Pedestrian Risk decrease with Pedestrian Flow and Increase with Vehicle Flow - A Case Study

Lars Leden

Adjunct Prof. Luleå University of Technology, SE-971 87 Luleå, Sweden
E-mail: Lars.Leden@ sb.luth.se or Lars.Leden@vtt.fi
September 8, 1997

Abstract

We used a unique database providing pedestrian accidents and estimates of pedestrian and vehicle flows for the years 1983 - 1986 for approximately 300 signalized intersections in Hamilton, Ontario, Canada. Pedestrian safety at semi-protected schemes, where left-turning vehicles face no opposing traffic but have potential conflicts with pedestrians, were compared with pedestrian safety at normal non-channelized signalized approaches, where right-turning vehicles have potential conflicts with pedestrians.

Four different ways of estimating hourly flows for left- and right-turning vehicles were explored. Parameter estimates were affected by the time period used for flow estimation. However, parameter estimates seem to be much more affected by the traffic pattern (left or right-turning traffic) even though approaches were selected so that the situation for left and right-turning turning traffic should be similar (no opposing traffic, no advanced green or other separate phases and no channalization). Left-turning vehicles causes higher risks for pedestrians than right-turning vehicles, probably due to higher speeds caused by larger radius for left-turning vehicles. At low vehicular flows right turns and semi-protected turns tend to be equal safe for pedestrians.

If risks for pedestrians are calculated as expected number of reported pedestrian accidents per pedestrian, risks decrease with increasing pedestrian flows and increase with increased vehicle flow.

1. BACKGROUND AND PURPOSE

My work focused on safety for pedestrians with conflicting left and right-turning turning vehicles and how the models are influenced by the choice of different time period for estimating pedestrians and vehicle flows. The conceptual framework in this context and the method used is explained in detail elsewhere (Leden, 1993). Data was also used to analyze how to typical schemes for accommodating left turning vehicles influence the safety for pedestrians (Quaye *et al*, 1993).

2. THE DATA

The data gathering part in my project was started by Almuina (1989). His work focused on pedestrian accidents and left-turning vehicles at signalized intersections. As a part of his work he prepared an Accident Database and an Intersection Geometry Database for the approximately 300 signalized intersections in the Regional Municipality of Hamilton-Wentworth (Region) located within south central Ontario. The majority of the intersections are located in the City of Hamilton in a typical American grid network with many one way streets.

To enhance the chances of success to find a "pure" relation between pedestrian accidents and pedestrian and vehicle flows I selected for analysis a set of signalized approaches which are similar in most respects except traffic flows and accident history. For accidents involving right-turning vehicles I gathered data for approaches which were not channelized, *i.e.* there was no extra island to exclude right-turning turning traffic from signal control and let them yield towards pedestrians. For accidents involving *left-turning* vehicles I used the following criteria to select approaches:

- no opposing traffic i.e. semi-protected scheme and
- no advanced green for left-turning traffic or other separate phase for left-turning traffic.

No opposing traffic in the approach could be due to the missing leg in a three-way intersection or due to one-way traffic in the opposite approach in the direction away from the approach. To decide which approaches fulfilled the criteria I used the information provided in the Intersection Geometry Database and in micro-fiches of the lay-out of the intersections. I got information about advanced green or other separate phase for left-turning traffic by interviewing the engineer in charge in the city of Hamilton (Mr. Hart Solomon).

The city of Hamilton provided us with stream counts of vehicles and pedestrian for fifteen minutes periods for 1983 to 1986. Typically there was one or two counts per year at each intersection. Counts were conducted from 7 to 10 a.m. and from 14 to 18 p.m. Monday to Friday and reported in fifteen minutes periods. From these counts hourly flows were estimated for *each* fifteen minutes period for *all* Mondays to Fridays during the study period. These estimations were used as basis for estimating average hourly flows in four different ways, see chapter 4. As available accident data was for 1977 to 1986 estimates of flows had to be extrapolated for the years 1977 to 1982 to correspond with the accident data. However these extrapolated estimates did not seem to be reliable, so this paper is restricted to describe the analysis of data from 1983 to 1986. The method to estimate flows is described by Quaye, Leden & Hauer (1993).

To correspond to the available traffic count information, I used only those accidents that occurred between 7 to 10 a.m. and between 14 to 18 p.m. Monday to Friday and which involved left-turning or right-turning vehicles and a pedestrian. 63 accidents from 1977 to 1982 and 66 accidents from 1983 to 1986 remained for the analysis. As already mentioned the analysis were restricted to describe the analysis of data from the years *1983 to 1986*. So I have a total of 66 accidents, 27 of them between left-turning vehicles and pedestrians and 39 between right-turning turning vehicles and pedestrians.

3 METHOD

The cause-and-effect relationship between accidents and traffic flows is of course crucial. For example does the straight-through flow not affect the safety for pedestrians at signalized intersections much and only those walking against red light. So the accidents were related to the flows to which the colliding vehicle and pedestrian belong, i.e. to specific accident patterns.

On the basis of an exploratory analysis, one can suggest functional forms for expressions that fit what has been observed. As I had only a small number of accidents my conclusion could not be very firm. Experience gained from previous work (Hauer et al, 1988) suggested that it could be reasonable to use the following form of the model:

$$x_i = b_0 F_1^{b_1} F_2^{b_2} + e_i = \hat{E}\{m\} + e_i \qquad (1)$$

where for each case i

x_i is the observed number of accidents per unit of time

$\hat{E}\{m\}$ is the estimated number of accidents per unit of time,

F_1 is the vehicle flow per hour (right-turning or left-turning)

F_2 is the pedestrian flow per hour and

e_i is the "error" variable, the residual.

My exploratory analysis supported this form.

The usual approach to the analysis is by multiple regression. To estimate parameters we have to transform the function in equation 1 by taking logarithmic values. The model, expressed in a traditional form can then be written:

$$y_i = \ln x_i = b_0 + b_1 \ln F_1 + b_2 \ln F_2 + e_i = m_i + e_i \qquad (2)$$

I estimated coefficients using the Generalized Linear Interactive Modeling (GLIM) software package (Aitkin *et al*, 1986). In this package it is possible to choose an appropriate error distribution. To be able to do the right choice it is necessary to understand the conceptual framework, which will be discussed below.

I have denoted m_i the safety of a specific intersection. Imagine a population of intersections that all have the *same traffic flows*. In this imaginary population, the m_i's would still vary from intersection to intersection because, although flows are identical, they involve different drivers in different cities, and so forth. Thus, one can speak of the expected value or mean of the m's ($E\{m\}$) in this imaginary population of intersections with *identical* traffic flows. This mean of m's is what describes the safety of a representative or an average intersection for this imaginary population of intersection with a specific traffic flow. Similarly, one can speak of the variance of the m's.

When fitting a model to accident data, I am trying to estimate $E\{m\}$ as a function of traffic flow (in this case). That is, I am trying to determine what the m is of some average or representative intersection and how it varies with traffic flow. However, the data used for estimation are not for average intersections. Each accident count I use is for one specific intersection from the imaginary population of intersections with the same flows. It follows that if $E\{m\}$ is what I wish to estimate, the accident count must be considered as a Poisson random variable that comes from a site with $E\{m_i\}$ as its expected value and that this m, in turn, is one of a distribution of m's characterized by $E\{m\}$ and $Var\{m\}$.

Thus, the distribution of accident counts around the $E\{m\}$ is one family of "compound Poisson distributions." In the special case in which the distribution of m's in these imaginary populations can be described by a gamma probability density function, the distribution of accident counts around the $E\{m\}$ must be taken as negative binomial, or with other words the error distribution e_i is *negative binomial*.

The variance of accident counts s^2 is given by $Var\{m\} + E\{m\}$ or

$$\hat{Var}\{m\} = s^2 - \bar{x} \tag{3}$$

Note that the relationships are not affected of the transformation into a logarithmic scale according to equation 2. In principle these relationships can be used to estimate $Var\{m\}$ for different subsets of the data with almost the same $\hat{E}\{m\}$. I have not enough data for this. I have to use results from work already done. It is described by Hauer et al (1991). As $\hat{Var}\{y_i\} = \Sigma e^2/n$ clues to $\hat{Var}\{m\}$ are contained in the residuals e_i. Hauer and Persaud (1987) found that there often is a relationship between $E\{m\}$ and $Var\{m\}$ and that it can usually be adequately represented by:

$$Var\{m\} = (E\{m\})^2/k. \tag{4}$$

where k is the first parameter for the negative binomial distribution.

This means that the same relationship is valid for subsets of data as for data for the whole Gamma distribution. From equation 4 we get for the whole distribution:

$$Var\{m\} = k/\lambda^2 = k^2/(\lambda^2 k) = (E\{m\})^2/k.$$

Hauer *et al* (1991) confirmed the validity of this empirical finding for many groups of their data base by using equation 3.

There are two different methods, which can be used to estimate k, the method of moments and the maximum likelihood method. I chose to use the latter one. However some examples calculated by both methods indicated that the two methods gave similar results.

Maycock and Maher (1988) suggests the method of moments to estimate k. As in equation 2, e_i is the residual (y_i - m_i), then

$$E\{e^2\} = m_i + m_i^2/k$$

and an estimate of k is given by:

$$k = \sum \hat{m}_i^2 / \sum (e^2 - \hat{m}_i) . \qquad (5)$$

Hauer *et al* (1991) describe the maximum likelihood method of estimating *k*. The iterative process of estimating $\hat{Var}\{m\}$ begins by estimating provisional model parameters on the assumption that $\hat{Var}\{m\} = 0$ or using some other starting guess. Once I have provisional parameter estimates I can find an estimate of *k* that maximizes the likelihood (L) of the data as follows. As showed above the accident counts can be assumed to be negative binomial distributed; the parameter a for the negative binomial model can be expressed as a function of k and $E\{m_i\}$ by using that $E\{m_i\}=b/a$ and $Var\{m_i\}=b/a^2= (E\{m_i\})^2/k$. So the probability of an accident count x_i for case number i can be written as:.

$$p(x_i) = [a/(a+1)]^k [k(k+1) \ldots (k+x_i-1)/x_i! [(a+1)^{x_i}] =$$

$$[k/E\{m_i\}]^k [k(k+1) \ldots (k+x_i-1)]/[x_i! (k/E\{m_i\}+1)^{x_i+k}].$$

The likelihood function *L* describes the probability of having the actual outcome of accident counts $x_1, \ldots, x_i, \ldots x_n$. If events are independent this probability can be calculated as Π p(xi). To facilitate calculations ln L is calculated instead of L, so,

$$\ln L = \ln \Pi\, p(x_i) = k [\Sigma \ln(k/E\{m_i\})] + \Sigma[\ln(k) + \ln(k+1) + \ldots + \ln(k+x_i-1)] - \Sigma(x_i+k) \ln(1+k/E\{m_i\}) + \text{constant}^1. \qquad (6)$$

Since we have estimates for $E\{m_i\}$, it is easy to find that value of k which maximizes ln*L* (and *L*) by calculating ln*L* for different values of k. This k is then used to calculate $\hat{Var}\{m\}$ by equation 4. The provisional error structure is revised, and model parameters are estimated anew. The process converges in two or three iterations.

As I had a very small database the estimates of k in my case were very uncertain and in some cases it was not even possible to calculate the estimate of k which maximizes the likelihood. However, the estimates of the model parameters were not very sensible for changes in k.

---

[1] not dependent on k

4 RESULTS

My work focused on how the models describing pedestrian safety are influenced by the choice of a different time period for estimating the hourly flows. I explored four different ways of estimating hourly flows:

1 average daily flow for each year, *daily model*

2 average hourly flows during a.m. and p.m. periods for each year, *i.e.* 7 to 10 a.m. and 14 to 18 p.m. for 1983, 1984 etc., *a.m./p.m. model*

3 average hourly flows for each hour and year, *i.e.* 7 to 8:00, 1983, 8:00 to 9:00, 1983 etc. *hourly model* and

4 average hourly flows for each fifteen minutes period and year, *i.e.* 7 - 7:15, 1983, 7:15 - 7:30, 1983 etc. *15 minute model.*

In all models estimated, the traffic flows are expressed in vehicles or pedestrians per hour. The dependent variable in each model was based on the number of police reported accidents occurring in the corresponding time periods (*e.g.* for the daily model the number of police reported accidents per day during study hours 7 to 10 a.m. and 14 to 18 p.m. was used etc. The parameter estimates obtained, after fitting the 8 models using the Generalized Linear Interactive Modeling (GLIM) software package are given in equation 7 and Tables 1 and 2.

$$\hat{E}\{m\} = b_0 F_1^{b_1} F_2^{b_2} \qquad (7)$$

m is the expected number of accidents per unit of time of a certain intersection with an hourly right or left-turning vehicular flow F_1 and hourly pedestrian flow F_2,

$\hat{E}\{m\}$ is the estimated number of accidents per unit of time of a certain intersection

b_0, b_1 and b_b are parameters to be estimated.

Table 1. Parameter estimates for *left-turning* vehicles.

Flow period	$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{k}$
1. day	2.62 10^{-7}	1,19	0,331	2,2
2. a.m./p.m.	4.85 10^{-8}	1,37	0,346	*
3. hour	1.82 10^{-8}	1,32	0,338	0,4
4. 15 minutes	3.61 10^{-9}	1,35	0,368	*

* not enough data for estimating k

Table 2. Parameter estimates for *right-turning* turning vehicles.

Flow period	$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{k}$
1. day	4.19 10^{-7}	0.864	0.475	*
3. a.m./p.m.	1.19 10^{-7}	0.919	0.570	*
4. hour	4.08 10^{-8}	0.913	0.514	*
5. 15 min.	2.43 10^{-8}	0.864	0.321	*

* not enough data for estimating k

It should be noted that the estimate from the daily model pertains to information aggregated over *2* a.m. or p.m. periods, *7* hours of the day (specifically: 7 to 10 a.m. and 2 to 6 p.m.), or *28* fifteen minutes periods. Ideally, one would expect that multiplying the 15 minute estimate of E{m} by 28, the hourly estimate by 7 or the a.m./p.m. estimate by 2 should yield the daily estimate. Figures 3 and 4 illustrate this. In these figures, the curves labeled 1 gives the daily estimates of E{m}, based on the daily model, for various values of left- or right-turning vehicular flow F_1, when pedestrian flow F_2 is 50 pedestrians per hour in Figure 3 and 500 pedestrians per hour in Figure 4. Curves 2, 3 and 4 are daily estimates obtained from the a.m./p.m., hourly and 15-minute models respectively, for the same traffic flow combinations.

It is evident from Figure 1 that estimates from the 4 different models gives similar estimates. Left-turning vehicles causes higher risks for pedestrians than right-turning vehicles, probably due to higher speeds and larger radius for left-turning vehicles. If risks for pedestrians are calculated as expected number of reported pedestrian accidents per pedestrian, *i.e.* equation 9 is divided by 7 F_2 (daily pedestrian flow), risks are decreasing with increasing pedestrian flow. Figure 2 shows estimates for F_1 = 50 vehicles per hour and Figure3 for 500 vehicles per hour. For small vehicle flows risk differences vanish between left- and right-turning models.

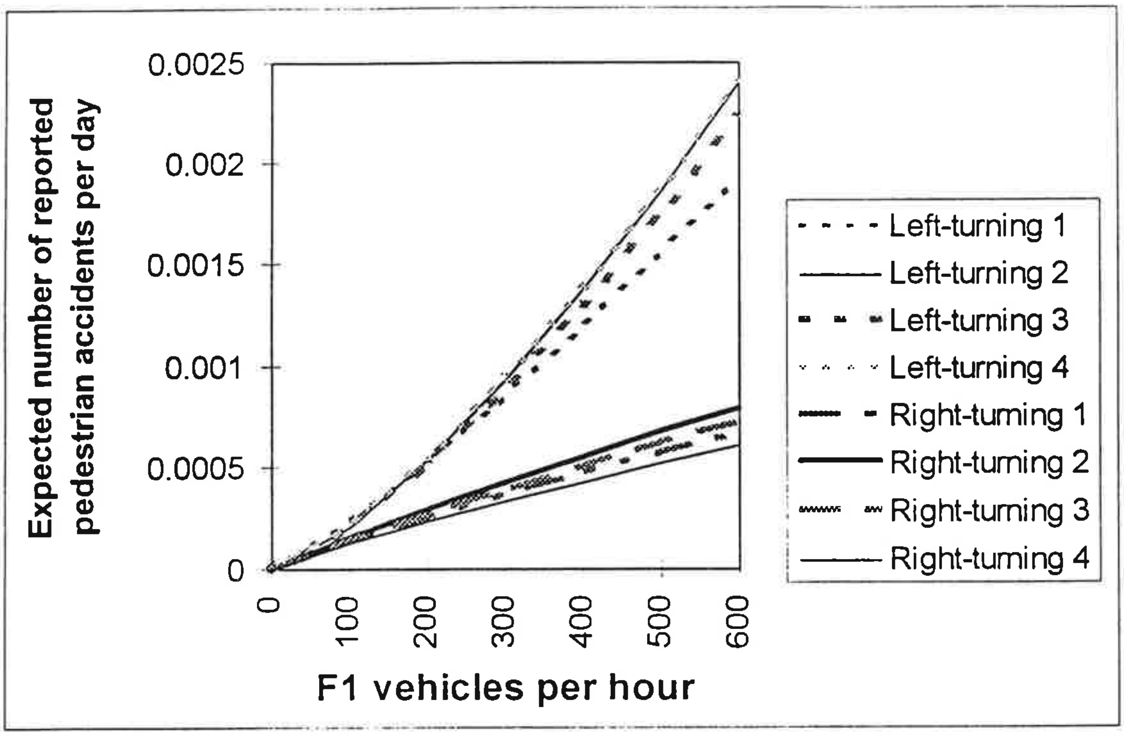

Figure 1. Estimates of E{m} for F_2 = 50 pedestrian per hour from the daily (1), a.m./p.m. (2), hourly (3) and 15-minute (4) models for left and right-turning turning vehicles

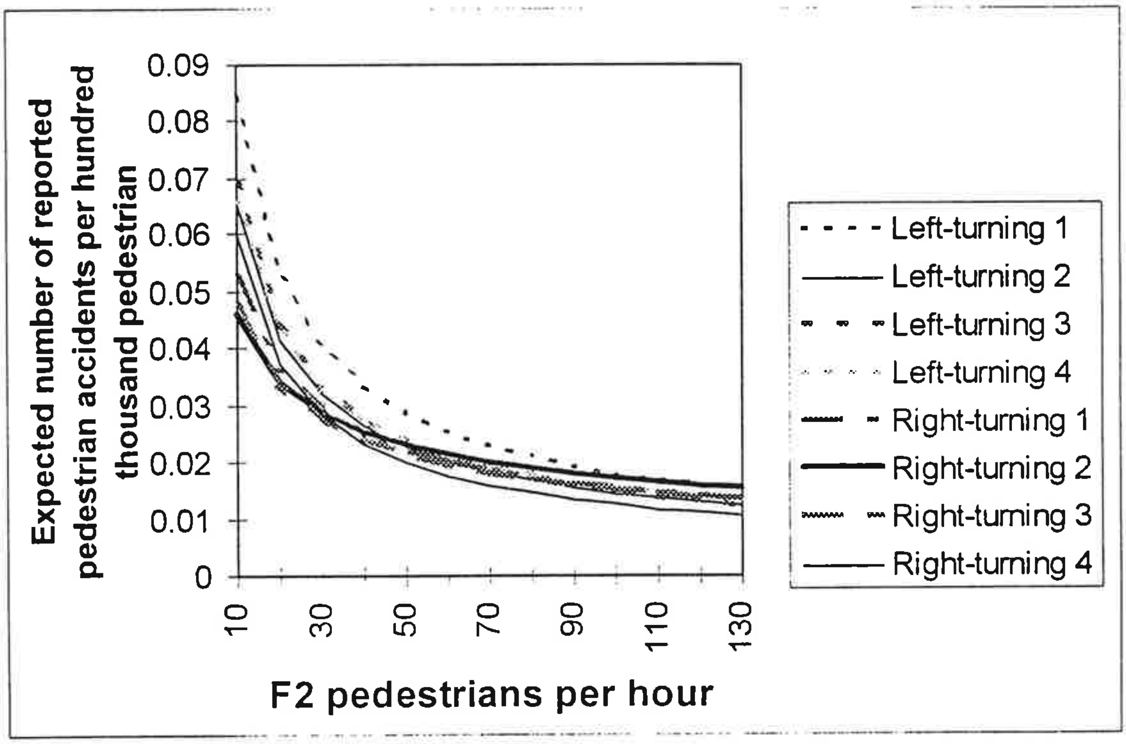

Figure2. Estimates of Expected number of reported pedestrian accidents per day for F_1 = *50 vehicles per hour* from daily (1), a.m./p.m. (2), hourly (3) and 15-minute (4) models for left and right-turning turning vehicles

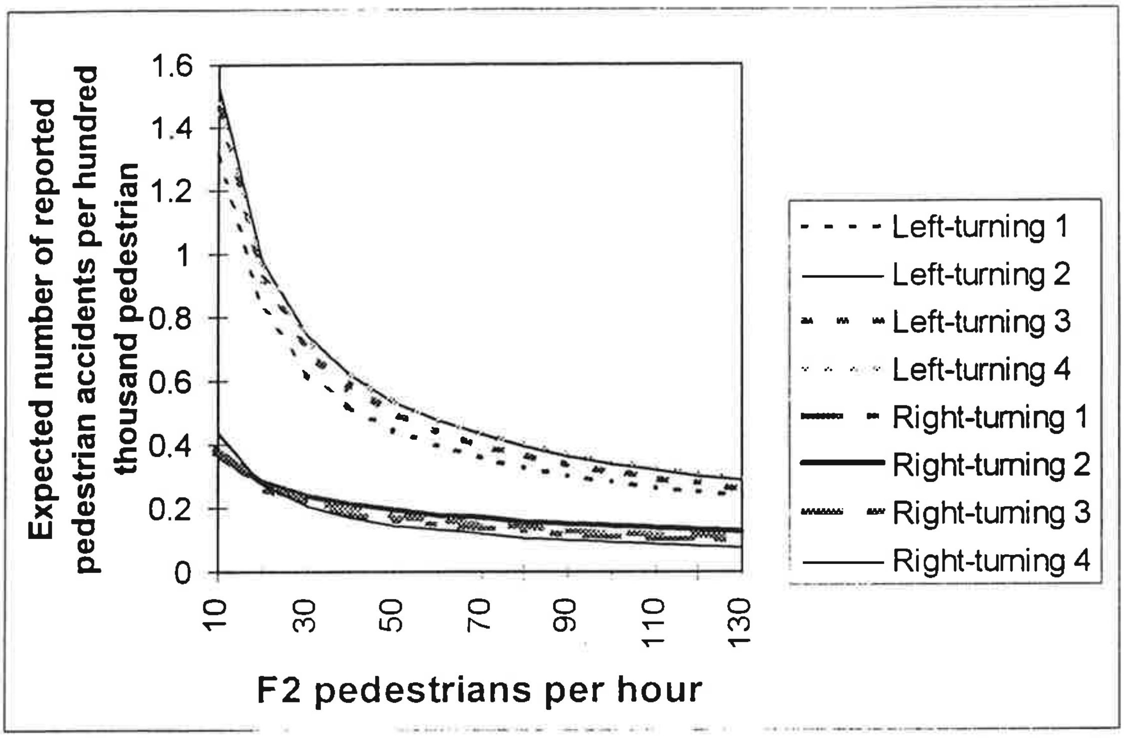

Figure 3. Estimates of Expected number of reported pedestrian accidents per pedestrian for F_1 = *500 vehicles per hour* from the daily (1), a.m./p.m. (2), hourly (3) and 15-minute (4) models for left and right-turning turning vehicles

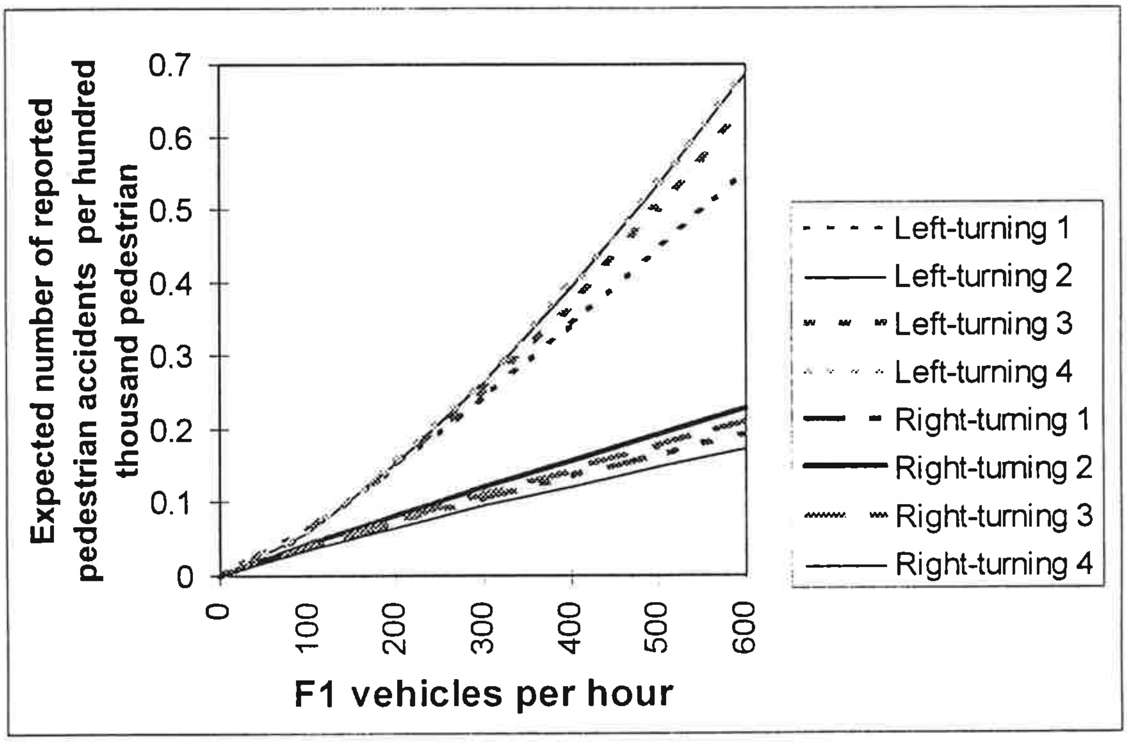

Figure 4. Estimates of Expected number of reported pedestrian accidents per pedestrian for F_2 = *50 pedestrian per hour* from the daily (1), a.m./p.m. (2), hourly (3) and 15-minute (4) models for left and right-turning turning vehicles

If risks for pedestrians are calculated as expected number of reported pedestrian accidents per pedestrian, risks are increasing when vehicle flow increases, see Figure 4.

SUMMARY AND DISCUSSION

We used a unique database providing pedestrian accidents and estimates of pedestrian and vehicle flows for the years 1983 - 1986 for approximately 300 signalized intersections in Hamilton, Ontario, Canada. Pedestrian safety at semi-protected schemes, where left-turning vehicles face no opposing traffic but have potential conflicts with pedestrians, were compared with pedestrian safety at normal non-channelized signalized approaches, where right-turning vehicles have potential conflicts with pedestrians.

Quaye, Leden & Hauer (1993) examine how two typical schemes for accommodating left-turning vehicles influences the safety for pedestrian. These are the semi-protected scheme described above and a permissive scheme, in which left-turning vehicles have to find suitable gaps in the opposing traffic. Semi-protected left turns tend to be safer for pedestrians at *low* vehicular flows. The opposite is true for high flows of left turning vehicles.
I explored four different ways of estimating hourly flows for left- and right-turning vehicles and fitted daily, a.m./p.m. hourly and 15 minute models to the data. Parameter estimates were affected by the time period used for flow estimation. However, parameter estimates seem to be much more affected by the traffic pattern (left- or right-turning traffic) even though approaches were selected so that the situation for left- and right-turning traffic should be similar (no opposing traffic, no advanced green or other separate phases and no channalization). Left-turning vehicles causes higher risks for pedestrians than right-turning vehicles, probably due to higher speeds caused by larger radius for left-turning vehicles. At *low* vehicular flows right turns and semi-protected turns tend to be equal safe for pedestrians, but right turns are safer for pedestrians than *permissive* left turns according to results quoted above.

If risks for pedestrians are calculated as expected number of reported pedestrian accidents per pedestrian, risks decrease with increasing *pedestrian* flows. Ekman (1996) has found for 95 non-signalized intersections in Malmö and Lund that the rate of pedestrian conflicts per pedestrian is not influenced by the pedestrian flow. According to Ekman this could be interpreted as follows: The individual pedestrian does not seem to benefit from presence of other pedestrians. Another interpretation is that the vehicle drivers do expect pedestrians, at least if pedestrian flow exceeds 30 pedestrians per hour. So in my case an explanation could be that vehicle drivers' alertness increases with increasing pedestrian flow.

If risks for pedestrians are calculated as expected number of reported pedestrian accidents per pedestrian, risks increase with increased *vehicle* flow. Ekman (1996) have found similar results for non-signalized intersections.

Flow fluctuates between and within each cycle of a traffic signal system. Flows are systematically higher at the start of the green period. However the time of the accident related to the start of the green period were not known, therefore it could not be studied how this influence the safety of pedestrian.

Acknowledgments

I started working on this study during my sabbatical year from August 1, 1991 to August 15, 1992 at University of Toronto. I had the pleasure being a temporary staff working for professor Ezra Hauer together with dr Kwei Quaye. I would like to thank Ezra especially for all brilliant clues and comments, which helped me to develop my own approach to analytical thinking and teaching.

It was a great experience for me and I would like to thank everybody who supported me and especially: The Technical Research Centre of Finland, VTT (Prof. Esko Kankare and Prof. Markku Salusjärvi), Swedish Road Administration (Mr. Thomas Lekander and Mr. Hans Rydgren), Åke and Greta Lissheds Foundation and Henry Ford Foundation. At last, but not at least I would like to thank dr. Risto Kulmala, VTT, for fruitful comments to a draft of this report.

I finished the study as Adjunct Prof. at Luleå University of Technology and would like to thank my colleagues there for a good start.

I would finally like to thank the members of the International Cooperation on Theories and Concepts in Traffic Safety (ICTCT) in general and in particular the organizer of the meeting in Lund, Mrs. Åse Svensson.

REFERENCES

Aitkin, M., Baker, R.J., Clarke, M.R.B., Francis, B., Gilchrist, R., Green, M., Nelder, J.A., Reese, R.A., Payne, C.D., Slater, M., Swan, A:V., White, R., 1986. the Generalised Linear Interactive Modeling system: Release 3.77, Manual. Oxford.

Almuina, A.L., 1989. Pedestrian Accidents and left-turning vehicles at signalized intersections. Toronto, University of Toronto, Department of Civil Engineering. 116 p + 8 App.

Ekman, L., 1996. On the Treatment of Flow in Traffic Safety Analysis a non-parametric approach applied on vulnerable road users. University of Lund, Department of Traffic Planning and Engineering. Bulletin 136.

Hauer, E., C.N., J. & Lovell, J., 1988. Estimation of Safety at Signalized Intersections. Washington, D.C., Transportation Research Board, Transportation Research Record 1185, p. 48 - 61.

Hauer, E., Persaud, B.N., Smiley, A. & Duncan, D., 1991. Estimating the Accident Potential of an Ontario Driver. Accid. Anal. & Prev. Vol. 23, pp. 133 - 152.

Leden, L., 1993. Estimation of safety . Some theoretical considerations applied to pedestrians´safety at signalized intersections.

Maher, M.,J., 1989. Fitting Predictive Equations to Accident Data when there is uncertainty in the Traffic Flow Estimates. London, PTRC, 17th Summer Annual Meeting, University of Sussex 11 - 15 September 1989. Traffic Management and Road Safety. P. 171 - 181.

Maycock, G. & Maher, M., J., 1988. Generalized Linear Models in the Analysis of Road Accidents - Some Methodological Issues. Crowthorne, Transport and Road Research Laboratory. Personal Information. Working Paper. 8 p.

Quaye, K., Leden, L. & Hauer, E., 1993. Pedestrian Accident and Left-Turning Traffic at Signalized Intersections. Toronto, University of Toronto. Safety Studies Group. 18 p.