

Pedestrian fatality risk as a function of car impact speed

Published article available at <http://dx.doi.org/10.1016/j.aap.2009.02.002>

Erik Rosén^{*}, Ulrich Sander

Autoliv Research, Wallentinsvägen 22, 447 83, Vårgårda, Sweden.

^{*}Corresponding author: Tel.: +46 322 626708, fax: +46 322 620118.

Email address: erik.rosen@autoliv.com

Abstract: Knowledge of the amount of violence tolerated by the human body is essential when developing and implementing pedestrian safety strategies. When estimating the potential benefits of new countermeasures, the pedestrian fatality risk as a function of impact speed is of particular importance. Although this function has been analysed previously, we state that a proper understanding does not exist. Based on the largest in-depth, pedestrian accident study undertaken to date, we derive an improved risk function for adult pedestrians hit by the front of passenger cars. Our results show far lower fatality risks than generally reported in the traffic safety literature. This discrepancy is primarily explained by sample bias towards severe injury accidents in earlier studies. Nevertheless, a strong dependence on impact speed is found, with the fatality risk at 50 km/h being more than twice as high as the risk at 40 km/h and more than five times higher than the risk at 30 km/h. Our findings should have important implications for the development of pedestrian accident countermeasures worldwide. In particular, the scope of future pedestrian safety policies and research should be broadened to include accidents with impact speeds exceeding 50 km/h.

Keywords: Pedestrian; fatality risk; logistic regression; impact speed

1. Introduction

Road traffic accidents are a global health problem claiming approximately 1.2 million fatalities per annum (WHO, 2004). The largest group of road user fatalities are pedestrians hit by motorised vehicles (Mohan, 2002; Otero et al., 1997; WHO, 2004), which will increase further with the motorisation of countries such as China and India (Kopits & Cropper, 2005; WHO, 2004). In the western world, typically 10 to 30% of all road accident fatalities are pedestrians (IRTAD, 2008; WHO, 2004). In many other countries, these proportions are substantially higher, although the exact figures are often difficult to assess (IRTAD, 2008; Mohan, 2002; Otero et al., 1997; WHO, 2004). Thus, there is a compelling need for worldwide implementation of effective pedestrian injury mitigation and crash avoidance countermeasures.

The concept of risk can be interpreted according to scientific context. In traffic safety literature, it is common to define the pedestrian fatality risk as the probability of death, given that the pedestrian was hit by a motorised vehicle and also injured. This is because very little data exist on crashes involving only uninjured pedestrians. We also note that pedestrian fatalities generally include only deaths occurring within 30 days as a result of a motor vehicle crash.

Within certain groups of the traffic safety community, there is presently a perceived consensus that the risk of pedestrian death is a well-known function of car impact speed. Typically, the fatality risk has been reported at 40 to 90% at an impact speed of 50 km/h

(ERSO, 2008; GRSP, 2008; OECD/ECMT, 2006; WHO, 2004). However, taking a deeper look into the traffic safety literature, we found that only a limited set of research articles, studying the risk of death as a function of impact speed, have been published during the past 30 years. Furthermore, most of the real-world data samples studied were either very small, substantially biased towards severe injury crashes, or more than 30 years old.

The scope of this study was to derive an improved function for adult pedestrian fatality risk based on real-world accident data. In addition to car impact speed, the effects of pedestrian age, height, weight, and gender were to be investigated. The resulting risk curves should offer assistance and guidance for future pedestrian safety strategies. In particular, they should be useful for benefit and effectiveness studies of proposed countermeasures. Child pedestrian mortality should be treated in a separate study due to the anatomical and biomechanical differences between children and adults (Tarri re, 1995). Therefore, only pedestrians aged 15 years or older were considered.

2. Literature review

Traffic safety reports that include pedestrian fatality risk curves most often cite the work of Anderson et al. (1995, 1997), Ashton (1982), Pasanen (1992)¹, Teichgr ber (1983), or Walz et al. (1983). For example, the reports mentioned in the introduction were all based on one or several of these articles. Rather surprisingly, we found that neither Ashton, nor Teichgr ber, or Walz et al. derived any risk curves in their articles. It is true that Ashton (1982) included fatality rates at different speed ranges from pedestrian accident investigations in Great Britain during the 1960s and 70s. However, Ashton et al. have specifically pointed out that, due to sample bias, these fatality rates did not give a fair description of the total population of accidents (Ashton et al., 1977; Ashton, 1982). This statement has been largely overlooked in subsequent studies using these data. Teichgr ber merely included a risk curve, which could be traced back to a report by Yaksich (1964)². Furthermore, we found that when Walz et al. (1983) were cited, people actually included the risk curve derived by Anderson et al. (1995, 1997), who applied an unconventional approach to deriving fatality risks. Furthermore, the analysis of Anderson et al. was based on only 56 accidents with a substantial bias towards severe injuries (IWGAM, 1986), thus yielding crude and exaggerated risk estimates. Finally, Pasanen (1992) applied regression analysis to the biased data presented by Ashton (1982), which inevitably rendered his risk estimates too high (Davis, 2001).

It was pointed out by Davis (2001) that the probable cause for the bias in the data presented by Ashton (1982) was the use of an outcome-based sampling scheme, meaning that the probability of an accident to be included in the sample depended on the injury severity of the accident. When outcome-based sampling is present, the data cannot be directly used to estimate fatality risks. This is easily realised by considering the following hypothetical example: Assume that 100 accidents have occurred at a certain impact speed and that 10 of these were fatal and 90 were non-fatal. The fatality rate, or empirical fatality risk, would then be 10%. However, if only a subset of these accidents were investigated, e.g., all fatal and every third non-fatal accident (outcome-based sampling), the sample would comprise 10 fatal and 30 non-fatal accidents. Hence, direct use of this sample would give an estimated fatality risk of 25%, which is 2.5 times higher than the actual fatality rate. Notice, that if one knew the sampling rates for both fatal and

¹ Sometimes Pasanen's PhD thesis from 1991 is cited instead. This thesis contained the same risk curve as the article by Pasanen (1992), but was written in Finnish.

² The report by Yaksich (1964) was based on police reported accidents in St. Petersburg, Florida, during 1958 to 1963. Furthermore, the victims included a high proportion of elderly pedestrians.

non-fatal accidents, the risk estimate could be adjusted by giving each non-fatal accident a weight factor equal to three (the inverse of the sampling rate).

Some other analyses on this topic have also been reported (Cuerden et al., 2007; Davis, 2001; Hannawald & Kauer, 2004; Oh et al., 2008a, 2008b), but these have not had the same impact on the traffic safety community as the articles mentioned above. Cuerden et al. (2007) reported fatality rates from the British crash study On-The-Spot. However, the data only included seven fatalities (and 101 survivors), thus making the results very approximate. Davis (2001) made use of the data presented by Ashton (1982) but tried to adjust for the bias by a weighting scheme, which was derived by comparing the distribution of pedestrian injury severities to national statistics. The results showed substantially lower fatality risks than those presented by Ashton (1982) and Pasanen (1992). However, these data are by now more than 30 years old and both car design and medical care have changed during these years. The analysis of Hannawald & Kauer (2004) concerned the risk of sustaining slight and severe injuries (maximum AIS2+ and AIS5+ respectively) and involved car-to-pedestrian collisions in which the forces through the car were coded as being more or less parallel to its longitudinal direction. (For information on the Abbreviated Injury Scale, AIS, see AAAM, 2001.) The analysis was based on a large, but unweighted, data set from the Medical University of Hanover and the German In-Depth Accident Study from the years 1991 to 2003. Their report did not describe in any detail how the statistical analysis was conducted, but it was claimed that the risk of sustaining a maximum AIS5+ injury was similar to the fatality risk. The analyses of Oh et al. (2008a, 2008b) were based on Korean data from 2003 to 2005, which included a substantially higher proportion of fatal accidents than Korean national statistics (Youn et al., 2005). The accident investigations were only briefly described, but it seems likely that the bias was an effect of outcome-based sampling. The results of previous work on this topic are summarised in Table 1. Comparing to the risk curve presented in Fig. 1, we see that the risk estimates of Cuerden et al. (2007), Davis (2001), and Hannawald & Kauer (2004) are in line with the findings of this study. A detailed literature review will be published in the near future.

Table 1: Summary of previous work

	Years of data	30 km/h	50 km/h	70 km/h
Anderson et al. (1997)	1978	8%	85%	100%
Ashton (1982)	1965–1979	≈ 5%	≈ 45%	≈ 95%
Pasanen (1992)	1965–1979	6%	40%	94%
Yaksich (1964)	1958–1963	≈ 22%	≈ 65%	100%
Cuerden et al. (2007)	2000–2007	≈ 2%	≈ 12%	≈ 33%*
Davis (2001)**	1965–1979	1%	7%	51%
Hannawald & Kauer (2004)	1991–2003	4%	14%	39%
Oh et al. (2008b)***	2003–2005	7%	34%	77%

Table showing years of data collection and fatality risks estimated at 30, 50, and 70 km/h in previous publications on the fatality risk for pedestrians struck by passenger cars.

*This estimate was based on private communication with Cuerden and was not given in Cuerden et al. (2007).

**Risk estimates regard pedestrians in the ages 15 to 59 years.

***Striking vehicles included passenger cars, SUV's, vans, trucks, and buses.

3. Data and methods

3.1 Data

We queried the German In-Depth Accident Study (GIDAS) for pedestrian accidents during the years 1999 to 2007. GIDAS contains the largest in-depth pedestrian accident sample ever collected, presently comprising 2127 pedestrians involved in road traffic accidents. The GIDAS teams operate in Dresden, Hanover and surroundings. The work shifts for the teams are specified by a statistically developed sampling plan and cover half the hours of each day and night (Otte et al., 2003; Pfeiffer & Schmidt, 2006). If an accident occurs and it is suspected that at least one person has been injured, GIDAS is contacted within minutes by the local police or fire department. GIDAS investigators then attend the crash scene with blue-lights and sirens. Accident reconstructions, which assess e.g. the impact speed, are later carried out based on on-scene information including collision point, pedestrian and vehicle end-positions, and brake marks, as well as interviews with the driver, pedestrian, and other eye-witnesses. (Further details on accident reconstructions are provided in Appendix A) This sampling procedure is intended to provide data that are representative of the sampling regions. However, if a crash occurs within these regions but police suspect no injuries, GIDAS is not contacted. In these cases, possible pedestrian injuries would later be reported to the police. This phenomenon leads to an over-representation of severe and fatal accidents in GIDAS (Pfeiffer & Schmidt, 2006). We compensated for this by considering German national statistics on pedestrian accidents (Verkehrsunfälle, 2003–2007). These statistics included all police reported injury accidents that occurred in Germany involving at least one motorised vehicle. To weight the GIDAS data, we made explicit use of a variable that coded each pedestrian as “ambulant,” “in-patient,” or “killed.” The definitions of these values were equivalent to the police definitions of “slight,” “severe,” and “fatal” injury. By matching the proportions of “ambulant,” “in-patient,” and “killed” pedestrians in GIDAS to the corresponding national proportions from 2003 to 2007, weight factors could be derived.

3.2 Final sample

The GIDAS pedestrian sample was then queried for pedestrians hit by the front of a passenger car with assessed impact speed (pedestrians lying on the ground prior to impact, as well as sport utility vehicles and other light trucks and vans were excluded). Four hundred and ninety two pedestrians aged 15 years or older were found, including 36 fatalities. All fatal accidents, crashes with impact speeds exceeding 65 km/h, and 20 randomly selected cases were studied in detail to certify the data quality³. As a result of these investigations, two pedestrians, surviving impact speeds of 77 and 108 km/h respectively, were excluded from the sample, due to interaction mainly with the side structure of the car. (In other words, these two pedestrians were “sideswiped” by the car and did not receive much impulse in the collision) Hence, the final sample consisted of 490 pedestrians aged 15 to 96 years. Finally, the weight factors were normalised so that the total size of the weighted sample also equalled 490 pedestrians. In Table 2, the distributions of pedestrian injury severities are given for the national statistics, the total GIDAS sample, and the final GIDAS sample. From those data it is straight forward to derive the final, normalised weight factors: $W_{\text{slight}}=1.4$, $W_{\text{severe}}=0.69$, and $W_{\text{fatal}}=0.50$.

³ In-depth information provided in the GIDAS database included sketches, photographs, police reports, medical records, etc.

Table 2: Distributions of pedestrian injury severities

	Slight	Severe	Fatal
National (2003–2007)	70.9%	27.0%	2.2%
Total GIDAS	53.2%	42.1%	4.7%
Final sample	44.4%	48.2%	7.3%

Table showing the distributions of slightly, severely, and fatally injured pedestrians in the national statistics (Verkehrsunfälle, 2003–2007), the total GIDAS sample, and the final GIDAS sample (N=490). These figures were used to derive the weight factors.

3.3 Statistical methods

The distributions of pedestrian age, gender, height, and weight were investigated for both the total sample (N=490) and the fatalities (N=36). These empirical investigations were aimed at providing us with an understanding of the data and the problem at hand.

Logistic regression analysis was applied to the weighted sample in order to derive an analytical expression for the pedestrian fatality risk as a function of impact speed. The probability of death, $P(v)$, was then assumed to have the following form

$$P(v) = \frac{1}{1 + \exp(-a - bv)} \quad (1)$$

where v is the impact speed and a , b , two parameters to be estimated by the method of maximum likelihood (Dobson, 2002; McCullagh & Nelder, 1989). The effects of pedestrian age, gender, height, and weight were investigated by applying multiple logistic regression analysis. The main objective of the latter analysis was to find an improved, multivariate function describing pedestrian fatality risk. The model selection was based on a subset of 353 cases, including 21 fatalities, for which all the additional variables were known. We treated age, height, and weight as continuous variables, while gender was nominal.

Model fit investigations were based on Akaike's information criterion (Akaike, 1974), likelihood ratio tests, and Wald chi-square statistics, as well as visual assessment of residuals and influence diagnostics. Some further details are provided in Appendix B.

Table 3: Summary of empirical data

Speed (km/h)	Cases	Fatalities	Rate (%)	Wgt rate (%)	Speed (km/h)	Cases	Fatalities	Rate (%)	Wgt rate (%)
1–9	35	0	0	0	60–69	18	5	28	22
10–19	93	2	2.2	0.92	70–79	8	6	75	69
20–29	99	1	1.0	0.44	80–89	2	1	50	42
30–39	103	4	3.9	1.9	90–99	4	4	100	100
40–49	99	5	5.1	2.9	100–109	1	1	100	100
50–59	27	7	26	18	110–119	1	0	0	0

The total number of observed cases and fatalities by impact speed interval for raw data. Fatality rates are provided for both raw and weighted data.

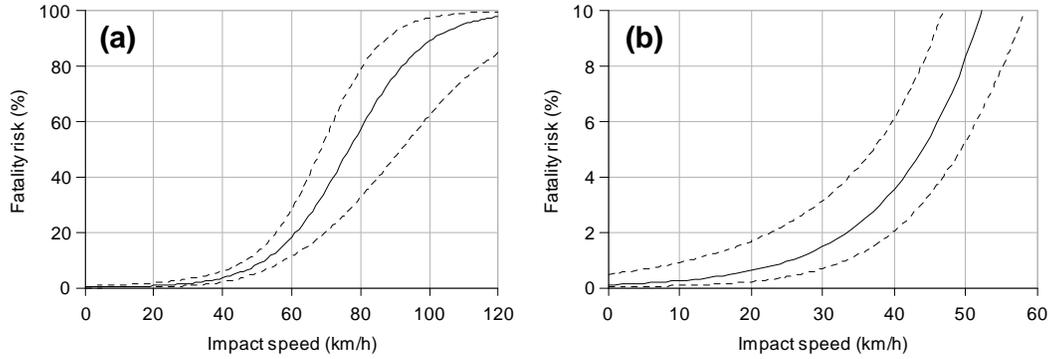


Figure 1: Pedestrian fatality risk

(a) The fatality risk as a function of impact speed for adult pedestrians hit by the front of a passenger car. The dotted curves show approximate 95% confidence limits. (b) Zoom in on the risk curve below 60 km/h.

4. Results

Table 3 shows the number of pedestrians and the fatality rates observed at different impact speed intervals. Comparing the raw and weighted fatality rates, we see that the effects of weighting decreased at higher impact speeds. We applied logistic regression analysis to fit an analytical function to the weighted fatality rates at all observed impact speeds. The resulting fatality risk function is presented in equation (2), where the impact speed, v , is measured in km/h.

$$P(v) = \frac{1}{1 + \exp(6.9 - 0.090v)} \quad (2)$$

The fatality risk function is also displayed in Fig. 1 together with an approximate 95% confidence band (see Appendix B for the mathematical formula). Zooming in on the risk curve at impact speeds below 60 km/h (Fig. 1b), we see that the relative risk increases rapidly with impact speed, which is in line with previous literature. However, the absolute risks are substantially lower than those generally reported (cf. the risk estimates by Anderson et al., 1997; Ashton, 1982; Pasanen, 1992; and Yaksich, 1964 given in Table 1, which have been the basis for the generally reported fatality risks).

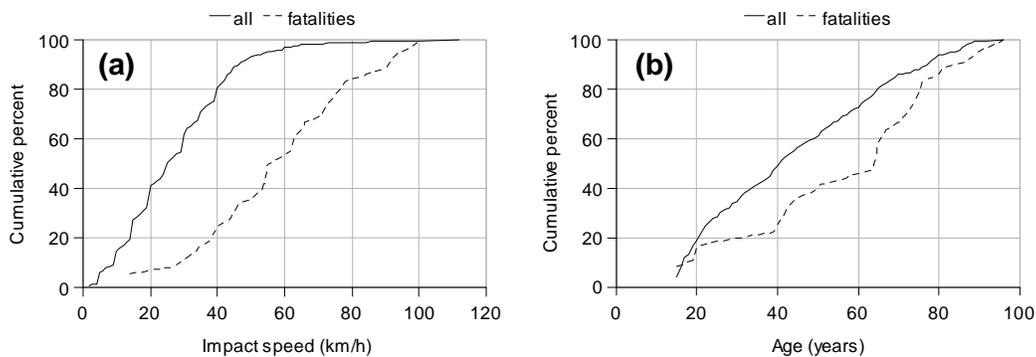


Figure 2: Distributions of impact speed and age

Cumulative distributions of (a) impact speed and (b) age for all pedestrians (N=490) and the fatalities (N=36).

We also investigated the effects of pedestrian age, gender, height, and weight on the fatality risk. Descriptive statistics for these explanatory variables for the weighted pedestrian sample and the fatalities are presented in Tables 4 and 5. Cumulative distributions of impact speed and pedestrian age are presented in Figs. 2a and b. The median impact speed and pedestrian age were 26 km/h and 42 years for the total sample, whereas they were 57 km/h and 65 years for the fatalities. This indicates that both these variables should be considered important predictors of pedestrian fatality risk. Earlier studies have also shown that the fatality risk increases with pedestrian age (see, e.g., Henary et al. (2006) and references therein). Furthermore, the percentage of females was 55% for the total sample, but only 36% for the fatalities. To choose the best fatality model, all possible linear combinations of the additional variables together with impact speed were tested. The model with the lowest value of Akaike's information criterion included only impact speed and pedestrian age. This was also the only model in which all explanatory variables were statistically significant according to the Wald chi-square test (two-tailed $\alpha=0.05$). The corresponding fatality risk function is presented in equation (3)⁴, where v is the impact speed measured in km/h, and age is the pedestrian age in years (the function is not applicable at ages below 15 years).

$$P(v, age) = \frac{1}{1 + \exp(9.1 - 0.095v - 0.040age)} \quad (3)$$

The above mentioned difference in gender was explained by females being exposed to lower impact speeds than males.

Some details of the results of the simple and multiple logistic regression analyses are shown in Table 6, while a more elaborate discussion on the model fit, including a brief analysis of residuals and influence diagnostics, is included in Appendix B.

5. Limitations

We stress that the fatality risk functions in this paper only apply to pedestrians aged 15 years or older. Furthermore, the results primarily describe the situation in Germany. In countries with less developed emergency and medical care, the fatality risk at any given impact speed would probably be higher.

6. Discussion and conclusions

Compared to the fatality risks generally reported in the traffic safety literature, equations (2) and (3) result in substantially lower estimates. However, a strong dependence on impact speed is present, with the risk at 50 km/h being more than twice as high as the risk at 40 km/h and more than five times higher than the risk at 30 km/h. This shows the importance of keeping impact speeds as low as possible within city areas where most pedestrian accidents occur. We also found that approximately 50% of all pedestrian fatalities had exposure to an impact speed between 50 and 80 km/h (Fig. 2). While these crashes are sometimes judged as virtually unsurvivable, the new risk curve shows that this is not the case. At an impact speed of 75 km/h, the fatality risk was estimated to approximately 50% (95% confidence interval: 26–68%). Hence, these high-speed crashes should be taken into account in future pedestrian safety activities.

⁴ Although the model selection was based on a subset of 353 cases (as described in subsection 3.3), the risk function in equation (3) was derived from the total sample, which comprised 490 cases.

Table 4: Descriptive statistics for the total sample

	N	Mean	Median	Std Dev	Min	Max
Impact speed	490	28 km/h	26 km/h	16 km/h	2 km/h	112 km/h
Age	490	45 years	42 years	22 years	15 years	96 years
Height	380	169 cm	169 cm	10 cm	116 cm	199 cm
Weight	380	70 kg	70 kg	14 kg	42 kg	120 kg
Car reg. year	452	1995	1995	4.6 years	1977	2006

Descriptive statistics for the weighted total pedestrian sample (N=490). The year of first registration for the cars is also included.

Table 5: Descriptive statistics for the fatalities

	N	Mean	Median	Std Dev	Min	Max
Impact speed	36	59 km/h	57 km/h	23 km/h	14 km/h	100 km/h
Age	36	57 years	65 years	24 years	15 years	96 years
Height	23	170 cm	173 cm	8.3 cm	153 cm	185 cm
Weight	21	73 kg	72 kg	16 kg	53 kg	102 kg
Car reg. year	35	1995	1995	6.0 years	1977	2006

Descriptive statistics for the fatalities (N=36). The year of first registration for the cars is also included.

Table 6: Logistic regression results

	Estimate	Lower limit	Upper limit	Standard error	Wald X²	P-value
Intercept	-6.9	-8.5	-5.3	0.81	72	<0.0001
v	0.090	0.060	0.12	0.016	34	<0.0001
Intercept	-9.1	-12	-6.6	1.3	49	<0.0001
v	0.095	0.063	0.13	0.016	34	<0.0001
age	0.040	0.012	0.068	0.014	7.7	0.0055

Details from the simple and multiple logistic regression analyses. The lower and upper limits are for a 95% Wald confidence interval.

Acknowledgements

We thank Thomas Svensson for comments and suggestions on the statistical analysis, and Torbjörn Andersson for comments on the manuscript.

References

- AAAM, 2001. The abbreviated injury scale: 1990 revision, update 1998. Association for the Advancement of Automotive Medicine. Barrington, IL, USA.
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–723.
- Anderson, R.W.G., McLean, A.J., Farmer, M.J.B., Lee, B.H., Brooks, C.G., 1995. Vehicle travel speeds and the incidence of fatal pedestrian crashes. In: *Proceedings of the 1995 International IRCOBI Conference on the Biomechanics of Impact*, Brunnen, Switzerland, pp. 107–117.
- Anderson, R.W.G., McLean, A.J., Farmer, M.J.B., Lee, B.H., Brooks, C.G., 1997. Vehicle travel speeds and the incidence of fatal pedestrian crashes. *Accid Anal Prev.* 29, 667–674.
- Ashton, S.J., Pedder, J.B., Mackay, G.M., 1977. Pedestrian injuries and the car exterior. SAE, Technical Paper 770092.
- Ashton, S.J., 1982. A preliminary assessment of the potential for pedestrian injury reduction through vehicle design. SAE, Technical Paper 801315.
- Cuerden, R., Richards, D., Hill, J., 2007. Pedestrians and their survivability at different impact speeds. In: *Proceedings of the 20th International Technical Conference on the Enhanced Safety of Vehicles*, Lyon, France, Paper No. 07–0440.
- Davis, G.A., 2001. Relating severity of pedestrian injury to impact speed in vehicle pedestrian crashes. *Transport Res Rec. Record No. 1773*, 108–113.
- Dobson, A.J., 2002. *An Introduction to Generalized Linear Models*. Chapman & Hall/CRC.
- ERSO, 2008 [homepage on the Internet]. European Road Safety Observatory [updated 2008 Jan 16; cited 2008 Jul 31]. Available from: http://www.erso.eu/knowledge/Fixed/40_Pedestrians/pedestrians.pdf.
- Funk, J.R., Cormier, J.M., Gabler, H.C., (2008). Effect of Delta-V errors in NASS on frontal crash risk calculations. In: *52nd Annals of Advances in Automotive Medicine*, San Diego, CA, USA.
- Global Road Safety Partnership, 2008. *Speed Management: A Road Safety Manual for Decision-Makers and Practitioners*. Geneva.
- Hannawald, L., Kauer, F., 2004. Equal effectiveness study on pedestrian protection. Technische Universität Dresden.
- Henary, B.Y., Ivarsson, J., Crandall, J.R., 2006. The influence of age on the morbidity and mortality of pedestrian victims. *Traffic Injury Prevention* 7, 182–190.
- Hugemann, W. (ed.), 2007. *Fachbuch Unfallrekonstruktion*. Autorenteam GbR, Münster.
- Interdisciplinary Working Group for Accident Mechanics, 1986. *The car-pedestrian collision: injury reduction, accident reconstruction, mathematical and experimental simulation: head injuries in two wheeler collisions*. University of Zürich and Swiss Federal Institute of Technology ETH.
- IRTAD, 2008 [homepage on the Internet]. International Road Traffic and Accident Database [cited 2008 Jul 7]. Available from: www.irtad.net.
- Kopits, E., Cropper, M., 2005. Traffic fatalities and economic growth. *Accid Anal Prev.* 37, 169–178.

- Kullgren, A., Lie, A., 1998. Vehicle collision accident data – validity and reliability. *J Traffic Med.* 26, 77–89.
- Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., 2004. *Applied Linear Statistical Models*. McGraw-Hill/Irwin, New York.
- McCullagh, P., Nelder, J.A., 1989. *Generalized Linear Models*. Chapman & Hall, London.
- McLean, A.J., Anderson, R.W.G., Farmer, M.J.B., Lee, B.H., Brooks, C.G., 1994. Vehicle travel speeds and the incidence of fatal pedestrian collisions: Volume I. The University of Adelaide, Report No. CR146.
- Mohan, D., 2002. Traffic safety and health in Indian cities. *Journal of Transport and Infrastructure* 9, 79–94.
- Odero, W., Garner, P., Zwi, A., 1997. Road traffic injuries in developing countries: a comprehensive review of epidemiological studies. *Tropical Medicine and International Health* 2, 445–460.
- OECD/ECMT Joint Transport Research Center, 2006. *Speed Management*. Paris.
- Oh, C., Kang, Y.S., Kim, W., 2008a. Development of probabilistic pedestrian fatality model for characterizing pedestrian-vehicle collision. *IJAT* 9, 191–196.
- Oh, C., Kang, Y.S., Youn, Y., Konosu, A., 2008b. Assessing the safety benefits of an advanced vehicular technology for protecting pedestrians. *Accid Anal Prev.* 40, 935–942.
- Otte, D., Krettek, C., Brunner, H., Zwipp, H., 2003. Scientific approach and methodology of a new in-depth investigation study in Germany so called GIDAS. In: *Proceedings of the 18th International Technical Conference on the Enhanced Safety of Vehicles*, Nagoya, Japan, Paper No. 161.
- Pasanen, E., 1992. *Driving speeds and pedestrian safety; a mathematical model*. Helsinki University of Technology, Transport Engineering, Publication 77.
- Pfeiffer, M., Schmidt, J., 2006. Statistical and methodological foundations of the GIDAS accident survey system. In: *Proceedings of the 2nd International Conference on ESAR*, Hanover, Germany, pp. 81–87.
- Pregibon, D., 1981. Logistic regression diagnostics. *Annals of Statistics* 9, 705–724.
- Tarri re, C., 1995. Children are not miniature adults. In: *Proceedings of the 1995 International IRCOBI Conference on the Biomechanics of Impact*, Brunnen, Switzerland, pp. 15–26.
- Teichgr ber, W., 1983. Die Bedeutung der Geschwindigkeit f r die Verkehrssicherheit. *Z f Verkehrssicherheit* 2, 53–63.
- Verkehrsunf lle 2003–2007 [database on the Internet]. Statistisches Bundesamt [updated 2008 Apr 3; cited 2008 Apr 14]. Fachserie 8 / Reihe 7. Available from: www.destatis.de.
- Walz, F.H., Hoefliger, M., Fehlmann, W., 1983. Speed limit reduction from 60 to 50 km/h and pedestrian injuries. In: *Proceedings of the 27th Stapp Car Crash Conference with International Research Committee on the Biokinetics of Impact (IRCOBI)*, pp. 277–285.
- World Health Organization, 2004. *World Report on Road Traffic Injury Prevention*. Geneva.
- Yaksich, S.J., 1964. *Pedestrians with milage: A study of elderly pedestrian accidents in St. Petersburg, Florida*. American Automobile Association, Washington D.C.
- Youn, Y., Kim, S., Oh, C., Shin, M., Lee, C., 2005. Research and rule-making activities on pedestrian protection in Korea. In: *Proceedings of the 19th International Technical Conference on the Enhanced Safety of Vehicles*, June 6–9, Washington D.C., USA.

Appendix A: Assessment of impact speed in GIDAS

The choice of reconstruction method for pedestrian accidents depends on the input data available (Hugemann, 2007). One of the key parameters for this article is the impact speed of the striking car. This can be calculated from the collision point and car end-position together with marks indicating braking and trajectories between those points. Additionally, the mass ratio between pedestrian and car can be used to calculate the vehicle's change of velocity during collision. Another approach makes use of the pedestrian throw distance to calculate the impact speed of the car. It is then necessary to account for the car front-end structure and the velocity of the pedestrian. Since the exact pedestrian end-position is not always known for GIDAS accidents, this method can lead to rather high uncertainty in the impact speed estimates. In cases where all the above mentioned input data are known, multi body computer simulations can be conducted using, e.g., the reconstruction software PC Crash. If neither of these methods can be applied due to lack of required input data, tables of experimental data with regard to car impact speed, pedestrian contact location, and wrap around distance are used. The upper bound method combines all of these methods to lower the uncertainty and also considers special restrictions and speed limitations due to environmental factors. In a last instance, witness statements are used to validate data for the accident reconstruction. If there is insufficient information available for any kind of reconstruction, impact speed will be coded as unknown. Note that pedestrian injury severity is never entered as a parameter in the reconstructions, which makes a potential systematic bias towards higher impact speeds for higher injury severities unlikely.

The derivation of risk curves is sensitive to errors in the impact speed assessments (Funk et al., 2008; Kullgren & Lie, 1998). A systematic error towards higher or lower impact speeds would inevitably shift the risk curve to the right or left respectively. In order to run a small test on the plausibility of the GIDAS reconstructions, the median impact speed for fatalities was compared to two other pedestrian real-world accident investigations. We therefore note from Table 5 that the median impact speed was 57 km/h for the 36 fatalities studied in this paper. In Adelaide, Australia, 181 fatally wounded pedestrians hit by motorised vehicles were investigated by the National Health and Medical Research Council Road Accident Research Unit between 1983 and 1991 (McLean et al., 1994). In Birmingham and Worcestershire, Great Britain, 81 pedestrian fatalities hit by passenger cars were investigated during the 1960s and 70s (Ashton, 1982). The median impact speed was between 50 to 60 km/h for both these studies. (We were not able to access any other source of such information.) Hence, the median impact speeds for the fatally wounded pedestrians in these three large real-world accident studies are in accord with each other. This provides an indication that the impact speed assessments made by GIDAS are comparable to other pedestrian studies around the world.

Appendix B: Details on the logistic regression

Some details of the results of the logistic regression analyses were provided in Table 6. Here we also note that Akaike's information criterion (AIC) decreased from 156 to 109 when including impact speed as a predictor and that the likelihood ratio chi-square test was highly significant (two-tailed $P < 0.0001$). AIC was further reduced to 103 for the multiple logistic regression model with both impact speed and pedestrian age as explanatory variables. Including an interaction effect between these variables made AIC increase. In practice, our model fit investigations also relied on visual inspection of regression diagnostics, including residuals and influence statistics (Pregibon, 1981). A subset of such diagnostics, for the simple logistic regression model, is displayed in Fig. B1, where they are plotted against the observation index. One can see from the Pearson residuals that two observations had poor fits compared to the others. These corresponded to two fatal cases with an impact speed of 14 km/h. From the confidence interval displacement diagnostics, it can be seen that one observation had a much

greater influence on the fitted parameters (a , b) in equation (1) than the others. This was a survivor with an impact speed of 112 km/h. These three accidents were studied in detail by considering the in-depth information provided in the GIDAS database, which included sketches, photographs, injuries, etc. From an epidemiological perspective it was then decided that these three cases should be kept in the study. As a test, the survivor at 112 km/h was temporarily removed, which changed the parameter a from -6.9 to -7.4 and b from 0.090 to 0.10 . These changes are within the standard errors of the parameters, see Table 6. Thus, from a statistical point of view, the temporal removal of this observation did not have a large effect on the risk curve.

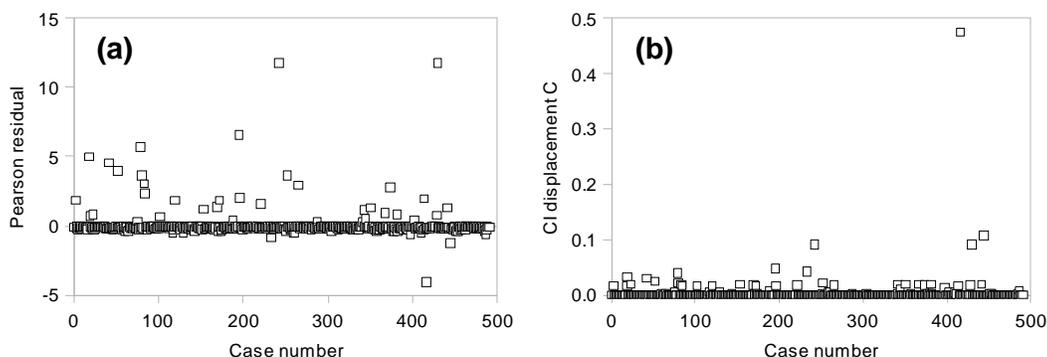


Figure B1: Residuals and influence statistics

Pearson residuals (a) and confidence interval displacement diagnostics (b) for the simple logistic regression analysis.

To investigate overdispersion (McCullagh & Nelder, 1989) the data were aggregated in speed intervals according to Table 3 (0-9 km/h, 10-19 km/h, 20-29 km/h, etc.) and logistic regression was conducted using the middle speed value for each speed interval (i.e. 5 km/h, 15 km/h, 25 km/h, etc.) as independent variable together with the weighted fatality rate at each interval as the dependent variable. For this logistic regression model, the deviance was 9.6 with 10 degrees of freedom, thus indicating that overdispersion was not a problem. The deviance goodness-of-fit test had a chi-square statistic of 0.96 giving a P -value = 0.48, which indicated a good model fit.

To investigate the effect of the weighting procedure, logistic regression was also applied to the raw/unweighted sample. The parameters a and b then became -6.0 and 0.083 , which is outside the standard errors of the parameters from the weighted analysis, but well within their 95% Wald confidence limits (see Table 6). The weighted and unweighted risk curves are compared in Fig. B2, from which it can be concluded that the weighting had an important effect at impact speeds below 50 km/h, but not at higher impact speeds. The raw and weighted empirical fatality rates at each speed band are included in Fig. B2, thus providing yet another way to assess the fit of the regression analysis to the data.

It is plausible that there is some under-reporting of especially slight injury accidents to the national statistics. This would mean that the weight factor for slightly injured pedestrians (called W_{slight} in subsection 3.2) would have to be a little larger in order to compensate for the under-reporting. To investigate the effect of such under-reporting, a very simple sensitivity study was conducted. We then multiplied the weight factor W_{slight} by a factor of 2, which

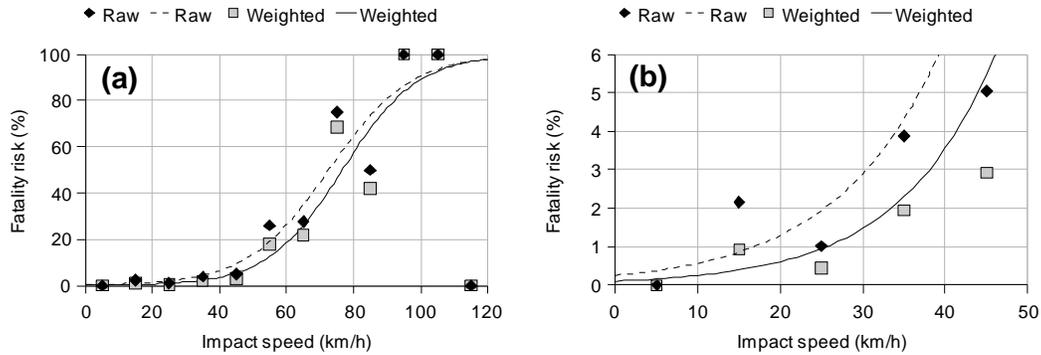


Figure B2: Comparison of the raw and weighted results

(a) Comparison of the empirical and analytical fatality risks for the raw and weighted data. The empirical fatality risks, or fatality rates, have been grouped in intervals of 10 km/h. (b) Zoom in on the risks below 50 km/h. The largest relative differences between the raw and weighted risks are at low speeds.

corresponded to an under-reporting of 50% of the slight accidents in the national statistics⁵. The parameters a and b then became -7.7 and 0.10 , which lies within the standard errors of the original estimates (see Table 6). Hence, this rather large change of the weight factors had only a small effect on the risk function. The analysis would be more sensitive to under-reporting of fatal accidents, but we see no reason for such a phenomenon.

A confidence band for the risk curve was derived following Kutner et al. (2004). In short, an approximate 95% confidence interval for the probability of death at an impact speed, v , is given by $1 / (1 + \exp(- a' - b' v \pm 1.96(X^T S^2 X)^{1/2}))$. Here X is a column vector with entries $(1, v)$; a' , b' are the maximum likelihood estimates of a , b ; and S^2 is the estimated approximate variance-covariance matrix for the regression parameters when the sample size is large. The final result of the confidence interval as a function of impact speed was $1 / (1 + \exp(6.9 - 0.090v \pm 1.96(0.66 - 0.024v + 0.00024v^2)^{1/2}))$.

⁵ It was not needed to re-normalise the weight factors, since only the main estimates for the logistic regression parameters were considered.