Pedestrian and Cyclist Death Rates: a Comparison between several Major Cities

Ian Dennis Miller

October 16, 2016

1   Introduction

Working Paper.

During the Summer of 2016 in Toronto, more than 1,000 pedestrians and cyclists were struck by automobiles (Spurr 2016). This finding grabbed headlines because, at a rate of nearly a dozen collisions per day, the number sounds shockingly high. In the previous year, Toronto saw more deaths by motor vehicle (CBC News 2015) than by gun (Gillis 2016), again shocking the city with a death toll that *seemed* high.

But is the automobile death toll *unreasonably* high? After all, isn't a bustling city with millions of trips per day bound to lead to a few collisions, if only by accident? Furthermore, how do we know Toronto even has a problem in the first place? It's possible that comparable cities have similar or even worse fatality rates. It's even possible these deaths are the inevitable cost of civic progress.

The question of pedestrian and cyclist safety is particularly important right now. There is a widely shared "gut feeling" that Toronto streets are becoming more chaotic (Hain et al. 2016), but there is uncertainty about the fundamental facts of the debate. This report seeks to provide a comparison between Toronto and other cities on the basis of pedestrian and cyclist fatalities. By looking at Toronto in the context of other major cities across North America and Europe, it will be possible to determine whether Toronto is relatively safe or not.

To perform this comparison, yearly pedestrian and cyclist fatality data were collected from seven major cities. The historical results were plotted and trend lines were fitted to the data, permitting a projection into the near future. The results of this report suggest that Toronto has become more dangerous to pedestrians and cyclists during the previous seven years. During the same time period, several other cities in the study exhibited safety improvements. If these trends continue for another 4 years, Toronto will become significantly more dangerous to pedestrians and cyclists than every other city in the study.

2   Methods

Toronto pedestrian and cyclist fatality data were collected for the 7-year period of 2009-2015. Similar data were collected for comparable cities. The cities in this study are:

- Amsterdam (NL)
- Chicago (US)
- Copenhagen (DK)
- London (UK)
- Los Angeles (US)
- New York City (US)
- Toronto (CA)

Four US and Canadian cities were selected for being the largest cities among those countries. London was selected partly for its political relationship to Canada, but otherwise for convenience since London transit data are extremely simple to obtain. Finally, Amsterdam and Copenhagen were selected for their good transit safety reputations, which would permit a baseline comparison.

2.1 Data Collection

The following data fields were collected for each year and city in this study:

- city
- metro area population
- year
- number of pedestrian fatalities in each year
- number of cyclist fatalities in each year

In addition to raw fields, citation data were collected for each observation. These data are stored as Comma Separated Values, and the resulting CSV file is available for download. Individual citations for every data point are available in the CSV file; data citations are not listed in this article's references.

2.2 Population and Road Use

Both city population and metro area population were collected for consideration. Population estimates provided by Wikipedia were held constant across all years for each city. City populations ranged from 0.6 million to 8.7 million (14.5:1), whereas metro area populations ranged from 2.0 million to 23.7 million (11.85:1). Between these population measures, the metro area provides a more regular range from smallest to largest city.

It is a common pattern for people to commute into cities during the workday even though they do not live in the city. Larger metro areas will generate more trips, so a population measure that includes outlying areas will be a better proxy for actual traffic and, therefore, opportunities for automobile, pedestrian, and cyclist interactions. For these reasons, Metro area population is used as an indicator of road use. Yearly *per capita* fatality rates were calculated based on Metro Area populations, and are reported as rates per 100,000 residents throughout the results.

2.3 Challenges

The task of data collection presented several challenges: 1) language translation; 2) local traffic safety data reporting policies; and 3) length of historical record.

The issue of language can be overcome with automated translation software. Traffic safety data varies across international jurisdictions due to subtle differences in the definition of key terms like "collision" and "cyclist." Another factor affecting the traffic safety record is the agency responsible for recording that data. In the case of the City of Amsterdam, this responsibility shifted from the police to the health system, resulting in a "hiccup" in the available data during the transition. In other cases, city-wide traffic mortality data were not stratified by transit method to permit the differentiation of cyclists from pedestrians. In yet other cases, there may have simply been no policy to record pedestrian and cyclist fatality data until recently. Together, these factors limit the number of years of data available for various cities.

2.4 Reproducibility

The R Statistical Environment (R Core Team 2016) was used for analysis and plotting. All scripts used to generate this report are available online[1]. Together with the raw data, which are available from the same source, this report may be reproduced in its entirety.

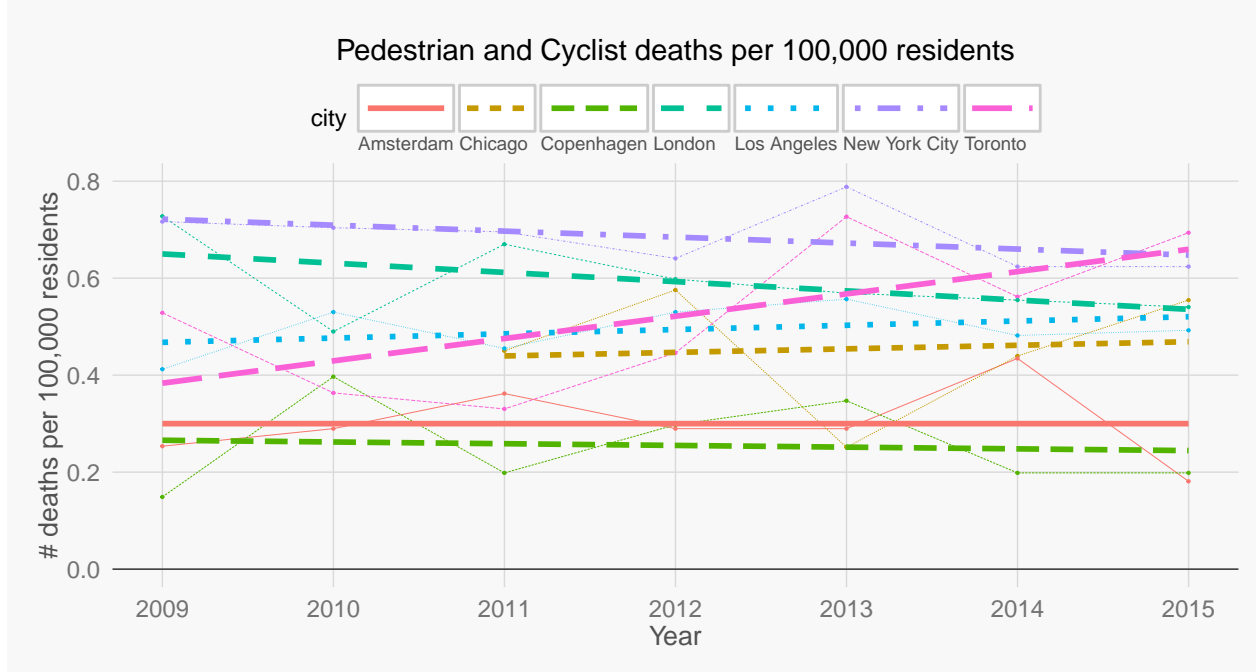---

[1]https://github.com/iandennismiller/road-safety/

Figure 1: Historical pedestrian and cyclist fatality data from 2009-2015 were plotted for seven major cities. Actual scores are plotted as points, which are then connected by thin lines for each city. Thick trend lines indicate the general direction of the raw data.

3 Results

The general analytic approach for this work is to visualize the historical record then fit trend lines to each city. This approach permits the basic exploration of past years as well as rudimentary predictions for the near future. Because there are many measurements for each city, it is possible to isolate variance both within- and between-cities. When this variance is visualized as 95% confidence interval bands, cities may be directly compared to determine the statistical significance of any differences.

3.1 Historical trends

The rate of pedestrian and cyclist fatalities per 100,000 residents are plotted in Figure 1. The trend line for each city is fitted to the raw values using a simple least squared error method (Wickham and Chang 2016)[2]. Due in part to the basic estimation process, the trends for New York City and Toronto appear to intersect in 2015. This result indicates Toronto is more dangerous to pedestrians and cyclists than each other city on a per capita basis, both in terms of the actual fatality rate in 2015 and in terms of the fatality rate trend.

3.2 Future trends

A more sophisticated regression model was built with the nlme library (Pinheiro, Bates, and R-core 2016) in order to obtain a more accurate line estimate that properly accounts for both within- and between-city variance. The fitted city intercepts and slopes are available in Table 1.

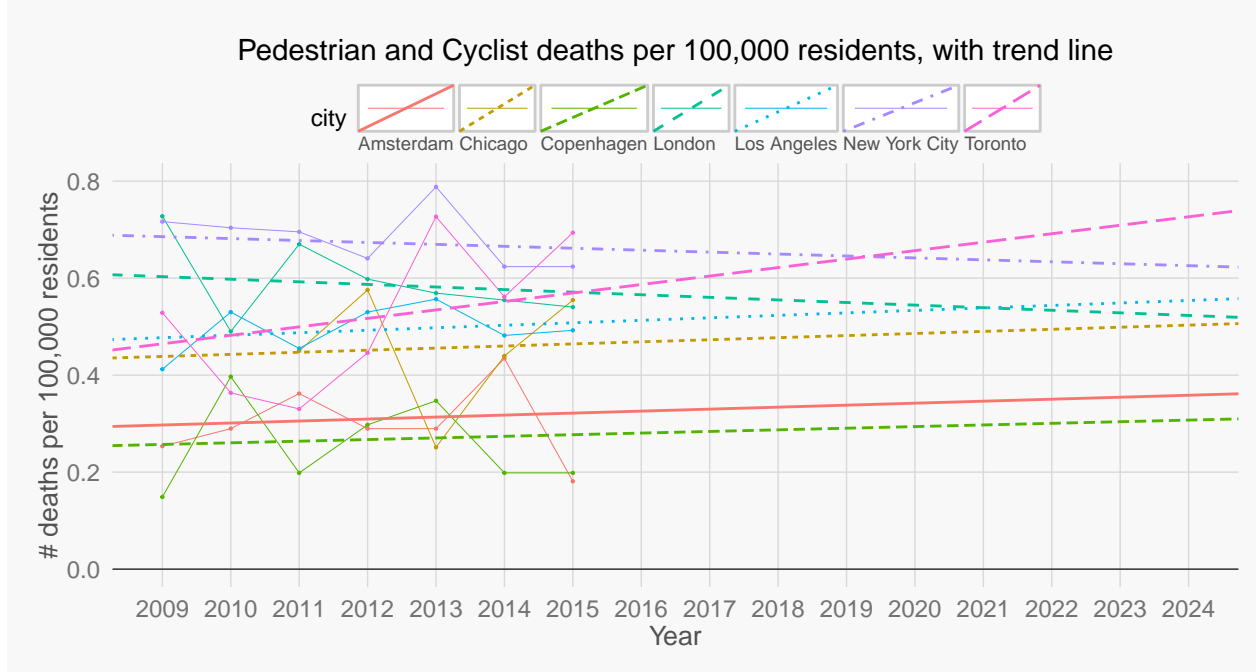[2]ggplot2 includes convenient options for quickly fitting a line to a set of points

Figure 2: Historical pedestrian and cyclist fatality rate trends are extrapolated into the near future.

$$fixed = \text{fatalities} \sim 1 + \text{year}$$
$$random = \quad \sim 1 + \text{year} \mid \text{city}$$

The fitted results of this model are plotted in Figure 2. The trend lines were then extended into the future in order to predict pedestrian and cyclist fatality rates. If these trends hold for the next four years, then Toronto is on track to become more dangerous to pedestrians and cyclists than every other city in the study. By 2015, Toronto has already become indistinguishable from London, and by 2019, Toronto matches New York City's rate. During the same prediction interval, both London and New York City are seen to become safer for pedestrians and cyclists.

3.3 Individual Cities

Because we have many observations for each city, it is possible to use the variance between these observations to test for statistically significant differences in average fatality rates between cities. A basic *t-test* of this kind may visualized with the help of 95% confidence interval bands that surround each city's trend line. When the confidence interval from one city does not contain the trend line of the other city, then the city fatality rate trends are significantly different. This sort of analysis may be interpreted similarly to a *t-test*.

3.3.1 Toronto and New York City

In 2009, New York City had the greatest pedestrian and cyclist fatality rate among all cities included in the study, with an average of 0.685 pedestrian and cyclist fatalities per 100,000 residents. The 7-year fatality rate trend for New York is slightly negative (slope = -0.003), whereas the 7-year fatality rate trend for Toronto is positive (slope = 0.017). These trends, along with raw data, are visualized in Figure 3.
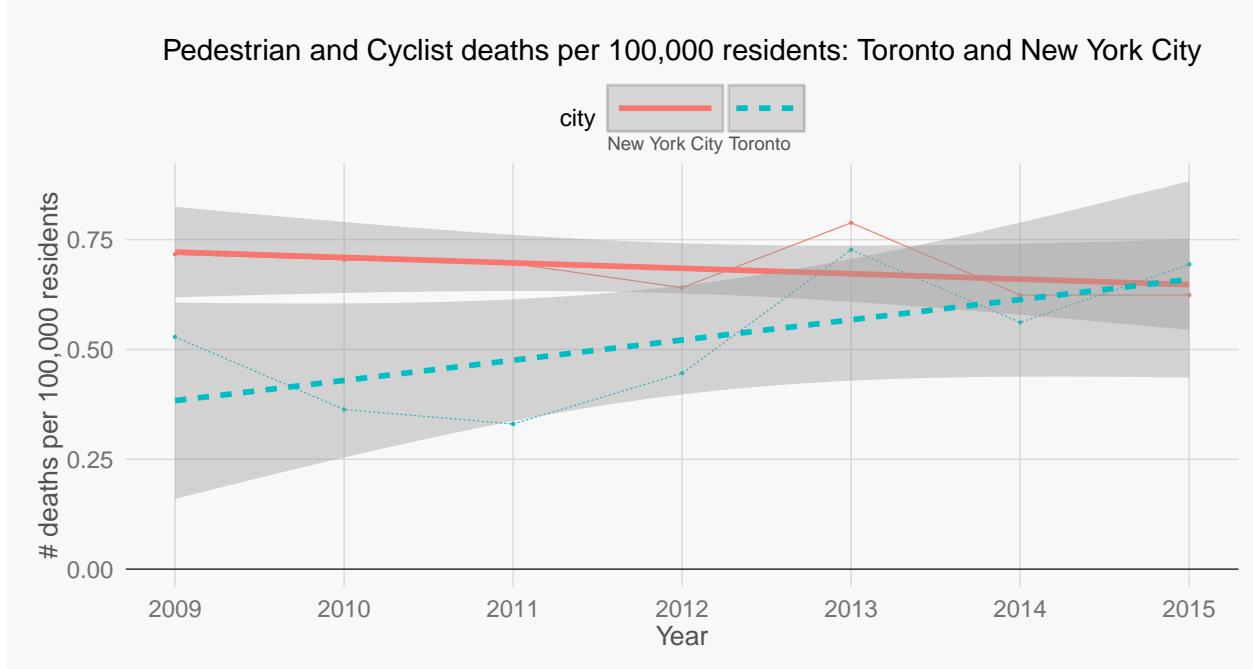
Figure 3: Toronto and New York City pedestrian and cyclist fatality rates are visualized, along with 95-percent confidence intervals.

Toronto caught up to New York City in 2013, at which point the two cities were no longer significantly different in terms of fatality rates. By 2015, Toronto actually surpasses New York City with the highest fatality rate that year, although the difference is still not yet statistically significant. When the trend lines are extended into the near future, we would expect to find that Toronto will become significantly more dangerous for pedestrians and cyclists than New York City. As New York City was previously the most dangerous, Toronto is therefore trending to become most dangerous among the 7 cities studied.

3.3.2 Toronto and Amsterdam

Amsterdam is an interesting city for comparison to Toronto because it is widely touted as a model transit city balancing the requirements of automobiles, pedestrians, and cyclists. Importantly, it is often claimed that the Amsterdam model delivers increased safety. If these safety claims are true, Amsterdam would provide a baseline fatality rate representing what might be obtained through transit policy alone.

To investigate, Toronto and Amsterdam are plotted in Figure 4. Based on our seven-year observation, only Copenhagen is safer than Amsterdam for pedestrians and cyclists (see Table 1). At least initially, Toronto does not appear to be significantly more dangerous than Amsterdam[3]. However, during the observation period, Toronto rapidly pulls away from Amsterdam to become significantly more dangerous to pedestrians and cyclists.

4 Discussion

This analysis suggests that in 2009 Toronto was among the safer cities in the study. However, in 2015 Toronto had the highest rate of pedestrian and cyclist fatalities in the study, and is trending to become the most dangerous of these cities within the next four years. This paints a rather grim picture for the future of

---

[3]although this is partly a statistical artifact due in part to the unreliability of estimates at the time domain limits
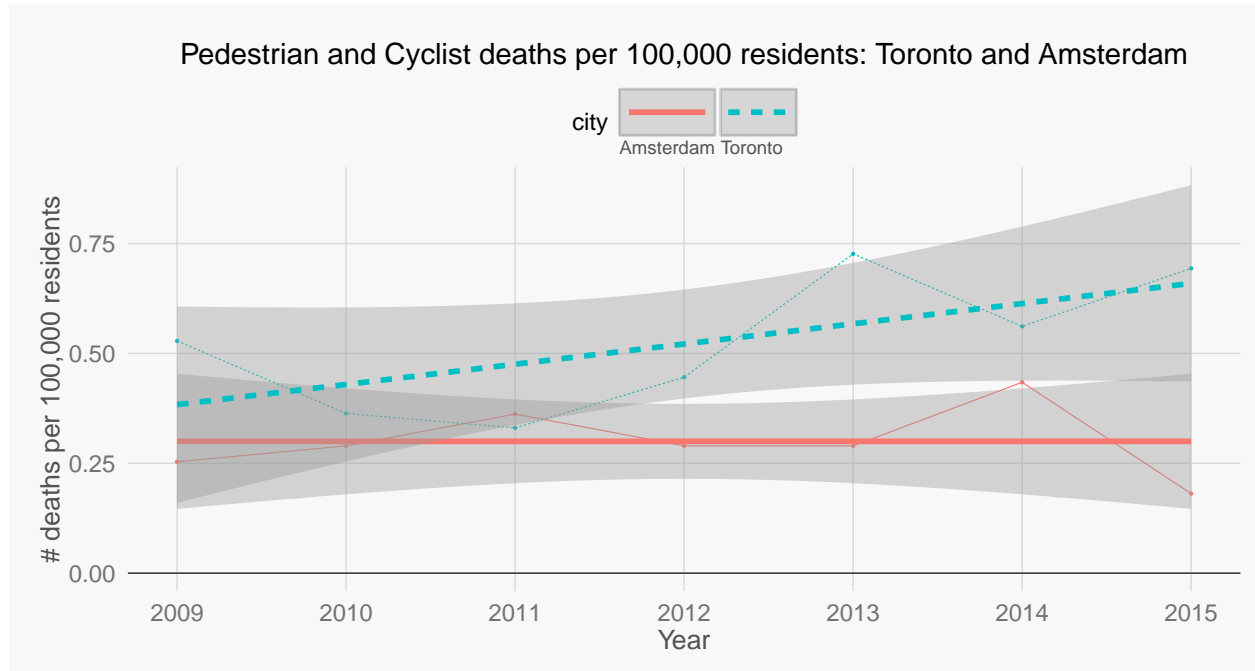
Figure 4: Toronto and Amsterdam pedestrian and cyclist fatality rates are visualized, along with 95-percent confidence intervals. In 2011, the rate in Toronto was actually lower than Amsterdam's, which illustrates the difference between a single measurement versus a trend.

Toronto: the trend is relatively steep in relation to the other cities and it is trending in the direction of greater danger. Toronto is somehow different from all the other cities in terms of this particular trend.

4.1 Limitations

The greatest limitation of this study is the underlying data, which could include more of the historical record for more cities. Although some cities provide archives stretching back to the 1990s, other cities seem to have no official record of pedestrian and cyclist fatalities whatsoever, at least not that the authors could find. Another major limitation to this study is the inability to explain *why* these changes have occurred. It is likely that public policy, road design, safety initiatives, and a variety of other factors influence safety differently in each of these cities, and without a more robust investigation into these factors, it is not possible to comment upon the causal forces that perturb the safety rates up and down. Alas, time is limited as well, and if there were more of it, the surely we could *solve it all*, if only ...

4.2 Future Directions

This work is open source; the data and the code are available for anybody to view and edit. Some immediate future directions are provided in the Todo.md file, which is in the root directory of the project.

Among these next steps is to control for factors that impact traffic: 1) kilometers of roadway; 2) kilometers of bike lanes; 3) number of crosswalks or intersections; 4) kilometers of sidewalk; 5) average speed limit; and 6) number of automobile trips per day. It will also be interesting to analyze the historical record in terms of safety policy interventions, such as "Vision Zero" initiatives. Risk ratios are a common tool used to communicate hazards in terms that are easy to comprehend, so these ought to be included in the final publication.

Table 1: Fitted Model. These slopes and intercepts specify fatality rate trend lines for each city.

	intercept	slope
Amsterdam	0.2971794	0.0040929
Chicago	0.4385113	0.0043063
Copenhagen	0.2571088	0.0033455
London	0.6031261	-0.0053332
Los Angeles	0.4772472	0.0051122
New York City	0.6856391	-0.0039906
Toronto	0.4646306	0.0174628

Another direction for future expansion is to include other cities. Of particular interest to Toronto will be other Canadian cities including Montreal and Vancouver. Lastly, although this report focused on the city of Toronto throughout its analysis and discussion, these data tell the story of every other city too. What are New York City and London doing to lower their pedestrian and cyclist fatality rates? Will the data-driven approach taken by Los Angeles to reform its most dangerous intersections improve safety?

As civic "open data" initiatives continue to spread internationally, these data will become available with greater frequency. Inevitably, a worldwide transit safety assessment shall emerge. Already, we are able to catch a glimpse of this future, and there do seem to be patterns among the cities that have achieved greater safety and those that haven't. Therefore, the ultimate future direction for this work is to inform policy that helps people move quickly and safely.

5 Conclusion

If we calibrate our safety expectations based upon other cities, then this study presents evidence that Toronto compares unfavourably to the other cities that were studied. Something about road safety is different in Toronto. Although it is not possible to conclude *why* Toronto looks so different, there is at least a good reason to begin this investigation in earnest. In conclusion, there is now evidence that the pedestrian and cyclist fatality rate is actually high in the City of Toronto.

6 Appendix

6.1 Trend lines

Slopes and Intercepts corresponding to the fitted trend model are presented in Table 1. Fatality rates are presented in units per 100,000 residents. The lines specified by these equations were then used to extrapolate historical trends into the near future.

6.2 License

References

CBC News. 2015. "Toronto Traffic Fatalities at Highest Level in 5 Years." *CBC News*, december. http://www.cbc.ca/news/canada/toronto/toronto-traffic-fatalities-1.3386126.

Gillis, Wendy. 2016. "Homicides down in 2015, Amid Rise in Fatal Stabbings." *Toronto Star*, January. https://www.thestar.com/news/crime/2016/01/01/homicides-down-in-2015-amid-rise-in-fatal-stabbings.html.

Hain, Charles, David Kister, David Crawford, Mike Sullivan, Rob Cowan, Scott Robinson, Richard Cassel, et al. 2016. "Making Toronto Streets Safer." *Toronto Star*, june. https://www.thestar.com/opinion/letters_to_the_editors/2016/06/18/making-toronto-streets-safer.html.

Pinheiro, José, Douglas Bates, and R-core. 2016. *Nlme: Linear and Nonlinear Mixed Effects Models*. https://CRAN.R-project.org/package=nlme.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Spurr, Ben. 2016. "More Than 1,000 Cyclists and Pedestrians Hit on Toronto Streets Since June 1." *Toronto Star*, sept. https://www.thestar.com/news/gta/transportation/2016/09/23/more-than-1000-cyclists-and-pedestrians-hit-on-toronto-streets-since-june-1.html.

Wickham, Hadley, and Winston Chang. 2016. *Ggplot2: An Implementation of the Grammar of Graphics*. https://CRAN.R-project.org/package=ggplot2.