rater reliability with regard to gender and the low rate of clear subcontractor errors. We will, however, assume that the raw proportions may be systematic underestimates of female representation.

Outcomes of interest

Table 2 summarises the outcomes we will look at, and also identifies as 'primary' the outcomes we consider to be of greatest interest.

In selecting our primary outcomes, we have followed these principles:

- Focus on outcomes where the present strand of research is the main source of evidence. For this reason, we have not included 'total number of people walking' and 'total number of people cycling' as primary outcomes here, since we will have substantially more power to address this point elsewhere in our research study using 24/7 data collected from VivaCity sensors.[7]
- Focus on the outcomes that we think are most important, and where we have an *a priori* position as to what a desirable change would look like. For example, we have included the age and gender of cyclists as a primary outcome, because cycling in London is currently very dominated by adult males[8]. London's 'Cycling action plan 2' (TfL, 2023) identifies an increase in the proportion of children and women cycling as key elements of the wider diversification of cycling necessary to deliver on its benefits.[9] By contrast, we have not included the age and gender of pedestrians as a primary outcome, as walking in London is much less demographically skewed and we are not aware of policies seeking to affect the age and gender diversity of pedestrians.

These principles led us to identify three sets of primary outcomes:

1) Use of mobility aids among adult pedestrians. We consider that an increase in this proportion would be desirable, potentially indicating that the streets had become more accessible for disabled people. We also consider that evidence of a decrease in this proportion would be a

[6] Our assumption is that the roadside enumerators coded this more accurately, as they benefitted from a closer view of the cyclists face. By contract, video raters were (for privacy reasons) working with images of low/pixellated quality such that facial features were blurred. In addition, if the cyclist was travelling away from the camera, the video rater would only see the side and back of the cyclist's face. Video raters therefore had to rely to a considerable extent on gender cues such as hairstyle, clothing and bicycle type. We speculate that video raters tended to default to 'male' for passing cyclists who did not have any 'female' gender markers

[7] https://westminsterresearch.westminster.ac.uk/item/w1q51/statistical-analysis-plan-low-traffic-neighbourhoods-in-london-interrupted-time-series-analysis-of-sensor-count-data

[8] See for example the Active travel trends section in https://content.tfl.gov.uk/travel-in-london-2023-active-travel-trends-acc.pdf

[9] https://content.tfl.gov.uk/cycling-action-plan.pdf

cause for concern, potentially indicating that streets had become less accessible for disabled people.[10]

2) Proportion of older children travelling on foot and by bike without adults; proportion of younger children travelling without holding hands. We consider that an increase in these proportions would be desirable, potentially indicating an increase in the independent movement and mobility of children.
3) Diversity of cyclists in relation to age and gender. We consider that an increase in the proportion of women and of children would be desirable, potentially indicating that cycling had become more diverse on these measures.

Note that most of our outcomes related to cycling exclude delivery cyclists. We decided to do this because delivery cyclists are by definition expected to be adults, and we also expected that the percent male and the percent on non-traditional bikes would be higher for delivery cyclists than for other types of cyclists.[11] As such, if LTNs increased the proportion of cyclists who were delivery cyclists this might mask any simultaneous increase in the proportion of non-delivery cyclists who were female or who were children. Our main interest in this research is to examine whether the introduction of LTNs has been associated with any change in the demographic characteristics and clothing of non-professional cyclists.

[10] This possibility was raised by some participants in the qualitative report 'Pave The Way', released by Transport for All in 2021 (https://www.transportforall.org.uk/wp-content/uploads/2023/11/Pave-The-Way-full-report.pdf). Specifically, some participants reported experiencing an increase in traffic danger following the introduction of an LTN due to experiences of 'road rage' and dangerous driving from drivers frustrated with the scheme, or due to dangerous cycling. The qualitative strand of the present NIHR study is examining in more detail how disabled people experience LTNs, and also ways that the schemes could be made more accessible.

[11] Preliminary evidence of this in relation to gender can be found in this study, which found that 91% of delivery cyclists appeared to be male (https://transportforqualityoflife.com/wp-content/uploads/2023/11/cycle-diversity-june-2021.pdf)

Table 2: Summary of our outcomes

Mode	Domain	Outcome	Details	Outcome type	Primary outcome?
Pedestrians	Count	Number of pedestrians	Number of pedestrians, including carried children, and including those wheeling or using mobility aids (e.g. scooters, wheelchairs, mobility scooters). Presented in total, and also split by adult male / adult female / child.	Daily count	No
	Demographics	% female among adults	% of all pedestrians age 17+ who are coded as female, excluding those where gender is coded ‘unclear’.	Binary	No
		% who are children	% of all pedestrians, including carried children, who are coded age 0-16. Repeated excluding school groups, as a sensitivity analysis.	Binary	No
		% using mobility aid among adults	% of all pedestrians age 17+ who are coded as using a wheelchair, mobility scooter, or ‘other mobility aid’ e.g. a stick. (NB it was extremely rare for use of a mobility aid to be recorded in a child: N=1/13,191 in 2021 data across the 4 sites).	Binary	Yes
	Independent mobility	% younger children not holding hands	% children aged 0-10 walking without holding hands, among those who are travelling with an adult but not being carried. Calculated excluding school groups (i.e. teachers escorting pupils).	Binary	Yes
		% older children travelling independently	% children aged 11-16 travelling without an adult (either alone or with other children). Calculated excluding school groups.	Binary	Yes
Cyclists	Count	Number of people on bicycles	Number of people on bicycles, including carried children. Presented in total, and also split by adult male / adult female / child.	Daily count	No
	Delivery cyclists	% delivery cyclists	% of bicycles ridden by delivery cyclists	Binary	No
	Demographics and cycle types (excluding delivery cyclists)	% female among adults	% of people on bicycles age 17+ who are coded as female, excluding those where gender is coded ‘unclear’.. Calculated excluding delivery cyclists	Binary	Yes
		% who are children	% of people on bicycles, including carried children, who appear age 0-16. Calculated excluding delivery cyclists.	Binary	Yes
		% special bicycles	% of bicycles that are special bicycles, e.g. cargo bikes, long bikes, bike trailer. Calculated excluding those coded as delivery cyclists.	Binary	No
	Independent mobility	% older children travelling independently	% children aged 11-16 travelling without an adult (either alone or with other children). Calculated excluding school groups (e.g. cycle training instructors escorting groups of pupils).	Binary	Yes
	Clothes among adult cyclists (excluding delivery cyclists)	% wearing a helmet	% of people on bicycles age 17+ who are wearing a helmet. Calculated excluding delivery cyclists.	Binary	No
		% wearing high viz	% of people on bicycles age 17+ who are wearing high viz clothing. Calculated excluding delivery cyclists.	Binary	No
		% wearing sports wear	% of people on bicycles age 17+ who are wearing cycling sports clothes (e.g. Lycra). Calculated excluding delivery cyclists.	Binary	No
E-scooters	Count	Number of e-scooter users	Number of people travelling by e-scooter. Not split by age and gender due to smaller sample sizes.	Daily count	No
All active travel	Count	Total number of active travel users	Total number of pedestrians, cyclists and e-scooter users. Presented in total, and also split by adult male / adult female / child.	Daily count	No

Statistical analysis

Units of analysis and regression modelling

For daily count data, the units of analysis will be days, as recorded separately at each observation point: thus 4 days before introduction of the LTN in the area, and 4-12 days after. For individual-level binary outcomes, the units of analysis will be people: for example, when looking at the age profile of pedestrians, each pedestrian will be one row of data, individually coded as 'adult' or 'child'.

For each outcome, we will present descriptive statistics for the relevant count or proportion before and after the introduction of the LTN, at the LTN sites and their matched controls.

We will then examine whether there is evidence of an impact of the LTN. We will do using Poisson regression models, fitting an interaction between before/after status and LTN/Control status. This interaction corresponds to a difference-in-difference analysis on the log scale, which we exponentiate to generate a ratio-of-ratios as a measure of change in the relative rate or relative risk.[12]

For daily count outcomes, we will use Poisson regression analysis with robust standard errors, and generate a ratio-of-ratio for change in the relative rate. We will confirm that our findings are consistent when instead using negative binomial regression.

For binary outcomes, we will use Poisson regression analysis with robust standard errors.[13] This has the advantage over logistic regression of approximating risk ratios rather than odds ratios, which are more intuitive to interpret for commonly occurring binary outcomes. This method will therefore allow us to generate a ratio-of-ratio for change in the relative risk.

Covariates for regression analysis

We will adjust our daily count and binary models for the following fixed effects:

- The observation point in question, as a 15-level categorical variable.
- Day of the week ,as a 4-level categorical variable.
- Hour of the day, as a 12-level categorical variable [only for the binary outcome models].
- Calendar year, as a 4-level categorical variable.
- Maximum temperature, as a linear plus quadratic variable.[14]
- Maximum wind, as a linear plus quadratic variable.
- Whether there was any rain, as a binary variable.

---

[12] We also considered conducting difference-in-differences analyses, for change in absolute risk. These analyses sought to use binomial regression model with an identity link, fitting an interaction between before/after status and LTN/Control status. In test analyses, however, these models frequently did not converge after adjusting for our planned fixed effects, such as the interaction between sensor and day of the week. In addition, we consider it *a priori* more likely that an LTN would have a consistent effect on relative change than absolute change. For example, we consider it more likely the implementation of an LTN would increase the share of child cyclists by a consistent relative amount in all sites (e.g. from 2% to 2.2% in a low-child-cycling site and from 20% to 22% in a high-child-cycling site; both an increase of 10% in relative terms), rather than by a consistent absolute amount (e.g. from 2% to 4% in a low-child-cycling site and from 20% to 22% in a high-child-cycling site; both an increase of 2 percentage-points in absolute terms).

[13] https://academic.oup.com/aje/article/159/7/702/71883

[14] Obtained from https://www.timeanddate.com/weather/uk/london/historic . In relation to the daily count models, these variables were defined across the entire day. In relation to the binary outcomes, the weather recording for 06:00 was used for observations in the period 7-8:59am, the weather recording for 12:00 was used for the period 9am-2:59pm, and the weather recording for 18:00 was used for the period 3pm-7pm.

We will include an interaction term between the observation point in question and the day of the week to take account of differences between observation points in their pattern of use across the week (e.g. according to factors like being near to a school or not). In addition, for the daily count model, we will include the number of minutes of recorded data across the day as an offset. This number was usually 12*60=720, but for 2% of days it was 715 because of battery changes to the video cameras.

Pooling results across analyses

We will initially run all regression analyses separately by LTN, to examine how far one sees heterogeneity between different schemes. We will then pool results across LTNs in two ways:

1) Using fixed effects meta-analysis techniques to synthesise the stratified results from across the four LTNs. This approach will estimate the average change in diversity on introduction of the LTN in **these four specific schemes**, and address the question of whether and how such changes varied across those schemes.
2) Using random effects meta-analysis techniques to synthesise the stratified results from across the four LTNs. This approach will estimate **how diversity would be expected to change in a 'typical scheme'.**

In interpreting the results from these models, we will record whether there was evidence of heterogeneity between the four LTNs, as measured using the Cochran's Q test statistic. If there is significant evidence of heterogeneity ($p<0.05$), then we will interpret the fixed-effect results with more caution, and give relatively more weight to the random-effects model than the fixed-effects model.

We will initially run these models for all four LTNs, and then conduct sensitivity analyses which exclude Camden Square as this is a much smaller scheme than the other three.

Example Stata code

An example of our Stata code for one of our daily count outcomes is as follows. In this analysis, days are the unit of analysis, and the outcome is a daily count integer, for example 'number of pedestrians that day'.

Step 1: stratified analyses:

```
xi: poisson count_outcome i.ltn*i.prepost i.observationpoint*i.dow i.year temperature
temperaturesq wind windsq rain if site==1, vce(robust) offset(lndailyminutes)
```

…and then likewise for sites 2, 3 and 4

Step 2: strata-specific estimates for the point estimate, lower confidence interval and upper confidence interval then extracted from the four models, and combined using the following code to run a) fixed-effect and b) random-effect models

```
metan point lci uci, eform
```

```
metan point lci uci, random eform
```

An example of our Stata code for one of our binary outcomes is as follows. In this analysis, individuals are the unit of analysis, and the outcome is a binary variable, for example 'child versus adult'.

Step 1: stratified analyses:

```
xi: poisson binary_outcome i.ltn*i.prepost i.observationpoint*i.dow i.hour i.year
temperature temperaturesq wind windsq rain if site==1, vce(robust)
```

…and then likewise for sites 2, 3 and 4

Step 2: strata-specific estimates for the point estimate, lower confidence interval and upper confidence interval then extracted from the four models, and combined using the following code to run a) fixed-effect and b) random-effect models

```
metan point lci uci, eform
```

```
metan point lci uci, random eform
```

Acknowledgements:

Thanks to Joe Maclaren, Jigisha Parekh, and other staff at CTS Traffic and Transportation Ltd for their work in collecting and coding the video data.

Appendix 1: Instructions for manual classification of Pedestrians and Cycles

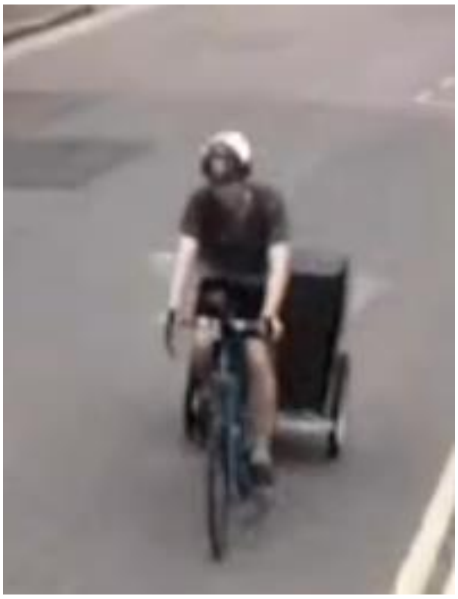
- Record the characteristics of each pedestrian or cyclist across the count line in either direction. Include pedestrians on the roadway, cycles on the footpath. Each person walking/wheeling/pedalling gets one line (children who are being carried are coded on that same line, see below). People can be counted even if just walking a very short distance, e.g. from their home to their car, if they cross the countline.
- If someone walks or cycles back-and-forth across the countline then count them each time they cross it – e.g. a binman going back-and-forth collecting bins, or a local resident who walks out to collect their bins and the returns home. If someone happens to spend ages standing/pacing right on the countline – e.g. on a phone call or standing in an outdoor queue - they can be counted just once, rather than a separate count each time they may sway one side of the line versus the other.
- Record the time they pass the count line.
- if visibility is temporarily very poor – e.g. a bin lorry is temporarily in the way – then it is better to skip people than trying to code them based on inadequate footage, e.g. the view of them in the far distance. I.e. quality more important than comprehensive coverage.
- coding age
  - use school uniforms to help. Pre-school = 0-3, primary school = 4-10, secondary school = 11-16.
- coding sex: male / female / unknown. Particularly for teenagers and adults, try to make a good guess if you possibly can.
  - **walk**: include here people walking while pushing a bicycle or anything else, e.g. a pushchair, wheelchair. Do not use this for pedestrians with an apparent mobility aid – they are coded as other ped mobility aid.
  - **jog**: jogging or running when they cross the count line. Do not count someone who is in jogging clothes but exclusively walking. Do count someone running while pushing something, e.g. a pushchair.
  - **Scoot/skate/rollerblade** – any other type of pedestrian not using a mobility aid.
  - **Wheelchair**
  - **Mobility scooter**

  - **Other ped mobility aid** – any apparent mobility aid apart from a wheelchair/mobility scooter, e.g. a walking stick, white cane for visual impairment

  - **Cargo bike** = the bike has a built-in substantial section for moving goods/kids, usually at the front of the bike. Could be 2 or 3 wheeled. Do not count baskets or panniers attached to an otherwise normal bike, i.e. things not integral to the structure. Examples of cargo bikes are below

- **Long bike** = built to be long, e.g. two kids can sit beside not just one, or a tandem. Examples below.

- **Trailer bike** = the bike has an additional unit attached, which could be pulling goods or kids. This would include a tagalong attachment that hooks a child's bike, back of an adult bike. Examples below



- **Mobility bike** = A wide range of designs, the common denominator being that they are adapted for use by disabled people. They differ from mobility scooters in that they involve someone using a foot or hand pedal. One example is below, see https://wheelsforwellbeing.org.uk/ for a wider range of examples.

- **Other modified/special bike** = any bike that does not fit in the above, a researcher will check all of these

- **Escooter** = electric powered scooters that you stand upright on.

- **Delivery rider** = Should be ticked AS WELL AS bicycle/escooter type – i.e. identify both whether the person is a delivery rider AND whether they are travelling on a normal bike, cargo bike, trailer bike, escooter etc. This covers anyone who seems to be cycling/scooting for the purpose of transporting things for business/commercial reasons. This could include:
  - take away food delivery – Deliveroo, Just Eat etc
  - delivering goods for a shop, e.g. Pop Florist delivering flowers, pedalme delivering a fridge

  Do not count a bike that just has a lot of stuff on it, unless there is an indication it is for business not personal transport. Do count a specialised delivery bike even if seems to be empty. Some examples are images below.

- **No. kids carried / pulled : E**.g. in arms, in a sling, in a pushchair, in a bike seat or trailer.

- **Children (age 0-16) only** – whether accompanied.
    - **Alone** = not seeming to be with anyone
    - **Kids only** = seem to be accompanied by other people aged <18, but no one aged 18+
    - **With an adult, not hold on** = seem to be accompanied by an adult (there may also be other kids and adults in the group), but not holding hands or otherwise physically attached to them
    - **With an adult, holding on** = holding hands with an adult, or otherwise attached to them (e.g. child reins, holding onto a push chair the adult is pushing). NB please apply this specifically at the countline, not e.g. for crossing a nearby road.

- **Cyclists & escooters only** – type of clothing worn.
    - **Helmet**
    - **High viz – see below**
    - **Sports wear – see below**

Hi-Viz

fluorescent or reflective clothes worn over the torso area that seem to be specifically designed for high visibility. So YES to a high viz jacket, vest, body straps, rucksack cover. No to a jacket that was just brightly coloured. NO to high viz items on other body parts, e.g. ankle straps, stripes on shorts, pannier, helmet.

WOULD INCLUDE

Hi-viz jacket and helmet

Hi-viz backpack cover

Hi-viz backpack cover and jacket

Hi-viz gillet

Hi-viz tabard or Sam Browne belt

Hi-viz vest open at front

NOT HIGH VIZ

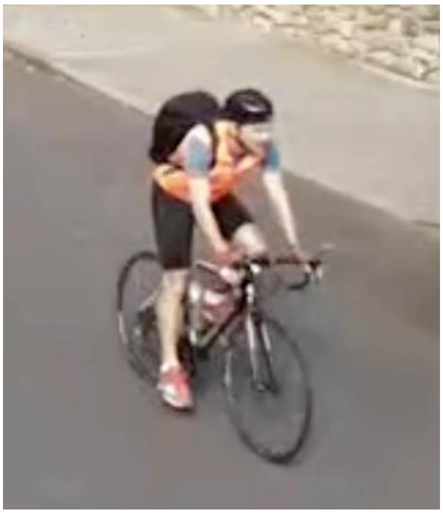
Hi-viz rack pack on bike, but not clothing or on backpack

Bright colour top (light mustard) but not hi-viz

Sportswear

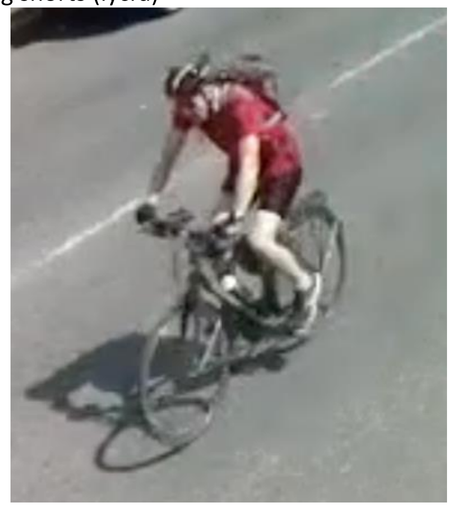
Sportswear means shorts, leggings or tops made from spandex/lycra type material and that look specially designed for cycling. It does not cover sportswear that is for other sports (e.g. cricket whites, jogging bottoms, yoga outfit) or shorts which are casual rather than sporty (e.g. denim / hot weather shorts). One rule of thumb is to ask whether the person is likely to wear those clothes only for cycling, or whether it would also be equally unsurprising to see them wearing these clothes for walking.

WOULD INCLUDE



Cycling shorts (lycra)

Cycling shorts and jersey (lycra)



Cycling Shorts (lycra) with casual t-shirt

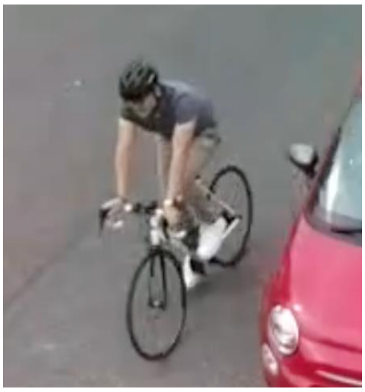
Cycling Shorts (lycra) with casual t-shirt

Cycling shorts and jersey (lycra) and cycling shoes

WOULD NOT INCLIDE

Looks more like yoga clothes than cycling clothes?

Denim shorts – not sportswear



T-shirt and trainers, but worn casually with trousers – not sportswear

Outdoors jacket with casual trousers and trainers – not sportswear

Outdoors jacket and baggy shorts – not sportswear

Edge case: but shorts look probably baggy not lycra, so don't include

Appendix 2: Example of data coding scheme

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
Date	Time	Age				Sex			Mode (tick all that apply)													Carried kids	ren only): with adult?				s + escooters only): what		
		0-3	4 to 10	11 to 16	17 plus	Male	Female	Unsure	Walk	Jog	Scoot t/skate/rollerblade	Wheelchair	Mobility scooter	Other Ped mobility aid	Normal bike	Cargo bike	Long bike	Bike trailer	Other modified bike	Delivery rider?	escooter	No. children pulled/carried	Alone	Only with kids	With adult, not hold hands	With adult, hold hands	Helmet?	High Viz?	Sports wear?
30/06/2021	07:02:21				1		1		1																				
30/06/2021	07:02:42				1	1									1												1		1
30/06/2021	07:04:10				1		1		1																				
30/06/2021	07:05:06				1	1									1												1		1
30/06/2021	07:07:06				1	1									1														
30/06/2021	07:08:04				1	1									1					1									
30/06/2021	07:09:38			1		1			1														1						
30/06/2021	07:14:01				1		1		1																				
30/06/2021	07:14:45				1		1								1												1	1	
30/06/2021	07:15:56				1	1									1												1		

Appendix 3: Assessment of the accuracy of manual classification of video data for pedestrian and cyclist characteristics

Manual classification of pedestrian and cyclist characteristics from videos is expected to yield some measurement error. This is particularly true since for privacy reasons the video images are somewhat pixelated such that one cannot recognise individuals. In addition the videos are mounted approximately 5 metres above the ground, and this aerial view gives a less clear view of people's faces.

The nature and extent of any errors is expected to be the same in both our before and after data, and in our intervention and control sites (note that the people classifying the videos were not aware of our study aims or hypotheses, nor were they told which sites were intervention versus control sites). As such, these errors are not expected to be systematically biased. Instead they are expected to be a form of random measurement error, that can be expected to reduce the size of any observed effects.

We took the following steps to try to minimise this measurement error:

4. Providing clear definitions/instructions, as shown in Appendix 1.
5. Our subcontractors employ raters with extensive experience in the manual classification of videos. These subcontractors also randomly spot check 10% of results, investigating and correcting any anomalies found before supplying data.
6. The category 'other modified/specialist cycle (including mobility bike)' was double checked by one of the research team for all baseline and follow-up data collection, to ensure comparability of these unusual and variable bicycle types.

In addition, to quantify the likely magnitude of any remaining measurement error we performed the following analyses.

1) Spot-checking for clear errors

At 4 pilot sites, the subcontractor coded videos on weekdays for three 2-hour periods: 08.00-10.00, 11.00-13.00, and 15.00-17.00. A member of the research team independently re-classified a random sample of 306 active travel users, and then subsequently compared these to the subcontractor coding. For each characteristic for each active travel user, we then determined whether the coding given by the subcontractor

d) was judged correct by the research team. This included a small number of cases where upon review we found a definite error in the research team coding.
e) disagreed with the research team but the footage was ambiguous, and we were unclear who was right.
f) did not agree with the research team, and we judged the subcontractor to have made an error

The distinction between b) and c) was made by two members of the research team in discussion.

A shown in Table 3, a high proportion (87-100%) of subcontractor classifications were judged correct by our research team for most characteristics. In addition, where there was disagreement it was usually judged that this was reasonable in the face of ambiguous footage. Across most variables, only 0-5% of classifications were judged to be definite subcontractor errors. A somewhat lower accuracy was seen for age and gender (1-5% clear error) than for sub-mode, carried children, independent mobility, or helmet use (0-3% clear error).

The only characteristic that did not show adequate accuracy was cyclist clothing, where the subcontractor was judged to have made an error on the high viz or sports status of the clothing in 23% of cases. This was subsequently improved substantially through clarification, with adequate accuracy achieved ('High viz / sports [V2]').

Note that we did not observe any escooters in the data used for this comparison. Here and in the following section, we believe the values for cyclists provide a realistic guide for accuracy in classifying escooters, as our experience from reviewing the videos is that all the characteristics identified are observable to a similar degree in both modes.

Table 3: Proportion of subcontractor errors identified for different characteristics at 4 Pilot sites

Mode	Characteristic	No. people	No. correct	No. unclear	No. subcontractor error
Pedestrian	Age	194	176	16	2
	Gender	194	184	7	3
	Sub-mode	194	191	1	2
	No. children carried	194	192	2	0
	Independent mobility	48	48	0	0
Cyclist / escooter	Age	112	98	8	6
	Gender	112	97	12	3
	Sub-mode	112	111	0	1
	No. children carried	112	110	1	1
	Independent mobility	17	16	1	0
	Helmet	112	109	0	3
	High viz / sports [V1]	112	75	11	26
	High viz / sports [V2]	112	99	9	4

Similarly during 2021 baseline data collection, we randomly sampled 18 5-minute segments of video data from across the 28 sites covered. A member of the research team independently re-classified 355 active travel users present in these segments. We found N=4 instances where a genuine active travel user had been missed by the subcontractor (i.e. they were present in the video but not entered on the spreadsheet), and N=4 instances where a non-existent active travel user had been recorded by the subcontractor (i.e. they were absent in the video but entered in the spreadsheet. In all four cases this seemed to reflect double-counting of genuine active travellers).

Among the remaining 351 active travel users recorded by both the researcher and the subcontractor, our findings are shown in Table 4. As this shows, across most variables, 0-3% of classifications were judged to be definite subcontractor errors. The exceptions were for gender in relation to cyclists and escooter users (5% errors, comprising 1 apparent man incorrectly coded as a woman, and 2 apparent women incorrectly coded as men) and sportswear in relation to cyclists and escooter users (5% errors, comprising 3 cyclists who were wearing shorts that were above the knee but that did not meet our strict definition of 'sportswear' as the shorts did not appear to be made of spandex/Lycra). This suggests that in general the coding by the subcontractor is of high quality, but that there may be some measurement error around the definition of 'sportswear'.

Table 4: Proportion of subcontractor errors identified for different characteristics across 18 randomly selected 5-minute segments of time from the 2021 data collected

Mode	Characteristic	No. people	No. correct	No. unclear	No. subcontractor error
Pedestrian	Age	292	283	8	1
	Gender	292	286	1	5
	Sub-mode	292	291	0	1
	No. children carried	292	292	9	0
	Independent mobility	37	36	0	1
Cyclist / escooter	Age	59	54	5	0
	Gender	59	55	1	3
	Sub-mode	59	57	0	2†
	No. children carried	59	59	0	0
	Independent mobility	3	3	0	0
	Helmet	59	58	0	1
	High viz	59	59	0	0
	High sports	59	56	0	3

†These N=2 instances were both cases where a cyclist pushing a bicycle on the pavement had incorrectly been classified as 'cyclist' not 'pedestrian'.

2) Subcontractor agreement with manual roadside classification for age and gender

At the four pilot sites, enumerators coded cyclists and/or pedestrians in person at the roadside, simultaneous to the video recordings being made and subsequently independently coded. At each site, enumerators made roadside observations for two 2-hour periods for cyclists (from 08.00-10.00, and from 11.00-13.00) and for two 2-hour periods for pedestrians (from 11.00-13.00 and from 15.00-17.00). At busier sites, roadside enumerators only coded pedestrians on one side of the pavement. All comparisons were made on weekdays. The roadside enumerators focused on capturing information on age and gender, which we judged to be the characteristics most likely to be difficult to discern in a video. This was subsequently borne out by the video comparisons described in the previous section.

We matched these manual classifications to those made by the subcontractors on the video data using each individuals recorded time stamp and comparing manual notes made on characteristics (e.g. 'red top') against the video footage. We were able to match 94% of individuals rated by the subcontractor to manually classified individuals (with the remaining 6% being uncertain cases). We identified 4/886 pedestrians and 3/738 cyclists who were missed by the subcontractor video rater despite being identified by the manual rater and confirmed on the video footage. In addition, 2 cyclists (and 0 pedestrians) were double-counted by the video raters.

We calculated the level of agreement between the roadside enumerator and the subcontractor video rater for matched individuals. We did this first at the level of the individual, in terms of the percent agreement (i.e. proportion of all individuals receiving the same classification), and also using the Cohen's Kappa statistic (which has the advantage of taking into account the possibility of the agreement occurring by chance: see Table 5). We then calculated agreement at the aggregate level by comparing the two methods in terms of the frequency distribution of characteristics (Table 6).

There was generally good agreement between the roadside enumerator and the video rater in judging age group except with respect to the category 65+. This showed poor agreement at the individual level (Kappa=0.39: Table 5) and also at the aggregate the level (6% pedestrians judged to be age 65+ by the roadside enumerator versus 2% by the video rater: Table 6). Even worse agreement was seen with regard to the category 65+ for cyclists, with individual-level agreement being no better than chance. In addition, feedback from both the roadside enumerators and the video raters was that they found this the hardest characteristic to judge. **We therefore determined that older age cannot be identified with sufficient accuracy, and determined not to use this category in our further work.**

The other problematic characteristic was gender for cyclists. First, for younger children on bicycles there was not good agreement on gender at the individual level (Kappa=0.37: Table 5). For older cyclists, the individual level agreement was acceptable (Kappa=0.66: Table 5) but there was a systematic bias towards the video rater classifying fewer cyclists as female than the roadside enumerator (26% vs 35% for those age 11+, Table 6; or 27% vs 36% if also including children aged <11).

A similar systematic skew towards video raters identifying fewer cyclists as female than roadside enumerators was also observed at a different data collection site in Southwark: at this location, two 12-hour roadside counts in March and April 2021 estimated that 37% of adult cyclists were female,[15] but two 12-hour video-rater counts in June 2021 estimated that only 31% of adult cyclists were female.[16] Based on a review of video footage, we believe that this discrepancy likely reflects 'male' being assigned by default to cyclists who have no obviously 'female' gender markings that are visible to a video rater, and whose detailed facial characteristics are not observable by the video rater (because of video pixelation for privacy reasons, and also because some cyclists are travelling away from the video). **We have therefore decided to continue to collect gender data for cyclists, but to focus only on how the gender distribution changes at sites over time.** We believe our data may be sensitive to such changes over time, given the acceptable inter-rater reliability with regard to gender (Table 5) and the low rate of clear subcontractor errors (Table 3). We will, however, assume that the raw proportions will be systematic underestimates of female representation.

[15] http://www.transportforqualityoflife.com/u/files/1_CycleDiversity_June2021.pdf .
[16] Data supplied courtesy of Southwark Council.

Table 5: Agreement for age and gender between roadside enumerator and video rater

	Variable	Levels	Ages included†	Agreement roadside enumerator vs video rater	
				% agreement	Cohen's Kappa
Pedestrians (N=825)	Age, all categories	1-10, 11-16, 17-64, 65+	All	89.7%	0.76
	Age, young child	1-10, 11+	All	98.4%	0.93
	Age, child	1-16, 17+	All	96.1%	0.89
	Age, elderly	17-64, 65+	17+	93.9%	0.39
	Age, 3 categories	1-10, 11-16, 17+	All	94.6%	0.85
	Gender for young children	Male, female	Age 4-10	87.2%	0.74
	Gender for teens	Male, female	Age 11-16	91.4%	0.82
	Gender for adults	Male, female	Age 17+	93.0%	0.86
	Gender for teens + adults	Male, female	Age 11+	92.9%	0.86
Cyclists (N=698)	Age, all categories	1-10, 11-16, 17-64, 65+	All	91.0%	0.63
	Age, young child	1-10, 11+	All	99.3%	0.94
	Age, child	1-16, 17+	All	93.7%	0.71
	Age, elderly	17-64, 65+	17+	97.7%	-0.00
	Age, 3 categories	1-10, 11-16, 17+	All	93.0%	0.69
	Gender for young children	Male, female	Age 4-10	68.8%	0.37
	Gender for teens	Male, female	Age 11-16	100%	1.00
	Gender for adults	Male, female	Age 17+	84.4%	0.64
	Gender for teens + adults	Male, female	Age 11+	85.4%	0.66

†When age subgroups are used, the analysis is restricted to individuals where both raters agree on the age category – e.g. where both raters agree the person is age 17+. The category 'unknown' was not used by manual raters and was used only 3 times by video raters, and so we excluded these people from our comparisons. The categories 0-3 and 4-10 were combined by the manual raters, and so are combined in these analyses.

Table 6: Distribution of age and gender recorded by the roadside enumerator and video rater

	Variable	Levels	Roadside enumerator	Subcontractor
Pedestrians	Age	1-10	14%	14%
		11-16	10%	7%
		17-64	70%	77%
		65+	6%	2%
	Gender for age 11+	Male	46%	47%
		Female	54%	53%
Cyclists	Age	1-10	7%	7%
		11-16	5%	6%
		17-64	86%	87%
		65+	2%	0.1%
	Gender for age 11+	Male	65%	74%
		Female	35%	26%