Periodica Polytechnica Civil Engineering, 67(2), pp. 619–627, 2023

Estimating the Underreporting Rate of Injured Cyclists

Gábor Pauer[2*], Nóra Krizsik[1,2], Szilárd Szigeti[1,2]

[1] Department of Transport Technology and Economics, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics, Műegyetem rakpart 3, H-1111 Budapest, Hungary
[2] Road Safety Research Centre, KTI Institute for Transport Sciences Non-profit Ltd., Than Károly u. 3-5., H-1119 Budapest, Hungary
[*] Corresponding author, e-mail: pauer.gabor@kti.hu

Received: 21 October 2021, Accepted: 28 February 2023, Published online: 06 March 2023

Abstract

International data reports indicated that underreporting is a frequent phenomenon in the case of cyclist accidents. However, determining the exact volume of underreported accidents or injured cyclists is a difficult task. Our research focused on a comprehensive examination of the personal injury accident data of cyclists. The primary aim was to elaborate a model determining the proportion of injured cyclists 'missing' from the official accident database and to explore the possible causes and processes leading to that. With the use of various data sets (accident database, data from hospitals and survey of cyclists), the model can alleviate the distortion effects arising from conclusions reached using only one data source. To demonstrate applicability, the surveys and examinations were carried out in relation to Hungary. The results showed that only 8.3% of injured cyclists were reported in the official accident database. The majority (62.6%) of injured cyclists received neither police action, nor medical treatment. The volume of the underreporting varied depending on the injury outcome and type of accident.

Keywords

underreporting, cycling, road safety, road accident, accident analysis

1 Introduction and literature review

The underreporting in data collection is a common problem in many scientific fields, including road safety related research. Underreporting occurs when the counting of some event of interest is for some reason incomplete or there are errors in recording the outcomes. The sources of error in recording accident data may vary according to time and location. The unavailability of data is a crucial gap that affects accident analysis [1] and the design, construction and management of roads [2]. Data loss or incorrect data recording may occur from the accident scene, through hospital data collection, to subsequent data recording by the police [3].

Underreporting is not a unique problem in Hungary, it is a worldwide issue [4–6]. Methods for entering fatal accidents in the police database vary by country, so data loss may occur at multiple points. The following list summarizes the data collection methods of the European Transport Safety Council PIN countries in the case of fatal accidents [7]:

- The police officer checks with the hospital whether the injured person has died within 30 days following the incident (Bulgaria, Israel, and Italy).
- The hospital or other healthcare institution notifies the police (on a mandatory or voluntary basis) when the victim has died within 30 days following an accident (Czech Republic, Estonia, Croatia, Lithuania, Portugal, and Serbia).
- A combination of the two – the police may visit the victim, though the hospital may also notify the police of a death in an accident (Cyprus, United Kingdom, Greece, Ireland, Poland, Lithuania, Hungary, Malta, Sweden, Switzerland, and Slovenia).
- Police and hospitals are not obliged to exchange data in the event of a fatal accident. For such cases, a responsible organization has been designated to collect fatal accident data from various sources (Belgium, Finland, Germany, and Spain).

In international terms, the most frequently reported accident types are those in which a motorized vehicle was involved. For example, in the Netherlands, only 4% of serious accidents were reported as a single-vehicle accident [8], whereas an estimated 70% of bicycle accidents

Cite this article as: Pauer, G., Krizsik, N., Szigeti, S. "Estimating the Underreporting Rate of Injured Cyclists", Periodica Polytechnica Civil Engineering, 67(2), pp. 619–627, 2023. https://doi.org/10.3311/PPci.19429

were single-vehicle accidents [9]. Similar conclusions can be made based on the accident data of Hungary, where more than half of the recorded bicycle accidents were car-cyclist collisions [10]. However, accidents are not always reported even when a motorized vehicle was involved. A study surveying 1,087 cyclists from Brussels showed that only 7% of minor bicycle accidents were reported to the police, although 19% of accidents involved a motorized vehicle [11]. An extreme case was also reported in Israel in a questionnaire survey, where 80% of the respondents suffered injury in a bicycle accident, but none of them reported it to the police [12].

In Vancouver, only 20% of cycling accidents were found in the official reports. The vast majority of those were collisions between bicycles and motorized vehicles in which the cyclists suffered serious injuries [13].

There are examples of even fatal accidents missing from the official databases. In the Netherlands, based on joint data from the Dutch Statistical Office and the Ministry of Infrastructure and Water Management (Rijkswaterstaat) in 2017 [14], 87% of fatal road accidents were included in the police database. The police database on fatal accidents is incomplete in Greece and Poland, also. In Greece, between 1985 and 2015, the number of fatal accidents in hospital reports was 15–25% higher than the values in the police database. In Poland, a study revealed that, based on even the most modest estimates, 5% of fatal road accidents were unrecorded in the police database [7].

There are also examples of the incorrect data recording of accidents. In Finland, fatal cycling accidents were included in the police database, but were not recorded as bicycle accidents. The rate of incorrect recording of the vehicle type in fatal accidents was 23% in 2015, 27% in 2016, 48% in 2017 and 33% in 2018 [15].

A study was undertaken between 2014 and 2015 to get a comprehensive picture of the underreporting of bicycle accidents. The online questionnaire was completed by 8,655 cyclists from 30 countries. After cleaning up and filtering the data, a total of 7,015 responses from 17 countries were analyzed, including at least 100 valid responses per country [16]. Overall, on average, only 10.5% of cyclists reported their accident to the police. As part of that research, the proportion of accidents reported to police by type of accident was also examined. Accidents involving bicycles and motorized vehicles were the most frequently reported (24.9% of the cases). The reporting rate of accidents was 10.2% in the case of bicycle-bicycle collisions, 4.3% in bicycle-pedestrian collisions and less than 5% in the case of single-vehicle accidents.

Reliable accident data forms the basis of any road safety related analysis [17, 18] and transport design task [19]. Although, the safety of individual infrastructure designs can be revealed through on-site inspections [20], it is essential to have a comprehensive historical accident database related to cycling in order to improve safety focusing on the most frequent risk factors [21]. On the reviewed literature, it can be concluded that reliability of data is affected negatively by the underreporting of accidents. However, the determination of the volume of underreported accidents or injured cyclists is a difficult task. Most studies on this topic focus on only one data source, undertake a survey among cyclists or study the accident database of the police, or hospitals [22]. Linking of data sources is not common in most of the countries [23]. Accordingly, in order to alleviate the possible distortion effects or incompleteness of the data sources, our goal was to develop a complex model that combines those approaches. The paper presents the input parameters, the methods for data correction and the model outputs. The results provided by the model are identified in the process of accident data recording. The application of the model is presented through a case study on Hungary.

2 Elaboration of the model estimating the underreporting of accidents in the case of cyclists

In the case of a bicycle accident, data recording may differ depending on the actions taken at the scene. If a serious or fatal outcome or a collision with a motorized vehicle occurs [24], the probability of police action is higher. In theory, those cases are included in the police official accident database. In other cases, the police are not called to the scene of the accident, but the injured person is treated in a hospital or other healthcare institute. The reporting of those cases depends on the policies of the various countries. However, data loss may occur even in cases when reporting to the police is mandatory. In some cases, the injured person visits the hospital only some days after the accident, and there are cases when the injured person does not admit the fact of the traffic accident to the doctor. In addition, in the case of many cyclist accidents, especially cases involving only minor injury, neither police action nor medical treatment may occur. Those possible post-accident scenarios must be borne in mind when elaborating the modelling framework.

2.1 Determining the inputs of the model

To be able to combine a number of data sources and therefore improve the quality of the estimation, the following approaches can be taken in model development:

- The most effective method for estimating the total number of injured cyclists proved to be online surveys undertaken among cyclists as, according to previous research, a significant majority of road accidents involving cyclists are not reported to the police and, in many cases, medical treatment is not necessary. Thus, the initial data of the elaborated model were derived from that data source.
- An adequate basis for exploring the number of injured cyclists who received medical treatment related to their accident is provided by examination of the databases of the health sector (medical centers, hospitals), which are assumed to include more data than police databases (since presumably many cyclist accidents causing personal injury are not reported to police).
- The results obtained from a search of the international literature indicate that lost [11, 16] or inaccurately recorded data [15] must also be taken into account.

In Table 1, the primary input data of the developed model have been determined and summarized in accordance with the above considerations. Data can be separated according to whether any medical treatment and/or police action was/was not taken in relation to the accident (either at the accident scene or later). Since some of the data may be derived from both questionnaire and hospital databases, the possible sources of the data are also indicated. The lower index of the notation indicates what kinds of actions were taken in relation to the bicycle accident as explained in Table 1. The upper index refers to the data source (*i*: Questionnare/Health database, referred to as *Q* and *H* later).

Note that the ratio of injured cyclists receiving neither medical treatment nor police action related to their accident can only be estimated on the basis of the self-reporting of the involved cyclists through a questionnaire. By questioning the cyclists, the other introduced ratios can also be determined. However, as mentioned earlier, our research aimed to combine more data sources, in order to improve the quality of the estimations related to the volume of underreporting. The questionnaire is also suitable for controlling and verifying accident data collected from other sources. According to our investigations, the health sector databases contain a broad range of information related to the treatment of injured cyclists, including personal data, the date of the first treatment, the severity of the injury and also the fact of whether or not police action was taken in relation to the accident, or the police have been notified by the doctor. As a consequence of that, the share of injured cyclists receiving only medical treatment, or both medical treatment and police action, can be calculated on the basis of the health sector data as well. Please note, that the only data source for determining the ratio of cyclists receiving only police action (and no medical treatment) was the self-reporting based questionnaire as, usually, the official police accident database contains no information about medical treatment, especially when a cyclist only visits the doctor at a later time.

Table 1 Notation and source of input data

Notation	Data	Possible data sources (*i*)
N^i_{NPM}	share of injured cyclists who received no medical treatment and no police action related to their accidents	Questionnaire
X^i_{OP}	share of injured cyclists who received only police action related to their accidents (no medical treatment)	Questionnaire
X^i_{OM}	share of injured cyclists who received only medical treatment related to their accidents (no police action)	Questionnaire Health database
X^i_{PM}	share of injured cyclists who received both medical treatment and police action related to their accidents	Questionnaire Health database

2.2 Improving reliability by data adjustment

Data from the initial data source (questionnaire) can be adjusted by comparing the same data from other sources, thus increasing the reliability of the results provided by the model.

The first proposed adjustment (k_1) uses the relative proportion of X_{PM} and X_{OM} that can be calculated using both the self-reported data and the data of the health sector (if the database of the health sector also contains information about the police action). Using both databases, the relative proportion can be calculated within that of injured cyclists who received medical treatment, as follows in Eqs. (1) and (2).

$$k^i_{1,PM} = \frac{X^i_{PM}}{X^i_{PM} + X^i_{OM}}, \tag{1}$$

$$k^i_{1,OM} = \frac{X^i_{OM}}{X^i_{PM} + X^i_{OM}}, \tag{2}$$

where $i = Q$ or $i = H$ depending on the applied data source.

That way, the values of $k^Q_{1.PM}$, $k^Q_{1.OM}$, $k^H_{1,PM}$, $k^H_{1,OM}$ can be obtained. Based on the comparison of those values, the adjustment factors $k_{1.PM}$ and $k_{1.OM}$ must be determined by the analyst with consideration to their knowledge of the accuracy of the applied data source. If additional information unavailable, the mean of the corresponding values is recommended.

The developed adjustment factor must then be applied related to the investigated initial input data (Eqs. (3) and (4)).

$$X_{PM}^{k1} = k_{1,PM} \times \left(X_{PM}^{i} + X_{OM}^{i} \right) \quad (3)$$

$$X_{OM}^{k1} = k_{1,OM} \times \left(X_{PM}^{i} + X_{OM}^{i} \right) \quad (4)$$

The second proposed adjustment (k_2) has been developed to consider lost or inaccurately recorded police data. The theoretical value of the proportion of cyclist injuries reported in the official databases would be given by the sum of the ratio of cases covered by police action ($X_{OP} + X_{PM}^{k1}$). However, that can be adjusted on the basis of the comparison of the health sector database and the official accident database of the police. The value of factor X_{PM}^{k1} expresses the rate of the injured cyclists who have received both medical treatment and police action and can be found in the official police database. The value of can then be further specified by simply multiplying it by .

Although, the adjustment factor k_2 has been calculated on the basis of the data related to injured cyclists who received both medical and police actions, the same adjustment factor can be applied in the case of the ratio of injured cyclists receiving only police action (X_{OP}). The authors recommend that alignment if the analyst has no reason to assume that the data on accidents affected only by police action would be more or less likely to be included in the official accident database than cases covered by both police and medical action.

2.3 Elaborating the outputs of the model

Following the introduction of input data and the possible methods of increasing reliability by data adjustment, the model's outputs have been determined.

The share of injured cyclists (Y) in the official accident database can be obtained according to the formula of Eq. (5).

$$Y = k_2 \times \left(X_{OP} + X_{PM}^{k1} \right) \quad (5)$$

The other cases that have been excluded from the official accident database can be described by the following considerations. The proportion of injured cyclists who received police action related to their accidents but could not be identified in the official database (Z_1) has been expressed by Eq. (6).

$$Z_1 = (1 - k_2) \times \left(X_{OP} + X_{PM}^{k1} \right) \quad (6)$$

The share of injured cyclists receiving only medical treatment (and no police action) is denoted by Z_2, which is equal to X_{OM}^{k1}. The proportion of injured cyclists receiving no medical treatment and no police action is expressed by X_{NPM} and is denoted by Z_3.

The output data of the elaborated model cover all the possible scenarios related to the data of injured cyclists, as presented in Eq. (7).

$$Y + Z_1 + Z_2 + Z_3 = 100\% \quad (7)$$

The elaborated model is suitable for investigating the aggregated accident data, but can also be applied for appropriate subsets of it (e.g., to separately investigate the data of injured cyclists according to different injury outcomes, different accident types, etc.). For that purpose, only the input data need be determined accordingly.

2.4 Identification of the results provided by the model in the process of accident data recording

The process of accident data recording with the different scenarios has been summarized in Fig. 1, also identifying the places of the outputs of the model in the event chain.

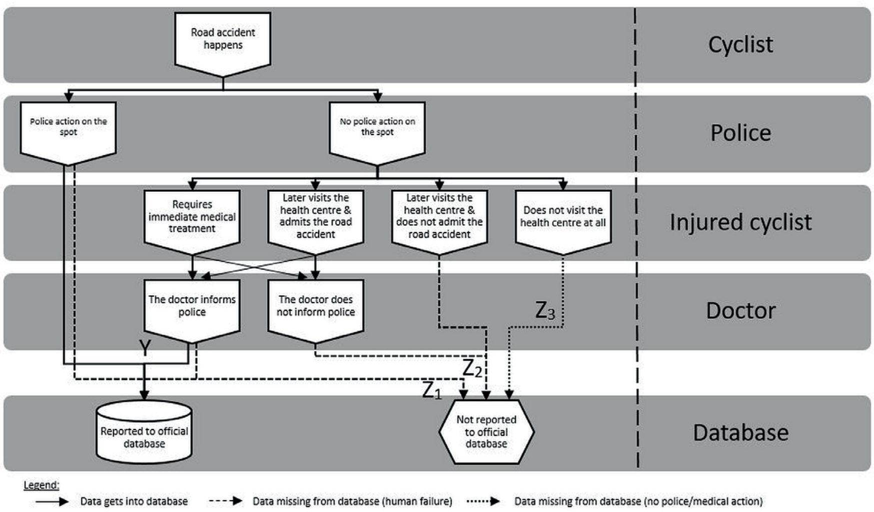

Fig. 1 Identifying the outputs of the model in the process of accident data recording

The routes of the output data can be well observed in the flowchart. In the official accident database, the represented proportion of injured cyclists is noted by *Y*. The database, in theory, could be complemented with data sets Z_1 and Z_2. To achieve that, it would be required, for one thing, to eliminate mistakes and deficiencies in the data recording and coding process (Z_1). Note that these cases do not include those that do not qualify as a road accident. However, self-reported initial data sets may include such cases, too.

Also, the databases of the health sector and the police should be connected, in order to facilitate data transfer between them. By the improvement of that process, the transferred data of injured cyclists could become more accurate and complete (bottom dashed line of Z_2). The most effective improvement in that regard would require significant effort and resources. It would require the police to collect and record the data of all injured cyclists receiving medical treatment (even in the case of minor injuries that might be treated by general practitioners days after the accident). Also, it would require the health sector to harmonize all data generated (from hospitals to general or specialist medical practitioners). To that end, the development of a common and standardized database, accessible to all doctors, is required.

Another part of the subset (upper dashed line) is formed by the cases that are missing from the official databases, since the injured person does not admit the road accident. By encouraging the elimination of those cases, the official database could also be expanded.

Z_3 indicates the share of the injured in minor bicycle accidents without any medical or police intervention. Those cases may only be detected by self-report, so would be practically impossible to include in the official database, although they represent a significant proportion of the accidents.

3 Application of the elaborated model – a case study on Hungary

To demonstrate the applicability of the elaborated model, the volume of underreporting in the case of the various groups of injured cyclists have been calculated based on data sets from Hungary. In 2019, which is the last complete year not affected seriously by the COVID-19 pandemic, the total number of road accidents involving personal injury was 16,627, of which 530 were fatal, 4,834 were serious and 11,263 were minor accidents. The ratio of accidents involving cyclists was 9.67%.

3.1 Presenting the inputs of the model

In the first step, data have been collected according to the combined data sources of the model. A questionnaire survey has been carried out among cyclists in 2020. In order to reach the target group, the survey was promoted and shared by the relevant public organizations and NGOs dealing with transport and cycling in Hungary (e.g., Hungarian Public Roads, KTI, Transport Science Association, Hungarian Cyclists' Club). The survey targeted only those who have suffered a cycling accident in the past 2 years.

The data of treated cyclists have been collected from three healthcare institutions. Using their own databases, the involved hospitals provided the basic data of the injured persons and their accidents. Finally, the official accident database was studied to compare and examine the accuracy of accident data obtained from the various sources. In accordance with the presented steps of our method, self-reported survey responses formed the basis of the analysis, which have then been adjusted by consideration of the data of the health sector and by the examination of the accident database.

3.1.1 Questionnaire survey

As mentioned, the questionnaire survey was carried out among the cyclists, who admitted to have at least one road accident between 2018 and 2019. The total number of relevant responses was 513. The majority of the respondents (80%) were between 26-55 years old and males were overrepresented (69.4%).

The survey aimed to reveal the number of accidents suffered by the cyclists and to determine the proportion of cases where different actions (medical treatment or police action) were taken. The results in that regard have been summarized in Fig. 2.

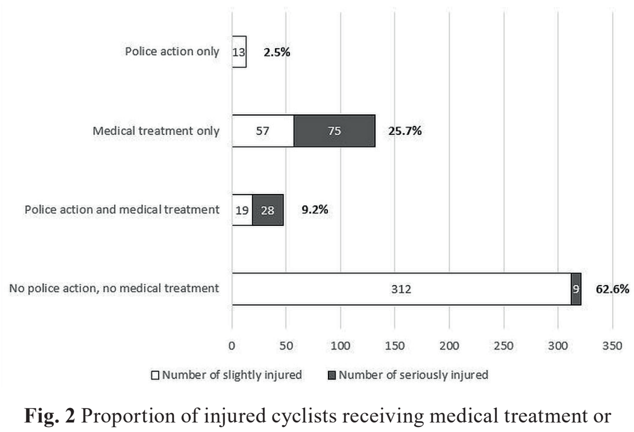

Fig. 2 Proportion of injured cyclists receiving medical treatment or police action related to their accidents (2018–2019)

In 62.6% of the total cases, the injured cyclists received no medical treatment or police action. However, that ratio was only 8% among the seriously injured. In every fourth case, medical treatment was provided but no police action was taken (at least, nor to the respondent's knowledge). Police action was taken in only 11.7% of the bicycle accidents. Those cases should be included in the official database.

3.1.2 Data of the healthcare institutions

Data collection involved three hospitals from different regions of the country. The hospitals provided data regarding all injured cyclists treated in their institution in the year 2018. The databases contained the primary information about the accidents: a short description of the circumstances, the data of the injured cyclist, the severity and type of the injury (ICD codes) and any involvement of alcohol; it was also indicated whether or not any police action was taken in relation to the accident. The main findings have been summarized in Table 2.

According to the provided data, a total of 1,716 injured cyclists were treated in the three hospitals. Only four riders died, 729 suffered serious injuries and 983 suffered minor injuries. Police action was registered by the doctors in only 218 cases (12.7%). Those cases have been further investigated by the authors, searching the official accident database. Of the 218 injured, 161 (73.9%) could be identified in the official database. That examination implies that many cases are missing from the database, even if police action occurred. Inaccurate and incomplete data recording also happens as some cases might be excluded from the road accident database due to the circumstances (e.g., the injury occurred on private property) though, in most cases, the description in the health sector database indicated the fact of a road accident.

3.2 Results of the model

Using the elaborated model, the volume of underreporting in the case of injured cyclists in Hungary has been estimated with consideration of the aggregated input data. Furthermore, detailed examinations have also been carried out considering subsets of the total data regarding the different injury outcomes and types of accidents. For that purpose, the necessary inputs have been derived from the collected databases. The type of the accident and injury outcome are consequences of different contributing factors and thus were presumed to have impacts also on the probability of the accident being reported [25]. The input data and the calculated correction factors have been summarized in Table 3.

The initial proportions (X_j) have been determined from the questionnaire. The first adjustment factors have been estimated on the basis of comparison of the ratios determined by the data of the questionnaire and the data of the hospitals. In Hungary, that medical reporting towards the police is not centralized indicates lower reliability of the data regarding police actions. Therefore, when determining $k_{1,j}$ factors, the data of the questionnaire were apportioned slightly greater weight. The factor k_2 was determined on the basis of the detailed investigation of the accident database in regard to the cases in which the data of the health sector noted the fact of police investigation.

Investigating the different scenarios, the probability of police action is much higher in the case of serious injury (25%) than in the case of minor injury (7.9%). That is also the case in relation to bicycle accidents involving a motorized vehicle, where police action occurred in 19.2% of the cases, compared to the 3.4% ratio in collisions with another bicycle or pedestrian. A ratio of 5.6% ratio is shown in the case of single-vehicle accidents. The proportion of injuries receiving neither medical nor police action was similar in the case of the different accident types, though was especially low (8%) in cases of serious injury. Considering the second correction factor, single-vehicle accidents were the most difficult to identify in the official accident database (41.7%).

The outputs have been calculated using the elaborated model and presented in Table 4.

Table 2 Data of the involved hospitals and the treated cyclists (2018)

Hospital	Location, service area	Population*	Slightly injured	Seriously injured	Fata-lities	Police action
North-Central Buda Centre, New St. John's Hospital and Clinic	The western region of Budapest and its agglomeration	approx. 800,000	484	320	2	95 (11.79%)
Szeged Traumatology Clinic	Szeged city and its agglomeration	approx. 350,000	398	270	1	94 (14.05%)
UP Clinical Centre	Pécs city and its agglomeration	approx. 240,000	101	139	1	29 (12.03%)

* The number of population in the service area of the hospitals were estimated by the experts of the referred hospitals

Table 3 Collected input data and calculated correction factors

	Slight injuries	Serious injuries	Single-vehicle accidents	Bicycle-bicycle and pedestrian-bicycle accidents	Motorized vehicle-bicycle accidents	Total
X_{OP}	0.032	0.0	0.0	0.0	0.054	0.025
X_{PM}	0.047	0.250	0.056	0.034	0.138	0.092
X_{OM}	0.142	0.670	0.299	0.322	0.204	0.257
X_{NPM}	0.778	0.080	0.645	0.644	0.604	0.626
$k^{O}_{1.PM}$	0.249	0.272	0.158	0.096	0.404	0.264
$k^{O}_{1.OM}$	0.751	0.728	0.842	0.904	0.596	0.736
$k^{H}_{1.PM}$	0.108	0.151	0.027	0.379	0.621	0.127
$k^{H}_{1.OM}$	0.802	0.690	0.973	0.621	0.379	0.752
$k_{1.PM}$*	0.2	0.3	0.1	0.25	0.5	0.25
$k_{1.OM}$*	0.8	0.7	0.9	0.75	0.5	0.75
X^{k1}_{PM}	0.038	0.276	0.036	0.089	0.171	0.087
X^{k1}_{OM}	0.152	0.644	0.320	0.267	0.171	0.262
k_2	0.717	0.763	0.417	0.760	0.809	0.739

* Estimated values based on the below-described considerations

Table 4 Outputs of the model related to the different scenarios

	Slight injury	Serious injury	Single-vehicle accident	Bicycle-bicycle and pedestrian-bicycle	Motorized vehicle-bicycle	Total
Y	5%	21.1%	1.5%	6.8%	18.2%	8.3%
Z_1	2%	6.5%	2%	2.1%	4.3%	2.9%
Z_2	15.2%	64.4%	32%	26.7%	17.1%	26.2%
Z_3	77.8%	8%	64.5%	64.4%	60.4%	62.6%

Based on the aggregated data, only 8.3% of cyclists with personal injury accidents were included in the official accident database. According to the outcomes, 5% of minor injuries and 21.2% of serious injuries were included. Regarding the types of accidents, underreporting was lowest (though still over 80%) in accidents between motorized vehicles and bicycles, while only 1.5% of the injured are included in the official database in the case of single-vehicle accidents. Those numbers highlight the significant underreporting of bicycle accidents in Hungary.

In order to illustrate the volume of the different (reported and non-reported) cases, an energy flow type chart based on the aggregated data is presented in Fig 3. Missing or incorrectly recorded data are indicated by darker shades, while the correctly recorded data are indicated by lighter shades.

As shown in the chart, only the data of one in twelve injured cyclists are included in the official accident database. However, as described earlier, the data of another 29.1% (2.9% + 26.2%) of the injured could be included in the database with the development of data recording and collecting processes. To express it numerically: in addition to the 3,001 injured cyclists reported in the official accident database in 2019, it is assumed that another 33,156 cyclists were injured in road accidents during the same year.

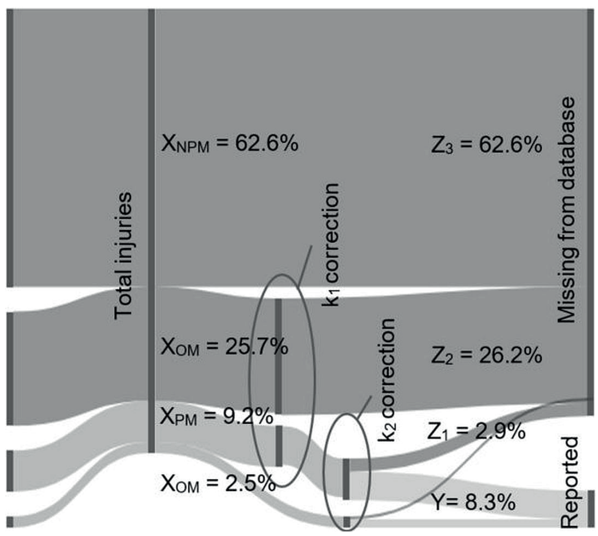

Fig. 3 Ratio of input and output data of the model regarding the total number of bicycle accidents

Comparing the official accident database with the health sector data, we are able to conclude that there is a significant difference between those two data sets. The volume of treated cyclists was around three times higher in the investigated hospitals than in the accident database. For example, the Szeged Traumatology Clinic (responsible for the treatment of about 350,000 people around the city of Szeged) treated 669 cyclists in 2018, while the reported number of injured cyclists in the whole of the surrounding Csongrád-Csanád County (with around 400,000 inhabitants [26], including the city of Szeged) was only 239 in the same year.

4 Conclusions

The underreporting of road accidents is a common problem in all countries. The volume of missing data can be especially high in the case of accidents involving cyclists, especially in accidents not involving motorized vehicles. According to the most comprehensive studies, the average ratio of injured cyclists reported in the police accident databases in Europe is around 10%.

This article presents a model to facilitate the estimation of those. Its main advantage is the combination of different data sources (e.g., questionnaire, health sector and police databases), thus enhancing the reliability of the results. However, it can be challenging to collect all the necessary input data. The outputs of the model were identified in the event chain of the accident data recording process.

The applicability of the model is demonstrated in the case study of Hungary. According to our results, only 8.3% of cyclists with personal injury were reported in the country's official accident database. In 2.9% of the cases, police action was taken at the scene, though there was no trace of the accident in the database. In those cases, the problem may have been incomplete or faulty police data recording. Another 26.2% of cyclists suffering accidents received only medical treatment and were not affected by any police action. The remaining and, at the same time, the most significant proportion (62.6%) were injured cyclists receiving neither medical treatment nor police action. The vast majority of those cyclists suffered only minor injury or were involved in single-vehicle accidents.

The elaborated model and the results of our study, facilitate the determination of multiplication factors that can be used to estimate the total number of injured cyclists based on the data recorded in the official accident databases. In addition, the critical points and main problems of the accident recording process are identified and potential opportunities for improvement indicated. The presented results can be used during accident data analysis for the detailed investigation of the safety of cyclists. The main limitation of the model is however that it is only able to support quantitative analysis by estimating the number of underreported accidents and/or injuries. Detailed characteristics of the accidents cannot be obtained from the databases of the health sector. Thus, the qualitative characteristics of accidents could only be determined if the number of cases missing from the database were reduced. Therefore, an important task of future work is to eliminate the deficiencies related to the data recording. For this, on the one hand, an important step is to connect the databases of different sectors, or to designate an organization that performs and coordinates data collection from several sources, as pointed out by our research. On the other hand, it would also be important to gain information about injured cyclists receiving neither police nor medical treatment. In the exploration of such cases, for example, the appropriate development of information services [27] can provide an opportunity.

The authors' future work will focus on the examination of road users' attitude towards voluntary data provision opportunities and the development of the related technical conditions.

References

[1] Mannering, F. L., Bhat, C. R. "Analytic methods in accident research: Methodological frontier and future directions", Analytic Methods in Accident Research, 1, pp. 1–22, 2014. https://doi.org/10.1016/j.amar.2013.09.001

[2] Castro, M., García-Espona, A., Iglesias, L. "Terrain Model Resolution Effect on Sight Distance on Roads", Periodica Polytechnica Civil Engineering, 59(2), pp. 165–172, 2015. https://doi.org/10.3311/PPci.7658

[3] James, H. F. "Under-reporting of road traffic accidents", Traffic Engineering & Control, 32(12), pp. 574–583, 1991. [online] Available at: https://trid.trb.org/view/364551

[4] Amoros, E., Martin, J.-L., Laumon, B. "Under-reporting of road crash casualties in France", Accident Analysis & Prevention, 38(4), pp. 627–635, 2006. https://doi.org/10.1016/j.aap.2005.11.006

[5] Samuel, J. C., Sankhulani, E., Qureshi, J. S., Baloyi, P., Thupi, C., Lee, C. N., Miller, W. C., Cairns, B. A., Charles, A. G. "Under-Reporting of Road Traffic Mortality in Developing Countries: Application of a Capture-Recapture Statistical Model to Refine Mortality Estimates", PLoS ONE, 7(2), e31091, 2012. https://doi.org/10.1371/journal.pone.0031091

[6] Nakahara, S., Wakai, S. "Underreporting of traffic injuries involving children in Japan", Injury Prevention, 7, pp. 242–244, 2001. https://doi.org/10.1136/ip.7.3.242

[7] Adminaite, D., Jost, G., Stipdonk, H., Ward, H. "An overview of road death data collection in the EU", European Transport Safety Council, Brussels, Belgium, PIN Flash Report 35, 2018. [online] Available at: https://etsc.eu/an-overview-of-road-death-data-collection-in-the-eu-pin-flash-35/

[8] Bos, N., Reurings, M., Derriks, H. "Correction for underreporting of road traffic casualties in the Netherlands; relevance and method", [pdf] In: 4th IRTAD Conference, Seoul, Korea, 2009, pp. 115–128. Available at: https://www.itf-oecd.org/sites/default/files/docs/4-bos.pdf

[9] van Kampen, L. T. B. "Gewonde fietsers in het ziekenhuis; Een analyse van ongevallen-en letselgegevens", (Injured cyclists in hospital, An analysis of accident and injury data), SWOV, Leidschendam, Netherlands, R-2007-9, 2007. [online] Available at: https://swov.nl/nl/publicatie/gewonde-fietsers-het-ziekenhuis (in Dutch)

[10] Glász, A., Juhász, J. "Car-pedestrian and car-cyclist accidents in Hungary", Transportation Research Procedia, 24, pp. 474–481, 2017. https://doi.org/10.1016/j.trpro.2017.05.085

[11] de Geus, B., Vandenbulcke, G., Panis, L. I., Thomas, I., Degraeuwe, B., Cumps, E., Aertsens, J., Torfs, R., Meeusen, R. "A Prospective cohort study involving commuter cyclists in Belgium", Accident Analysis & Prevention, 45, pp. 683–693, 2012. https://doi.org/10.1016/j.aap.2011.09.045

[12] Haworth, N., Shinar, D., Oron-Gilad, T. "Developing an international survey of bicycle and helmet usage", In: Shinar, D. (Ed.) Proceedings of the 3rd International Cycling Safety Conference (ICSC2014), 3rd International Cycling Safety Conference Organizing Committee, Gothenburg, Sweden, 2014, pp. 1–12. [online] Available at: https://eprints.qut.edu.au/80806/

[13] Winters, M., Branion-Calles, M. "Cycling safety: Quantifying the under reporting of cycling incidents in Vancouver, British Columbia", Journal of Transport & Health, 7(A), pp. 48–53, 2017. https://doi.org/10.1016/j.jth.2017.02.010

[14] ITF "Road Safety Annual Report 2018", International Transport Forum, OECD Publishing, Paris, France, 2018. https://doi.org/10.1787/1c884dcb-en

[15] Adminaité-Fodor, D., Jost, G. "How safe is walking and cycling in Europe?", European Transport Safety Council, Brussels, Belgium, - PIN Flash Report 38, 2020. [online] Available at: https://etsc.eu/how-safe-is-walking-and-cycling-in-europe-pin-flash-38/

[16] Shinar, D., Balero-Mora, P., van Strijp-Houtenbos, M., Haworth, N., Schramm, A., …, Tzamalouka, G. "Under-reporting bicycle accidents to police in the COST TU1101 international survey: Cross-country comparisons and associated factors", Accident Analysis and Prevention, 110, pp. 177–186, 2018. https://doi.org/10.1016/j.aap.2017.09.018

[17] Sipos, T. "Spatial Statistical Analysis of the Traffic Accidents", Periodica Polytechnica Transportation Engineering, 45(2), pp. 101–105, 2017. https://doi.org/10.3311/PPtr.9895

[18] Ghadi, M., Török, Á., Tánczos, K. "Integration of Probability and Clustering Based Approaches in the Field of Black Spot Identification", Periodica Polytechnica Civil Engineering, 63(1), pp. 46–52, 2019. https://doi.org/10.3311/PPci.11753

[19] Miskolczi, M., Földes, D., Munkácsy, A., Jászberényi, M. "Urban mobility scenarios until the 2030s", Sustainable Cities and Society, 72, 103029, 2021. https://doi.org/10.1016/j.scs.2021.103029

[20] Makó, E., Hóz, E., Miletics, D., Borsos, A. "Road Safety Audit experiences related to cycling infrastructure", [pdf] In: 6th Annual International Cycling Safety Conference, Davis, California, USA, 2017. Paper: 60, 3 p. Available at: https://icsc2017.figshare.com/articles/journal_contribution/Road_Safety_Audit_experiences_related_to_cycling_infrastructure/5405089

[21] Jaber, A., Juhász, J., Csonka, B. "An Analysis of Factors Affecting the Severity of Cycling Crashes Using Binary Regression Model", Sustainability, 13(12), 6945, 2021. https://doi.org/10.3390/su13126945

[22] Yannis, G., Papadimitrou, E., Chaziris, A., Broughton, J. "Modeling road accident injury under-reporting in Europe", European Transport Research Review, 6, pp. 425–438, 2014. https://doi.org/10.1007/s12544-014-0142-4

[23] Yannis, G., Laiou, A. "Current Road Safety Trends in Greece", Periodica Polytechnica Transportation Engineering, 42(2), pp. 145–151, 2014. https://doi.org/10.3311/PPtr.7511

[24] Levulyte, L., Baranyai, D., Török, Á., Sokolovskij, E. "Bicycles' role in road accidents, A review of literature", Transport and Telecommunication, 17(2), pp. 122–127, 2016. https://doi.org/10.1515/ttj-2016-0011

[25] Wang, Y.-G., Peng, H., Xiang, W.-S., Pei, Y.-L. "Examining injury severity in left turning crashes at intersections", Periodica Polytechnica Civil Engineering, 55(2), pp. 191–197, 2011. https://doi.org/10.3311/pp.ci.2011-2.11

[26] Hungarian Central Statistical Office (KSH) "Administrative Place Name Book of Hungary, Gazetteer of Hungary – 1st January, 2015", [pdf] Hungarian Central Statistical Office, Budapest, Hungary, 2015. ISSN 1217-2952. Available at: https://www.ksh.hu/docs/hun/hnk/hnk_2015.pdf

[27] Földes, D., Csiszár, C. "Personalised information services for bikers", International Journal of Applied Management Science (IJAMS), 10(1), pp. 3–25, 2018. https://doi.org/10.1504/IJAMS.2018.089944