DTU Library

Improving knowledge of cyclist crashes based on hospital data including crash descriptions from open text fields

Møller, Mette; Janstrup, Kira Hyldekær; Pilegaard, Ninette

Published in:
Journal of Safety Research

Link to article, DOI:
10.1016/j.jsr.2020.11.004

Publication date:
2021

Document Version
Peer reviewed version

Improving knowledge of cyclist crashes based on hospital data including crash descriptions from open text fields

Møller, M.[a,*], Janstrup, K.H.[a], Pilegaard, N.[a]

[a]Technical University of Denmark, Department of Technology, Management and Economics, Division of Transport. DK-2800 Kgs. Lyngby, Denmark

***Corresponding author**

Technical University of Denmark, Department of Technology, Management and Economics, Section of Transport, Bygningstorvet, Building 116B, DK-2800 Kgs. Lyngby, mette@dtu.dk, Tel. +45 45256537

Acknowledgments

The authors would like to thank the Municipality of Aarhus, the Danish Asphalt Pavement Association and the Danish Safe Roads Association for their support to this study.

Improving knowledge of bicycle crashes based on hospital data including crash descriptions from open text fields

Abstract

Introduction: In this study we explore the added value of bicycle crash descriptions from open text fields in hospital records from the Aarhus municipality in Denmark. We also explore how bicycle crash data from the hospital complements crash data registered by the police in the same area and time period. *Method:* The study includes 5,313 Danish bicycle crashes, of which 4,205 were registered at the hospital and 1,078 by the police. All crashes occurred from 2010 to 2015. We performed an in-depth analysis of the open text fields on hospital records to identify factors associated with each crash using four categories: bicyclist, road, bicycle, and the other party. We employed the chi-squared test to compare the distribution of variables between crashes registered at the hospital and by the police. A binary logit model was used to estimate the probability that a crash factor is identified, and that each crash factor is associated with a single-bicycle crash. *Results:* The open-ended text fields in hospital records provide detailed information about crash factors not available in police records, including riding speed, inattention, clothing, specific road conditions and bicycle defects. The factors alcohol and kerb had the highest odds of being identified in relation to a single-bicycle crash. Crash data registered at the hospital included a larger number of bicycle crashes, particularly single-bicycle crashes and crashes with slight injuries only. *Conclusion:* Crash information registered at the hospital in Aarhus Municipality contributes to a better understanding of bicycle crashes due to detailed information about crash-associated factors as well as information about a larger number of bicycle crashes, particularly singe-bicycle crashes. *Practical implication:* Efforts to improve access to detailed information about bicycle crashes are needed to provide a better basis for bicycle crash prevention.

Keywords: Underreporting, In-depth analysis, Bicycle safety, Crash analysis, binary logit model.

1. Introduction

The individual positive environmental (Xia et al., 2013) and health benefits of bicycling generally outweigh the possible negative effects of increased exposure to air pollution (Woodward & Samet, 2016). These and other positive impacts of bicycling have contributed to a substantial increase in initiatives to promote bicycling over the past decade (Schepers & Heinen, 2013), as bicycling is considered a sustainable means of daily transport and is often mentioned as an important element to ensure a sustainable transport sector. Unfortunately, bicycling is related to a comparably high risk of road traffic injury (Elvik, 2009). In Denmark, the risk of being injured or killed in a bicycle crash is 13 times higher per travelled kilometre than the risk related to car travel (Christiansen & Warnecke, 2018). Research supports the idea of safety in numbers for bicyclists, and thus that the number of bicycle crashes does not increase proportionally with an increasing number of cyclists (Elvik & Bjørnskau, 2017; Jacobsen, 2003). However, changes in bicycling patterns and frequency do lead to changes in the type and severity of crashes (Schepers et al., 2017). Information on risk factors associated with bicycle crashes remains relevant to develop targeted preventive measures.

Traditionally, knowledge of road traffic crashes relies upon police registered data. However, due to the high level of underreported bicycle crashes in such crash data (e.g. Elvik & Mysen, 1999) information about the occurrence and characteristics of bicycle crashes is incomplete, particularly regarding single-bicycle crashes (e.g. Juhra et al., 2012). In Denmark it is estimated that the level of underreporting regarding bicycle crashes is approximately 90%, as only approximately 10% of bicyclist crashes are registered by the police and thus appear in the official national road traffic crash statistics (Janstrup et al., 2016). Hospital data do not include all bicycle crashes either, but previous studies have shown that the level of underreporting of bicycle crashes in hospital data is lower than in crash data registered by the police (e.g. Janstrup et al., 2016; Watson et al., 2015). Further insight into differences and similarities between bicycle crashes registered at a hospital and by the police is needed to fully understand the contribution and limitations of these data.

The Danish national road traffic crash database includes information about a number of factors systematically registered by the police. The registration follows a template that includes aspects such as time of day, weather condition, vehicles involved, crash situation, and crash location. However, information regarding the behaviour of the bicyclist, the road conditions and the interaction between the bicyclist and the surroundings is limited and restricted to aspects included in the template. Consequently, a detailed understanding of factors associated with the crash is difficult. Moreover, exploring influence from aspects not included in the template – such as distracted bicycling or an uneven surface – is not possible. To further improve the understanding of bicycle crashes, data sources that include more detailed information about the crash, the surroundings and the behaviour of the bicyclist and allow a more explorative approach is relevant. In line with the conceptual framework for bicycle safety outlined by Schepers et al. (2014b), previous studies applying an explorative approach on more detailed data sources (e.g. Møller & Haustein, 2016) have successfully categorized crash associated factors into three categories: (a) the road and its surroundings, (b) the vehicle and (c) the behaviour and condition of the road user. This categorization will also be applied in this study.

Based on the above, this study sets out to explore the added value of bicycle crash descriptions in open text fields in hospital records from Aarhus municipality. The study also explores how bicycle crash data from the hospital complement police registered crash data from the same area and time period, thereby improving the basis for increasing bicycle safety. Exploring crash descriptions in open text fields adds to previous research on bicycle crashes including hospital data (e.g. Langley et al., 2003; Lujic et al., 2008; Juhra et al. 2012; Watson et al., 2015) by facilitating a more detailed understanding of factors associated with a crash.

2. Method

2.1 *Road traffic crash registration by the police and at Aarhus municipality*

In Denmark, the police only register a bicycle crash if the police are notified about the crash and the crash occurred on a public road. Bicycle crashes occurring in a forest or on a private road are not registered. When notified, the police are only obligated to register the crash if injuries or material damage on a motor vehicle exceeds DKR 50000 (EUR 6700) or other material damage exceeds DKR 5000 (EUR 670). Due to these criteria, a high number of bicycle crashes are not registered by the Danish police. Regarding injury severity, the police uses a scale with four levels: *Fatalities*, the person dies within 30 days of the day of the crash; *severe injuries*, temporary or permanent incapacity; *slight injuries*, the injuries require medical treatment; and *no injuries*, bruises and/or property damage only.

Only a few Danish hospitals register detailed information about road traffic crashes and include open text fields. The hospital in Aarhus Municipality has registered detailed information about road traffic crashes systematically for many years. In addition to open text fields, the hospital follows a template that partly overlaps with the template used by the police, including the four-level injury scale. Road traffic crashes registered at the hospital include crash-involved persons arriving in an ambulance or similar and persons with less severe injuries able to travel to the hospital on their own for a medical check. There are no restrictions regarding crash location for a road traffic crash to be registered at the hospital.

2.2. Road traffic crash data in this study

This study includes all crashes involving bicyclists registered by the police (N=1,078) and at the hospital (N=4,205) in the Danish municipality of Aarhus from 2010 to 2015. Both data sources include information on crash characteristics such as day of the week, mode types, crash type (e.g. single-bicycle crash), road user information (gender, age and injury severity) and road conditions (e.g. surface and light conditions). In addition, the hospital data included an open-ended description of the crash containing detailed information about the crash location and cyclist behaviour prior to the crash.

2.3 Analysis

To explore the added value of bicycle crash descriptions in open-ended text fields in hospital records, we performed a manual three-step in-depth analysis; first, careful reading of the text; second, identification of crash factors, and third, categorizing the crash factors according to four categories: (a) the behaviour and condition of the bicyclist, (b) the road and its surroundings, (c) the bicycle and (d) the other party. A crash factor is a specific circumstance associated with the crash identified by the bicyclist or hospital staff. Table 1 provides a description of the crash factors regarding the bicyclist, the road and the bicycle. Regarding the category "the other party", the available information indicated that the behaviour of the other party was associated with the crash, but the information was too limited to specify a crash factor. We therefore decided to not include those crashes in the detailed analysis of the crash factors. In cases where the description of the crash was missing or the information provided was too limited, a "no crash factor" code was registered.

Table 1.
Description of Identified Crash Factors for Bicyclist, Road and Bicycle Categories.

Category	Crash factor	Description
Bicyclist	Alcohol	Cycling under the influence of alcohol
	Inattention	Non-cycling related activities or unspecified inattentiveness
	Bicycling speed	Bicycling speed too high for the conditions or 25^{+} km/h
	Handling the bicycle	Foot slipping on the pedal, stumbling getting on/off the bicycle, etc.
	Clothing, etc.	Clothing, bags, etc., getting stuck in the wheel
	Crowding	Too little space between the cyclist and other road users
	Violations	Not respecting right of way, red light riding, etc.
	Loss of control	Losing control over the bicycle for no specified reason
	Illness	Acute indisposition
Road	Slippery	Wet leaves, icy, etc.
	Kerb	Pedal or wheel hitting the kerb
	Design	Bike path too narrow for a bicyclist
	Objects on road	Wire, stone, etc., on the road,
	Road surface	Holes, bumpy surface
	Road works	Hitting or colliding with equipment related to road works
	Weather	Blinded by the weather condition (hard rain, snow, bright sunshine)
	Crossing animal	Crossing cat, etc.
Bicycle	Bicycle chain	Chain breaks
	Various bicycle defects	Saddle comes off, handlebar breaks, etc.
	Brakes	Brakes not working
	Gear	Shift of gears causes turbulence, etc.

To explore how bicycle crash data from the hospital complement crash data registered by the police in the same area and time period, we compared the distribution of relevant variables (e.g. injury severity, gender, age group, crash type, light condition, year and season) between the two data sets using the chi-squared test. We then employed a binary logit model to estimate the probability that a crash factor, and a crash factor in a specific category, is identified, as well as the probability that each crash factor was associated with a single-bicycle crash. The binary logit model estimated the probability, based on a function of a vector X_i of observable variables. The observable variables included person and crash characteristics. A significance level of 0.05 was used in all analyses.

3. Results

3.1. Differences between crash data registered at the hospital and by the police

The number of crash-involved bicyclists is significantly higher in the hospital data (N=4205, 45%) compared to the police data (N=1078, 11%). We also identified significant differences regarding the injury severity of the involved bicyclists. In the hospital data, the percentage of bicyclists with slight injuries (N=2410, 57.3%) was significantly higher than in the police data (N=113, 10.5%). The police data include a higher percentage of bicyclist fatalities, but the absolute number of bicyclist fatalities is higher in the hospital data (Table 2). The only variable where no significant differences were found was gender, but here the percentage of bicyclists where gender information is missing is higher in the police data (N=108, 10.0%) than the hospital data (N=9, 0.2%); the same is true for. In general, the hospital data include a higher percentage of bicyclists in all age groups.

Table 2
Bicycle Crashes Registered by Police and at the Hospital in Aarhus Municipality from 2010 to 2015.

Variable	Category	Police data N	Police data %	Hospital data N	Hospital data %	χ^2-test, p-value
Severity	Fatality	10	0.9	12	0.3	
	Severe injury	209	19.4	832	19.8	
	Slight injury	113	10.5	2410	57.3	
	No injury/N/A	746	69.2	951	22.6	<0.001
Gender	Male	491	45.5	2115	50.3	
	Female	479	44.5	2081	49.5	
	N/A*	108	10.0	9	0.2	0.9046
Age	0-8 years old	4	0.4	63	1.5	
	9-17 years old	51	4.7	351	8.3	
	18-29 years old	455	42.2	1636	38.9	
	30-45 years old	203	18.9	823	19.6	
	46-65 years old	219	20.3	1083	25.8	
	Older 65 years old	38	3.5	249	5.9	
	N/A*	108	10.0	0	0.0	<0.001
Crash type	Single-bicycle crash	36	3.3	2287	54.4	
	Collision with pedestrian or animal	23	2.1	148	3.5	
	Collision with vehicle on straight road	140	13.0	601	14.3	
	Intersection, collision with vehicle on same road	494	45.8	656	15.6	
	Intersection, collision with vehicle from crossing road	385	35.8	513	12.2	<0.001
Light condition	Daylight	872	80.9	2827	67.2	
	Twilight	35	3.2	411	9.8	
	Darkness	168	15.6	926	22.0	
	N/A*	3	0.3	41	1.0	<0.001
Surface	Dry	775	71.9	2096	49.8	
	Wet	224	20.8	679	16.1	
	Slippery	23	2.1	378	9.0	
	N/A*	56	5.2	1052	25.1	<0.001
Day	Weekday	941	87.3	3348	79.6	
	Weekend	137	12.7	857	20.4	<0.001
Season	Winter	185	17.2	807	19.2	
	Spring	247	22.9	1010	24.0	
	Summer	284	26.3	1176	28.0	
	Autumn	362	33.6	1212	28.8	0.008
Year	2010	130	12.1	616	14.6	
	2011	170	15.8	846	20.1	
	2012	195	18.1	723	17.2	
	2013	208	19.3	757	18.0	
	2014	188	17.4	696	16.6	
	2015	187	17.3	567	13.5	<0.001

Note: * The N/A category is not included in the statistical test.

Regarding crash type, the distribution is significantly different in the two data sets. The hospital data include a larger percentage of single-bicycle crashes (N=2287, 54.4%) compared to the police data (N=36, 3.3%), whereas the police data include a larger percentage of intersection crashes. The percentage of crashes occurring in twilight and darkness is higher in the hospital data (31.8% vs 18.8%), as is the percentage of crashes occurring during the weekend (20.4% vs 12.7%). Most crashes are registered in the autumn for both data sets, but the distribution across seasons is different.

3.2 Crash factors derived from crash descriptions in open text fields

For 1,274 (30%) of the bicycle crashes registered at the hospital, a crash factor could not be identified due to missing or limited information. Out of the 2,931 crashes with an identified crash factor, 1,025 (33%) regarded the bicyclist, 956 (31%) the road, 96 (3%) the bicycle and 1,038 (33%) the other party. Table 3 provides an overview of the crash factors identified in the first three categories, as well as the severity degree. In the bicyclist category, alcohol (N=262, 26%) and inattention (N=259, 25%) were the most frequently identified crash factors. Clothing, bags and similar factors were identified in 9% of cases and crowding in 9% of the cases.

Table 3
Overview of Crash Factors and Injury Severity of the Involved Bicyclist

		Fatality or severe injury		Slight injury		No injury		Total		χ^2-test,
Category	Crash factor	N	%	N	%	N	%	N	%	p-value
Bicyclist	Alcohol	55	21.0	137	52.3	70	26.7	262	100	0.786
	Inattention	60	23.2	148	57.1	51	19.7	259	100	0.120
	Bicycling speed	41	36.6	51	45.5	20	17.9	112	100	<0.001
	Handling the bicycle	24	25.5	56	59.6	14	14.9	94	100	0.061
	Clothing, etc.	15	16.7	55	61.1	20	22.2	90	100	0.454
	Crowding	22	24.7	51	57.3	16	18.0	89	100	0.248
	Violations	6	12.0	27	54.0	17	34.0	50	100	0.207
	Loss of control	10	20.4	25	51.0	14	28.6	49	100	0.853
	Illness[a]	3	15.0	9	45.0	8	40.0	20	100	-
	Total	236	23.0	559	54.5	230	22.4	1025	100	0.011
Road	Slippery road	54	17.5	170	55.2	84	27.3	308	100	0.469
	Kerb	40	17.3	146	63.2	45	19.5	231	100	0.034
	Road design	14	11.5	67	54.9	41	33.6	121	100	0.021
	Objects on the road	28	23.9	68	58.1	21	17.9	117	100	0.182
	Road surface	25	31.6	39	49.4	15	19.0	79	100	0.040
	Road works	13	24.5	31	58.5	9	17.0	53	100	0.362
	Weather[a]	4	16.7	14	58.3	6	25.0	25	100	-
	Crossing animal[a]	4	18.2	13	59.1	5	22.7	22	100	-
	Total	182	19.0	548	57.3	226	23.6	956	100	0.002
Bicycle	Bicycle chain	6	14.0	28	65.1	9	20.9	43	100	0.363
	Various bicycle defects	10	32.3	14	45.2	7	22.6	31	100	0.254
	Brakes[a]	6	37.5	8	50.0	2	12.5	16	100	-
	Gear[a]	2	33.3	2	33.3	2	33.3	6	100	-
	Total	24	25.0	52	54.2	20	20.8	96	100	0.1302

[a]Category not included in χ^2-test.

Few crash factors were significantly associated with any particular severity degree (Table 3). For crashes where high riding speed (N=41, 36.6%) or road surface (N=25, 31.6%) were identified as crash factors, significantly more bicyclists than expected were fatally or severely injured. For crashes where problems with a kerb (N=45, 19.5%) or the road design (N=41, 33.6%) were identified as crash factors, more cyclists than expected sustained no injury after the crash.

Regarding gender, we found a significant difference in the distribution of crash factors in the bicyclist category (p>0.001), but no significant gender differences in the distribution of crash factors in the two other studied categories (road and bicycle). Regarding the specific crash factors alcohol (N=184, 70.2%) and bicycling speed (N=67, 59.8%), these are more common in crashes with a male bicyclist, whereas handling the bicycle (N=59, 62.8%) and violations (N=34, 68%) are more frequent crash factors in crashes involving a female bicyclist (Table 4). For the categories inattention, crowding and illness, we found no significant gender differences.

Table 4
Overview of Crash Factors in Bicyclist Category by Gender

		Male		Female		Total		χ^2-test,
Category	Crash factor	N=1463	Pct.=49.9	N=1461	Pct.=49.8	N=2931	Pct.=100	p-value
Bicyclist	Alcohol	184	70.2	78	29.8	262	100	<0.001
	Inattention	125	48.4	133	51.6	259	100	0.626
	Bicycling speed	67	59.8	45	40.2	112	100	0.042
	Handling the bicycle	35	37.2	59	62.8	94	100	0.015
	Clothing, etc.	45	50.0	45	50.0	90	100	0.995
	Crowding	38	43.2	50	56.8	89	100	0.205
	Violations	16	32.0	34	68.0	50	100	0.014
	Loss of control	24	49.0	25	51.0	49	100	0.884
	Illness	8	40.0	12	60.0	20	100	0.371

The odds of identifying a crash factor are lower than the odds of not doing so (Table 5). However, compared to crashes with no injury, the probability that a crash factor is identified is higher for injury crashes, particularly for crashes involving severe injury. Regarding crash type, the probability is lowest for single-bicycle crashes and for crashes in “intersections, collision with vehicle from same road”. The odds of identifying a crash factor are highest for crashes with “collision with pedestrian or animal” (OR=3.557), but the 95% confidence interval is high (2.151-5.883) which indicates some uncertainty in the model.

Table 5

Factors Influencing Probability a Crash Factor Was Identified.

Variable	Category	Estimate	p-value	OR	95% CI
Intercept		1.122	<0.0001	—	—
Severity	Fatality	-0.934	0.1524	0.393	(0.109-1.412)
	Severe injury	-0.370	0.0008	0.690	(0.556-0.858)
	Slight injury	-0.589	<0.0001	0.555	(0.464-0.663)
	No injury/NA	—	—	—	—
Type	Single-bicycle crash	—	—	—	—
	Collision with pedestrian or animal	1.269	<0.0001	3.557	(2.151-5.883)
	Collision with vehicle on straight road	0.334	0.0013	1.398	(1.140-1.714)
	Intersection, collision with vehicle on same road	0.164	0.0944	1.178	(0.972-1.427)
	Intersection, collision with vehicle on crossing road	0.343	0.0021	1.409	(1.133-1.753)
Log-likelihood					-2491
AIC					4998
Pseudo R-square					0.110

Note: 59 observations were deleted due to missing information.

The binary logit model was also employed to see if some crash factors were more likely to be associated with single-bicycle crashes (see Table 6) compared to bicycle crashes involving other parties (vehicle, animal or pedestrian). Some observations were removed from the model due to missing information. The final model is based on information from 2,989 bicyclists of which 1,499 were involved in a single-bicycle crash.

Table 6

Odds of a Single-Bicycle Crash Compared to Bicycle Crashes involving Several Parties for Each Crash Factor

Variable	Category	Estimate	p-value	OR	95% CI
Intercept		-2.939	<0.0001	—	—
Bicyclist	Alcohol	5.421	<0.0001	226.035	(130.651-391.055)
	Inattention	1.425	<0.0001	4.159	(2.751-6.287)
	Bicycling speed	3.364	<0.0001	28.890	(17.767-46.978)
	Handling the cycle	4.575	<0.0001	97.063	(52.601-179.107)
	Clothing, etc.	5.761	<0.0001	317.639	(124.094-813.046)
	Loss of control	4.165	<0.0001	64.374	(30.338-136.595)
Road	Slippery	6.314	<0.0001	552.183	(278.889->999.999)
	Kerb	5.555	<0.0001	258.486	(139.067-480.45)
	Design	3.624	<0.0001	37.486	(23.157-60.683)
	Objects on the road	6.220	<0.0001	502.536	(179.026->999.999)
	Road surface	6.145	<0.0001	466.597	(142.819->999.999)
	Road works	5.366	<0.0001	213.973	(74.489-614.646)
	Weather	3.405	<0.0001	30.124	(12.617-71.921)
Bicycle	Bicycle chain	5.960	<0.0001	387.575	(91.504->999.999)
	Other defects	5.604	<0.0001	271.450	(81.805-900.74)
Log-likelihood					-770
AIC					1572
Pseudo R-square					0.889

Note: 1307 observations were deleted due to missing information.

Many of the identified crash factors were significantly associated with the odds of a single-bicycle crash, but the odds ratio values, together with the 95% confidence intervals, show that the uncertainty of these numbers is very high, especially for the variables *slippery*, *objects on the road*, *road surface* and *bicycle chain*. This problem is due to the very low number of non-single-bicycle crashes for which these variables were identified as a crash factor.

Regarding the bicyclist category, we see that the crash factor *clothing* has the highest estimate (5.761), and thus that the probability that the crash is a single-bicycle crash is highest when this crash factor is identified. However, the 95% CI is very wide for this crash factor. The crash factor *alcohol* also has a high estimate (5.421), and the odds of having a single-bicycle crash is more than 200 times higher compared to a crash with another party involved. For this crash factor, the 95% CI is not that big (130.651–391.055). The lowest estimate is found for the crash factor *inattention* (1.425), but the odds of a single-bicycle crash increases by a factor four compared to a crash with another party involved. The small 95% CI (2.751–6.287) for the OR-value for *inattention* shows that the certainty of this number is quite high.

Regarding the road category, the highest probability of a crash being a single-bicycle crash is found when the road is *slippery* (6.314), there is an *object on the road* (6.220) or there are other problems with the *road surface* (6.145). However, the 95% CI for the OR-values of these variables shows a very high uncertainty. Instead the focus should be on the estimates and OR-values for *kerb* (5.555, OR=258.486), *design* (3.624, OR=37.486), *road works* (5.366, OR=213.973) and *weather* (3.405, OR=30.124). The odds of being a single-bicycle crash are highest when the *kerb* is identified as a crash factor.

For the bicycle category, the *bicycle chain* has the highest probability (5.960) but the 95% CI is very big and associated with a very high degree of uncertainty.

4. Discussion

The purpose of this study was to explore the added value of bicycle crash descriptions from open text fields in hospital records from Aarhus municipality in Denmark. An additional purpose was to explore how bicycle crash data from the hospital complement crash data registered by the police in the same area and time period. The results suggest that crash data registered at the hospital complement information about bicycle crashes, as these data allow the identification of crash factors not available in the police registered crash data. In addition, hospital data include information about a larger number of bicycle crashes than police crash data, particularly regarding single-bicycle crashes.

The information provided in the open text fields of the medical records allowed detailed insight into factors associated with each crash, such as inattention and specific road conditions. Previous studies have shown that errors are associated with crashes, as well as near crashes, among bicyclists (e.g. Puchades et al., 2017; Twisk et al., 2015). Our results support this and add to the existing literature by identifying specific error types not available for police registered crashes, such as clothing and bags getting stuck in the wheel by mistake, feet slipping on the pedal when trying to get on/off the bicycle and bicycle defects. The available information did not allow a detailed understanding of the reasons behind the different behaviours associated with each crash, and this therefore remains a relevant topic for future research. In general, our results support the need for measures to reduce unsafe bicyclist behaviour and ensure a bicycle-friendly environment which reduces errors and related consequences in case of a crash, as suggested by the safe systems approach (Wegman et al., 2012).

Regarding bicyclist behaviour, our results confirm the results of previous studies (e.g. Billot-Grasset et al., 2016) which identified bicyclist behaviour as a key factor in bicycle crashes. Alcohol impairment and inattention were the most frequently identified factors. Like Orsi et al. (2014), we found that female bicyclists were less likely to be influenced by alcohol at the time of the crash, whereas inattention was more prevalent among female bicyclists. Alcohol is known to decrease the ability to behave safely in traffic due to factors such as decreased reaction times, increased error rates, tunnel vision and slower visual information processing (e.g. Friedman et al., 2011). Evidence has also been established for the harmful effects of being inattentive to bicycling tasks (e.g. de Waard et al., 2011; Terzano, 2013). Our results indicate that efforts to prevent drunk bicycling, as previously suggested by Andersson and Bunketorp (2002), are still highly relevant, particularly for male bicyclists. The data do not allow conclusions regarding the reasons behind drunk cycling; however, underestimation of personal risk has previously been identified as a contributing factor (e.g. Hagemeister & Kronmeier, 2017).

Regarding the environment, a few previous studies have found that roads with poor surface conditions may lead to traffic disruption and increased crash risk (e.g. Corazza et al., 2016; Janstrup et al., 2019; Pulugurtha et al., 2013). Our results add to this by showing that aspects such as road work sites and slippery and/or bumpy surfaces are also associated with crash involvement among bicyclists. The data do not allow us to draw conclusions regarding specific causal effects; however, in line with the safe systems approach (e.g. Wegman et al., 2012), the data do highlight the importance of designing road environments which support safe behaviours and allow bicyclists to compensate safely in case of human error. Bicycle defects were only identified as a crash factor in a small number of crashes, but nevertheless indicate that better bicycle maintenance has the potential to improve bicyclists' safety.

Our results indicate that women are involved in bicycle crashes related to some type of violation (e.g. no lights, illegal manoeuvre) at the time of the crash more frequently than men. This is somewhat surprising, as male road users are generally found to engage in more traffic violations (e.g. Varet et al., 2018), and a previous study found male bicyclists to be less compliant with road traffic rules (Johnson et al., 2011). It is possible, however, that the gender difference regarding violations is partly related to the data registration procedure. In general, persons identifying with the feminine gender role expectations are more prone to guilt (Benetti-McQuoid & Bursik, 2005; Efthim et al., 2001). This may cause female cyclists to be more willing to admit engaging in traffic violations compared to male cyclists. However, additional studies looking specifically into gender-based differences in bicycling skills, trip purpose, gender role expectations and other aspects that may explain or contribute to the identified gender differences are relevant, as limited information about gender differences and bicyclist safety exist (Stipancic et al., 2016). Our results show that information about gender is more complete in hospital data, which is thus a relevant data source for studies on gender differences in bicycling crashes.

Significantly more crash-involved bicyclists were registered in the hospital data than the police data, thereby confirming that bicycle crash data from the police suffer from a high level of underreporting (e.g. Aslop & Langley, 2001; Langley et al., 2003), and that both data types are relevant for improving bicyclist safety (e.g. Short & Caulfield, 2014). This is particularly true for single-bicycle crashes and crashes involving only slight injuries. Half of the bicyclists registered at the hospital were injured in single-bicycle crashes. This is in line with results from previous studies (e.g. Beck et al., 2016), although the share of single-bicycle crashes in our results was lower than elsewhere (e.g. Schepers et al., 2014a). Importantly, our results also showed that, for some crash types, the number of crashes registered by the police was higher than the crash data registered at the hospital. Thus, both data sources are needed to get a complete picture of the prevalence of bicyclist crashes, as neither source is complete.

In our study, the possibility of identifying a crash factor strongly depended on the information provided by the hospital. The results showed that the level of detail provided decreased with increasing crash severity. The data did not allow conclusions regarding the reasons behind this, but it is possible that the severely injured bicyclists' need for immediate care allowed hospital staff less time to focus on gaining a description of the crash situation. Nevertheless, if trying to gain a detailed understanding of the factors associated with bicycle crashes, our results indicate crash data registered at a hospital may be more suitable for less severe crashes than for police registered crashes. However, the probability of identifying a crash factor was lower for single-bicycle crashes compared to all other crash types, which indicates that although more crash details are available in hospital data than police data, efforts are needed to ensure that more information about these and other types of bicycle crashes is registered. In addition, initiatives to ensure that crash information is made available for road safety researchers and practitioners are relevant to improve road safety for cyclists.

5. Limitations

National hospital data regarding cyclist crashes is not available in Denmark. The data included in this study were therefore limited to Aarhus municipality, one of the few Danish municipalities that registered bicycle crash information at the hospital for many years. To validate the results obtained in the present study on a national level, studies including data from a larger share of Danish municipalities would be relevant, but not possible at the moment. Assessing the degree to which the results are transferrable to other countries is difficult; first, because access to the information included is generally highly restricted, and second, because data collection procedures most likely differ across countries. However, similar in-depth analysis would be relevant in other countries as well.

As in other studies (Imprialou & Quddus, 2017), our results are based on information registered by the hospital staff, which has not been verified for accuracy. It has been shown that there is some inconsistency between road traffic crash information registered by the police and by hospital staff (e.g. Lopez et al., 2000), partly due to the different purpose for which the data are collected (e.g. civil claims and trauma treatment), which may influence the reporting (Imprialou & Quddus, 2017). Furthermore, like other studies including qualitative data (see Aust et al., 2012; Møller & Haustein, 2016), some variation was found in the type of information and the level of detail provided. These variations may stem from several known and unknown sources (e.g. time pressure or misinterpretations on the part of the hospital staff and social desirability; Lajunen et al., 1997) on the part of the injured bicyclist, influencing him or her to leave out certain types of information and prioritize the inclusion of others. However, it was not possible to control for this influence in the present study.

6. Conclusion

This study showed that crash data registered at the hospital contribute significantly to our understanding of bicyclist crashes both in terms of the number of crashes and associated factors. The results showed that preventive measures to reduce inappropriate bicyclist behaviour and poor road conditions are of key importance to improve bicyclist safety. Measures to ensure well-maintained bicycles are less crucial but are also relevant for crash and injury prevention. This study confirmed that more bicyclist crashes are registered at the hospital than by the police and that hospital data are important to improve information about bicycle crashes. As hospital data and police data are reported differently, both are necessary to increase the understanding of crash factors and improve crash prevention efforts.

Acknowledgement

Financial support from the Danish Asphalt Pavement Association and the Danish Safe Roads Association is gratefully acknowledged. The authors would like to thank the Municipality of Aarhus for support regarding the data, and the reviewers for their comments, which significantly improved the paper.

References

Andersson, A.L., & Bunketorp, O. (2002). Cycling and alcohol. *Injury*, *33*, 467–471.

Aslop, J., & Langley, J. (2001). Under-reporting of motor vehicle traffic crash victims in New Zealand. *Accid. Anal. Prev.*, *33*, 353–359.

Aust, M.L., Fagerlind, H., & Sagberg, F. (2012). Fatal intersection crashes in Norway: patterns in contributing factors and data collection challenges. *Accid. Anal. Prev.*, *45*, 782–791.

Beck, B., Stevenson, M., Newstead, S., Cameron, P., Judson, R., Edwards, E.R., Bucknill, A., Johnson, M., & Gabbe B. (2016). Bicycling crash characteristics: An in-depth crash investigation study. *Accid. Anal. Prev.*, *96*, 219–227.

Benetti-McQuoid, J., & Bursik, K. (2005). Individual differences in experiences of and responses to guilt and shame: Examining the lenses of gender and gender role. *Sex Roles*, *53*, 133–142.

Billot-Grasset, A., Amoros, E., & Hours, M. (2016). How cyclist behavior affects bicycle accident configurations? *Transp. Res. F*, *41*, 261–276.

Christiansen, H., & Warnecke, M.L. (2018). *Risiko i trafikken 2007-2016. DTU-report 2018*. http://orbit.dtu.dk/files/146185353/Risiko_i_trafikken_2007_16.pdf

Corazza, M.V., Mascio, P.D., & Moretti, L. (2016). Managing sidewalk pavement maintenance: a case study to increase pedestrian safety. *J. Traffic Transp. Eng.*, *3*, 203–214.

De Waard, D., Edlinger, K., & Brookhuis, K. (2011). Effects of listening to music, and of using a handheld and handsfree telephone on cycling behavior. *Transp. Res. F*, *14*, 626–637.

Efthim, P.W., Kenny, M.E., & Mahalik, J.R. (2001). Gender role stress in relation to shame, guilt, and externalization. *Journal of Counselling & Development*, *79*, 46–54.

Elvik, R. (2009). The non-linearity of risk and the promotion of environmentally sustainable transport. *Accid. Anal. Prev.*, *41*, 849–855.

Elvik, R., & Bjørnskau, T. (2017). Safety-in-numbers: A systematic review and meta-analysis of evidence. *Safety Science*, *92*, 274–282.

Elvik, R., & Mysen, A.B. (1999). Incomplete accident reporting: Meta-analysis of studies made in 13 countries. *Transp. Res. Rec.*, *1665*, 33–140.

Lujic, S., Finch, C., Boufous, S., & Hayen, A. (2008). How comparable are road traffic crash cases in hospital admissions data and police records? An examination of data linkage rates. *Aust N Z J Public Health*, *32*, 28–33.

Friedman, T.W., Robinson, S.R., & Yelland, G.W. (2011). Impaired perceptual judgment at low blood alcohol concentrations. *Alcohol*, *45*, 711–718.

Hagemeister, C., & Kronmeier, M. (2017). Alcohol consumption and cycling in contrast to driving. *Accid. Anal. Prev.*, *105*, 102–108.

Imprialou, M., & Quddus, M. (2017). Crash data quality for road safety research: current state and future directions. *Accid. Anal. Prev.*, *130*, 84–90. http://dx.doi.org/10.1016/j.aap.2017.02.022

Jacobsen, P.L. (2003). Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Inj. Prev*. *9*, 205–209.

Janstrup, K.H., Kaplan, S., Hels, T., Lauritsen, J., & Prato, C.G. (2016). Understanding traffic crash under-reporting: Linking police and medical records to individual and crash characteristics. *Traffic Inj. Prev.*, *17*, 580–584.

Janstrup, K.H., Møller, M., & Pilegaard, N. (2019). A clustering approach to integrate traffic safety in road maintenance prioritization. *Traffic Inj. Prev.*, *20*(4), 442–448. https://doi.org/10.1080/15389588.2019.1580700.

Juhra, C., Wieskötter, B., Chu, K., Trost, L., Weiss, U., Messerschmidt, M., Malczyk, A., Heckwolf, M., & Raschke, M. (2012). Bicycle accidents – Do we only see the tip of the iceberg? A prospective multi-center study in a large German city combining medical and police data. *Injury*, *43*, 2026–2034

Lajunen, T., Carry, A., Summala, H., & Harley, L. (1997). Impression management and self-deception in traffic behavior inventories. *J. Pers. Ind. Dif.*, *22*, 341–353.

Langley, J.D., Dow, N., Stephenson, S., & Kypri, K. (2003). Missing cyclists. *Inj. Prev.*, *9*, 376–379.

Lopez, D.G., Rosman, D.L., Jelinek, G.A., Wilkes, G.J., & Sprivulis, P.C. (2000). Complementing police road-crash records with trauma registry data – an initial evaluation. *Accid. Anal. Prev.*, *32*, 771–777.

Møller, M., & Haustein, S. (2016). Factors contributing to young moped rider accidents in Denmark. *Accid. Anal. Prev.*, *87*, 1–7.

Orsi, C., Ferraro, O.E, Montomoli, C., Otte, D., & Morandi, A. (2014). Alcohol consumption, helmet use and head trauma in cycling collisions in Germany. *Accid. Anal. Prev.*, *65*, 97–104.

Puchades, V.M., Pietrantoni, L., Fraboni, F., Angelis, M.D., & Prati, G. (2017). Unsafe cycling behaviors and near crashes among Italian cyclists. International journal of injury control and safety promotion. *International Journal of Injury Control and Safety Promotion*, *25*(1), 70–77https://doi.org/10.1080/17457300.2017.1341931

Pulugurtha, S.S., Ogunro, V., Pando, M. A., Patel, K. J., & Bonsu, A., (2013). Preliminary results towards developing thresholds for pavement condition maintenance: Safety perspective. *Procedia – Social and Behavioral Science*, *104*, 302–311.

Schepers, P., Agerholm, N., Amoros, E., Benington, R., Bjørnskau, T., Dhondt, S., de Geus, B., Hagemeister, C., Loo, B.P.Y, & Niska, A. (2014a). An international review of the frequency of single-bicycle crashes (SBCs) and their relation to bicycle modal share. *Inj. Prev.*, *21*, e138–e143.

Schepers, P., Hagenzieker, M., Methorst, R., van Wee, B., & Wegman, F. (2014)b. A conceptual framework for road safety and mobility applied to cycling safety. *Accid. Anal. Prev.*, *62*, 331–340.

Schepers, P., & Heinen, E. (2013). How does a modal shift from car trips to cycling affect road safety? *Accid. Anal. Prev.*, *50*, 1118–1127.

Schepers, P., Stipdonk, H., Methorst, R., & Oliver, J. (2017). Bicycle fatalities: Trends in crashes with and without motor vehicles in the Netherlands. *Transp. Res. Part F*, *46*, 491–499.

Short, J., & Caulfield, B. (2014). The safety challenge of increased cycling. *Transp. Pol.*, *33*, 154–165.

Stipancic, J., Zangenehpour, S., Miranda-Moreno, L. Saunier, N., & Granié, M.A. (2016). Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology. *Accid. Anal. Prev.*, *97*, 19–27.

Terzano, K. (2013). Bicycling safety and distracted behavior in The Hague, the Netherlands. *Accid. Anal. Prev.*, *57*, 87–90.

Twisk, D.A.M., Commandeur, J.J.F., Vlakveld, W.P., Shope, J.T., & Kok, G. (2015). Relationships amongst psychological determinants, risk behaviour, and road crashes of young adolescent pedestrians and cyclists: implications for road safety education programmes. *Transp. Res. Part F*, *30*, 45–56.

Varet, F., Granié, M.A., & Apostolidis, T. (2018). The role of individualism, gender and situational factors on probabilities of committing offences in a French drivers sample. *Transp. Res. Part F*, *56*, 293–305.

Watson, A., Watson, B., & Vallmuur, K. (2015). Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accid. Anal. Prev*., *83*, 18–25.

Wegman, F., Zhang, F., & Dijkstra, A. 2012. How to make more cycling good for road safety. *Accid. Anal. Prev.*, *44*, 19–29.

Woodward, A., & Samet, J. (2016). Active transport: Exercise trumps air pollution, almost always. *Prev. Med*., *87*, 237–238.

Xia, T., Nitschke, M., Zhang, Y., Shah, P., Crabb, S., & Hansen, A. (2013). Traffic-related air pollution and health co-benefits of alternative transport in Adelaide, South Australia. *Env. Int*., *74*, 281–290.