From injury to risk: Using case-crossover and case-control methods to study cycling injury risk in London

Dr. Rachel Aldred
Reader in Transport
University of Westminster
rachelaldred.org  @RachelAldred
Acknowledgements

- Co-authors: Anna Goodman, James Woodcock, John Gulliver

- Collaborators: David Morley, TfL (datasets), Rahul Goel (comments)

*Paper published in Accident Analysis & Prevention*
Injury risk analysis  

Cycling injury risk analysis limited by many problems, two in particular

- Often lack of good data on infrastructure/road environment characteristics, particularly over time

- Usually no good data on cycling exposure/flows – so we can look at characteristics of injury sites but not risk
Injuries (Stats19)
Measuring risk at a route segment level

- New approaches learn from epidemiological research
- Involves comparing places where injuries happen to where they don’t happen (i.e. control group)
- The control group of places represents where we might expect injuries to happen, if all roads were equally risky
  - i.e. we need some measure of where people are cycling to create the control group
The case-control method

- Aggregate-level

- Needs two data sources: data on where injuries happen and data on where people cycle (from which we generate control sites)
  - Usually don’t have the latter, but in London TfL have the ‘Cynemon’ model
In this study, 6,244 injury sites come from Stats19 police injury data for London; 6,046 control sites were randomly selected from Cynemon model, based on cycling volumes.

Time period covered: 2013-4, weekdays 7am-7pm.
Identifying route environment characteristics

Control and injury sites matched to a range of available datasets using QGIS

- Road classification (OS)
- Junction status (OS)
- Motor traffic volumes (via Imperial College London model)
- TfL 2014 speed limit data
- TfL 2015 bus lane data
- ONS deprivation data (area-level income data)
- London borough boundaries
- Also Cynemon for cycling flows at segment level

*could not include new cycle infra*
Data analysis

- Descriptive analysis: comparing the balance of control and injury sites, based on a particular factor/variable
  - A higher proportion of injury sites than we would expect (as measured via control sites) suggests a higher risk at places where that factor is present, than places where it isn’t
Road type (primary = ‘main road’)

- Residential
- Tertiary
- Secondary
- Primary
- Unclassified

Injury sites (%) vs Control sites (%)
Speed limits (mph)

<table>
<thead>
<tr>
<th>Speed Limit (mph)</th>
<th>Injury sites (%)</th>
<th>Control sites (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>30</td>
<td>80</td>
<td>70</td>
</tr>
<tr>
<td>40+</td>
<td>5</td>
<td>10</td>
</tr>
</tbody>
</table>

- Red: Injury sites (%)
- Green: Control sites (%)
Motor vehicle volumes (vehicles per day)

Injury sites (%)  Control sites (%)

- <2000
- 2000-9999
- 10,000-19,999
- 20,000-29,999
- 30,000+
Cyclist volumes (cycles per day)

- <1000
- 1000-1999
- 2000-2999
- 3000-3999
- 4000+

Injury sites (%)
Control sites (%)

Legend:
- Injury sites (%)
- Control sites (%)
Data analysis

– Descriptive analysis: comparing the balance of control and injury sites, based on a particular factor/variable
  – A higher proportion of injury sites than we would expect (as measured via control sites) suggests a higher risk at places where that factor is present, than places where it isn’t
– Regression modelling: separating out the impacts of different factors
  – Three different regression models built, here presenting the one using all variables
Road type

Road type (residential = 1)

<table>
<thead>
<tr>
<th>Road type</th>
<th>Odds ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tertiary</td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>2.0</td>
</tr>
<tr>
<td>Primary</td>
<td>1.5</td>
</tr>
<tr>
<td>Unclassified</td>
<td></td>
</tr>
</tbody>
</table>
# Odds ratios for intersections and speed limits

## Intersection (95% confidence intervals)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1</td>
</tr>
<tr>
<td>Yes</td>
<td>3.33 (3.07-3.61)</td>
</tr>
</tbody>
</table>

## Speed limit (mph, 95% confidence intervals)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>1.26 (1.12-1.42)</td>
</tr>
<tr>
<td>40+</td>
<td>0.64 (0.46-0.89)</td>
</tr>
</tbody>
</table>
Safety in numbers

Injury odds vs. cycles per day

Injury odds (relative to 1,000 cycles per day) vs. Cycles per day
Danger in (motor vehicle) numbers

Injury odds vs. motor vehicles per day

Injury odds (relative to 10,000 MVPD)
Conclusions

These data suggest that speed limits of 20 mph may reduce cycling injury risk, as may motor traffic reduction. Further, building cycle routes that generate new cycle trips should generate ‘safety in numbers’ benefits.

Need for more injury risk studies including more infrastructural characteristics (especially new cycle infrastructure, but also car parking presence/ control) and covering more contexts…
Next steps: a case-crossover analysis for all UK

- Individual-level
- Need one data source – injured cyclists and the route they followed
- New project: uses algorithms to impute routes, based on cycle trips that are very likely to have started at home locations (on which we have data)
- Then matching to spatial data and regression analysis as with case-control study (but control and injury points are matched for each individual, allowing interaction analysis)
- Funded by the Road Safety Trust
From injury to risk: Using case-crossover and case-control methods to study cycling injury risk in London

Dr. Rachel Aldred
Reader in Transport
University of Westminster
rachelaldred.org  @RachelAldred