Sri Lanka Society of Transport and Logistics

Journal of South Asian Logistics and Transport
Volume 2 Issue 2 September 2022
JSALT 2.2.4: pp. 79-98
DOI: http://doi.org/10.4038/jsalt.v2i2.50

REPORTING BEHAVIOUR OF PEDESTRIAN INVOLVED ACCIDENTS IN SRI LANKA

A G H J Edirisinghe [a]*, E M P Ekanayake [b] and B D S Madushani [b]

[a] Dept. of Civil Engineering, University of Peradeniya, Sri Lanka.

[b] Dept. of Mathematical Sciences, Wayamba University of Sri Lanka.

* Correspondence should be addressed to *jayalath@eng.pdn.ac.lk*

ABSTRACT

According to the World Health Organisation, the majority of victims of road traffic accidents are vulnerable road users, including pedestrians. However, under-reporting of such accidents leads to inadequate availability of reliable data in this area of research.

The present study was an attempt to investigate the severity of this problem of under-reporting of accidents involving pedestrians. Primary data were collected for the purpose, and statistical techniques such as univariate and bivariate analyses, Chi-square testing, and Binary Logistic Regression were used to examine the association of several diverse factors including demographic features, severity of injury, lost productivity, and receipt of compensation, with accident reporting behaviour.

The results revealed that the most common accidents, nearly 30% of all, involved motorcycles, while accidents due to slipping on roads or pavements and roadside structures together accounted for 33% of all reported accidents. These, according to this study, appeared to largely go under-reported, precluding further action.

The study found the presence of a strong relationship between reporting behaviour and the nature and severity of injuries sustained in the accidents. The lowest chance of reporting an accident was found in regard to slipping on a pavement while the injuries treated without visiting a hospital had the lowest likelihood of being reported to the Police.

Keywords: *Reporting behaviour, Accident underreporting, Vulnerable users of roads, Pedestrians, Accident severity*

1. INTRODUCTION

According to the Global Status Report on road safety (2018) by the World Health Organisation (WHO) [1], approximately 1.35 million people die every year due to road traffic accidents and nearly 50% of traffic accident-related deaths impact the vulnerable road user group comprising of pedestrians, cyclists and motorcyclists. Further, it also reports that males are more likely to be victims of road traffic crashes than females. As walking is an important activity in maintaining the physical fitness of individuals, as well as to complete the first and the last miles of a trip, more attention of responsible infrastructure development agencies is required to the safety of pedestrians.

A comprehensive study on safety of pedestrians on roads is not possible due to various issues pertaining to records on pedestrian involved accidents. This subject, though was addressed by many research groups in other countries, not many studies could be found in the Sri Lankan context. Therefore, the current research was undertaken to examine the nature of pedestrian involved accidents and accuracy of reporting to police in Sri Lanka and to propose appropriate methods to improve the situation.

As pedestrian crashes have been identified as a serious issue, many researchers have conducted pedestrian crash analyses in detail [2], [3], [4], [5], [6], [7]. There were a number of studies on the characteristics of pedestrian involved accidents. According to a study carried out on multi-level hot zone identification for pedestrian safety in Florida, United States of America (USA) [8], nearly 90% of accidents involving pedestrians had happened within 14 miles of the residence of the victim and approximately 50% of pedestrian involved accidents had occurred within residential zones. One of the major challenges faced in carrying out scientific research related to road traffic accidents is the non-availability of a comprehensive databases. Under-reporting of traffic accidents is a well-known issue in this regard in many countries. In late 1980s, the average level of error in the reporting of minor injury accidents among countries was 50% [9], which increased to 75% by late 1990s [10]. In Australia, the level of error was reported to be 69% [11]. Errors in reporting traffic accidents were found to be caused by lack of motivation, knowledge, and benefit, and due to others' influence too [12].

Furthermore, the severity of an injury has been found to have a direct relationship with under-reporting [13], and under-reporting varied within 93% - 98% for slight injuries, 32.5% - 96% for non-fatal injuries, and 0.5% - 89.5% for fatal injuries in middle income countries. Reluctance to summon the police [14] especially by the pedestrians [15] and younger road users [6] contribute heavily to under-reporting.

Lack of awareness and the fear of being caught by the parents are reasons for the teenagers to keep away from reporting accidents [16]. According to a model developed using a Poisson process for the under-reporting problem in underprivileged areas with low income and minority populations in Illinois, USA, it was found that pedestrian crashes were contributed by the absence of separate paths for walking and vehicular traffic, income disparities, and the presence of children [4]. A study conducted in Denmark revealed that under-reporting level of fatally injured accident data for pedestrian, cyclists, moped, motorcyclist, car, bus and other victims of road crashes in Denmark was 16.42% [12]. According to that study, the pedestrian injury rates were associated with population density, age composition of the local population, their level of unemployment, gender composition and level of education.

According to a study conducted by Mako and Szakonyi, the knowledge and attitude of pedestrians have a direct relationship with accidents [17]. That study further revealed that 44% of pedestrian engaged accidents were mainly due to the fault of the pedestrians and 27% occurred due to the speeding of drivers. In retrospect, the common causes behind errors in accident data reporting were cited as weaknesses in the policing system, distance to the nearest police station, victims not following the "within 30 days" rule to report, their ignorance of the law, limited legal requirements for reporting an accident, intentional non-compliance by road users, reluctance of people to summon the police and file a report, settlement (agreement between the parties), absence of serious injuries, hit and run cases, and political vs. militia conflicts [13]. Though there are a number of studies/ research on pedestrian accidents, there is a need for new knowledge in relation to under-reporting accidents in which only pedestrians are involved such as slipping on roads or on pedestrian walkways as they cause untold hardships to the pedestrians.

In Sri Lanka, the percentage of deaths of vulnerable road users who met with road traffic accidents was as high as 78% in the years 2017, 2018 and 2019. Even though there has not been any systematic study to estimate the severity of accident under-reporting in Sri Lanka, there is evidence to perceive increased incidence of traffic accident under-reporting from the year 2003 onwards. The ratio of fatal accidents to property-damage-only accidents dropped from 20.2 to 5.3 and the ratio of fatal accidents to minor injury accidents dropped from 7.1 to 5.1 during the period from 2003 to 2008. As there was no significant change in the annual increment of fatal accidents and grievous accidents, it could be observed that accident under-reporting has become increasingly common since the year 2003.

There could be numerous reasons for accident under-reporting in Sri Lanka. One such cause behind could be the new insurance schemes introduced by various parties facilitating claims without a police report, even though, according to the law of the country, all traffic accidents should be reported to police [17]. Lack of knowledge among accident victims is another major reason for accident under-reporting. If an accident does not involve another party, many pedestrians tend not to report such incidents assuming that the fault is with the pedestrians themselves. It is noteworthy that without such vital information, organisations responsible for road maintenance become unable to take corrective measures to minimise future incidence of accidents.

In the light of the above, this study was carried out to investigate the rate of accident under-reporting with respect to pedestrian involved accidents in Sri Lanka, and the significance of associated factors, namely, gender, employment, severity, days lost, and the nature of the accident, on such under-reporting. The study was designed based on primary data collected by students of the Faculty of Engineering, University of Peradeniya, in an island-wide structured questionnaire survey.

In Section 2, the nature of data and the criteria behind filtering data for the study, coding, and the variables are described in detail because comprehensive primary data were collected following a time-consuming process. Further, the statistical techniques used for data analysis, namely, the univariate and bivariate analyses, Pearson's Chi-square test, and Binary Logistic Regression, are also presented in this section. Section 3 elaborates the results obtained from the above statistical methods and highlighting the reasons for under-reporting of pedestrian involved accidents with possible follow-up action in the discussion. The conclusions are summarised in Section 4.

2. METHODOLOGY

2.1. Data Collection

Primary data were collected through a comprehensive survey using the Convenience Sampling Method covering the whole country and gathering information from 1242 respondents, which included four demographic characteristics of the population: age, gender, highest educational qualification, and profession. Questionnaire for the study was developed using the knowledge and information gathered from similar preceding studies [12], [16]. The survey was administered using a google form; however, the respondents were interviewed in person by a large group of students of the Faculty of Engineering at the University of Peradeniya, Sri Lanka who came from all over the country. Of these respondents, only those who had met with at least one accident were filtered out to make a preliminary sample of size 422. As some of the respondents had experienced more than one accident and as separate questionnaires were administered to them, this sample comprised of 804 different accidents to carry out the statistical analyses. In addition to the demographic details, the questionnaire sought elaborative information on the nature of the accident, its seriousness, lost productivity, compensation received, complaint made as well as whether the accident was reported or not. Accordingly, all those aspects were represented by the variable names of Gender, Age, Profession, Education, Nature, Severity, Loss, Compensation, Complaint and Report.

2.2. Missing Values

Dealing with missing values is an important preliminary step in statistical analysis based on raw data. Some data may not be available, either due to the refusal of a respondent to answer or the irrelevance of a particular question to the person. The percentage of missing values was examined in terms of the number (and the percentage) of variables, cases (respondents), and the values. All ten variables (100%) had some missing values, 106 (13.18%) cases had at least one missing value, and 214 (2.66%) of the values were missing. The questionnaire included the nine independent variables listed in Table 1, which provided all the mean variables having at least 0.01% of missing values arranged according to their descending order. The tenth variable, Report, (dependent and binary) was introduced to identify whether the accident was reported or not for the purpose of statistical analysis.

Table 1: Variable Summary

	Missing		Valid N	Mean	Standard. Deviation
	N	Percent			
Loss	44	5.5%	760		
Age	30	3.7%	774	30.74	12.294
Complaint	29	3.6%	775		
Compensation	26	3.2%	778		
Education	17	2.1%	787		
Profession	13	1.6%	791		
Severity	9	1.1%	795		
Gender	9	1.1%	795		
Nature	8	1.0%	796		

As the overall percentage of missing values (2.66%) was below the cut off mark (5%) adopted for removing the missing values in a statistical analysis, the examination was extended to check whether the missing values were randomly distributed. The distribution of missing values for the ten variables indicated no particular pattern among the missing values of different variables to create any bias in them, which confirmed their random occurrence. Further, it was equally important to understand the most frequent patterns of missing values, for which the percentages of cases (respondents) against the decreasing order of the ten most frequent patterns were examined. This revealed that over 80% of the cases had no missing values and the missing values on Loss was less than 5%. As the overall percentage of missing values was under 5% and distributed randomly, listwise deletion method [18] could be used to remove them and to make a revised representative sample of 698 cases, used for all analyses of this study. The distribution of *reporting behaviour* in the modified sample is shown in Table 2, which indicates that 80.5% of the respondents had not reported accidents.

Table 2: Distribution of *reporting behaviour* in the modified sample

		Frequency	Percentage	Valid percentage	Cumulative percentage
Valid	No	562	80.5	80.5	80.5
	Yes	136	19.5	19.5	100.0
	Total	698	100.0	100.0	

2.3. Method of Analysis

As the main objective of the study was to describe the accident reporting behaviour, the information on reporting an accident was treated as the dependent variable while all the others were considered as independent variables. A univariate analysis was carried out for the categorical variables of Nature, Severity, and the Complaint itself while bivariate analyses were conducted to study the reporting behaviour with each of Gender, Profession, Age, Educational qualifications, Nature of the accident, Severity of the accident, Loss of productivity, and the Compensation received. Further, the reporting behaviour was considered with the nature of the accident, its severity, and the lost productivity. Also, the same was studied with the nature of the accident and the profession involved.

The Pearson's Chi-square Test was used to identify whether there is any significant association between the dependent variable of reporting behaviour and the categorical independent variables at 5% level of significance using the following hypothesis:

H_0: There is no association between the independent variable considered and the reporting behaviour.

H_1: There is an association between them.

The null hypothesis would be rejected if the p-value was found to be less than 0.05 [19]. Further, the Cramer's V Test was used to identify the level of association as weak [< 0.2], moderate [0.2, 0.3], or strong [> 0.3]. The Pearson's correlation coefficient was used instead of the Chi-square test in assessing the relation between the continuous variable of age and the reporting behaviour.

Although bivariate analysis (Chi-square test and correlation) is useful in identifying the association of independent variables with the accident reporting behaviour, there is a possibility that variables found as insignificant in the bivariate analysis could emerge non-negligible in an advanced complex multivariate analysis. In order to explore this situation statistically, all independent variables were used to fit the binary logistic model.

Binary Logistic Regression (BLR) is a versatile analysis tool of regression that can extract relationship between a categorical dependent variable and a set of independent variables. It is used when the dependent variable is of Bernoulli type, i.e., assuming only two values such as 0 and 1 or success and failure. Also, the BLR can deal with qualitatively different values that cannot be tested by other widely used statistical models like Multiple Linear Regression for binary response data [20].

The simple binary logistic model has the form,

$$\text{logit } P(x) = ln\frac{P(x)}{1-P(x)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \quad (1)$$

where $X_1, X_2, \ldots, X_n$ are the explanatory (independent) variables, $P(x)$ is the probability of the outcome of interest (success), and β_0 is the intercept term.

Further, the odds ratio was calculated as follows, to measure the impact of a sub-category of an independent variable on the accident reporting behaviour, relative to that of the reference category of the same variable.

Odds, the ratio of probability of success to the probability of failure $= \frac{P(x)}{1-P(x)}$ (2)

where $P(x)$ is the probability of success (reporting an accident).

Accordingly,

if Odds1 $= \frac{P_1(x)}{1-P_1(x)}$ and Odds2 $= \frac{P_2(x)}{1-P_2(x)}$, then,

$$\text{Odds Ratio} = \frac{P_1(x)/1-P_1(x)}{P_2(x)/1-P_2(x)} \quad (3)$$

where Odds1 is the Odds of the sub-category and Odds2 is that of the reference category.

All possible answers to each question posed to the respondents were separated into one reference category while retaining the others as the dummy variables for the analysis in the SPSS package. Accordingly, female (Gender), government (Profession), postgraduate (Education), accident with a vehicle (Nature), full-cost/ adequate amount (Compensation), admitted to hospital (Severity), and one or two days (Loss) were chosen as the reference category of each categorical variable indicated within parenthesis while considering the other responses as its dummy variables.

The final model was decided after a few steps by carrying out the BLR analysis with forward conditional method within the SPSS package and the Hosmer – Lemeshow Test was used to find the goodness of fit of the model with the following hypothesis [19].

H_0: Model is a good fit for data

vs.

H_1: Model is not a good fit for data

The Hosmer- Lemeshow Test produces a Chi–square value and the p–value, which is compared at 5% level of significance to determine the goodness of fit. Further, the Cox & Snell R square and Nagelkerke R Square values generated by the SPSS explain the percentage range of variation in the dependent variable (accident reporting behaviour) explained by the fitted model.

Having considered all the variables at the outset, the final model (regression equation) retained only the most significant variables whose significance was quantified in terms of the p-values associated with each predictor while indicating the Odds Ratio for each sub-category, which was the probability of reporting an accident relative to the reference category.

Finally, hypothesis testing was conducted to check the significance of each predictor in the regression equation. The results obtained through this process are explained in the next section.

3. RESULTS AND DISCUSSION

Univariate analysis was carried out for the categorical variables of Nature, Severity, and Complaint and the continuous variable of age, depicted in Figure 1.

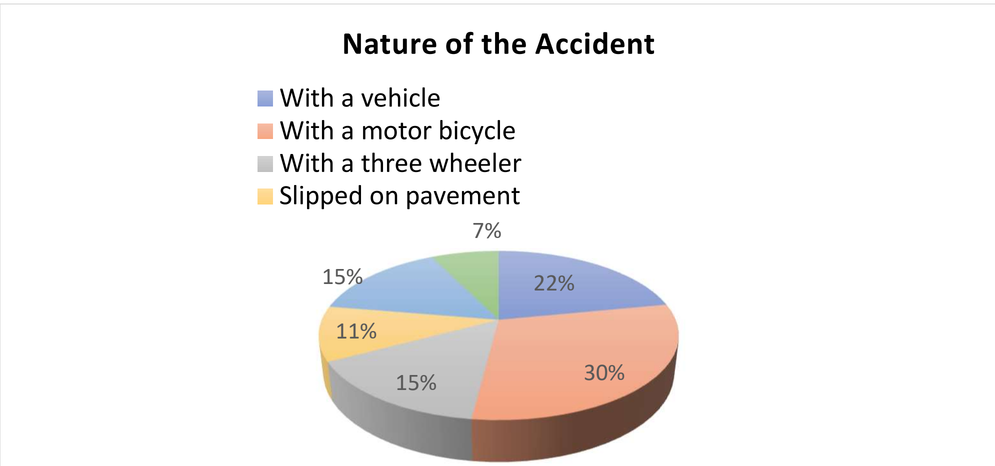

Figure 1(a): Percentage Shares of the Accidents by Nature

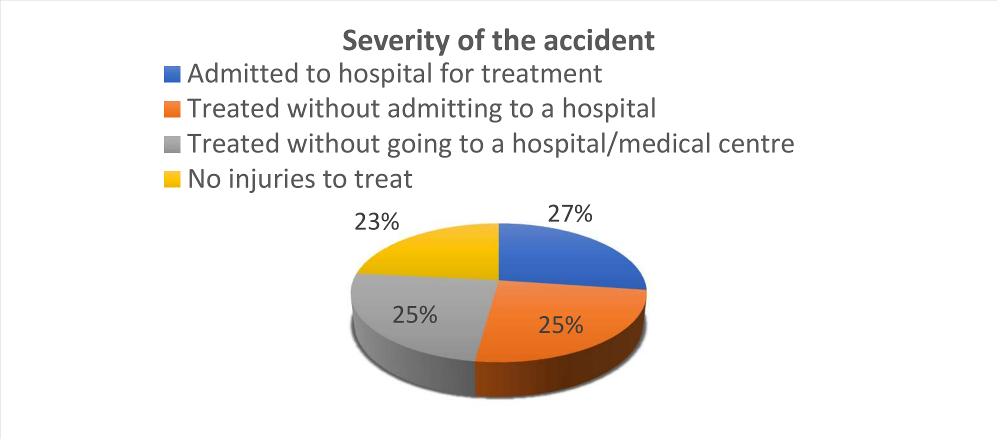

Figure 1(b): Percentage Shares of recorded Accidents by Severity

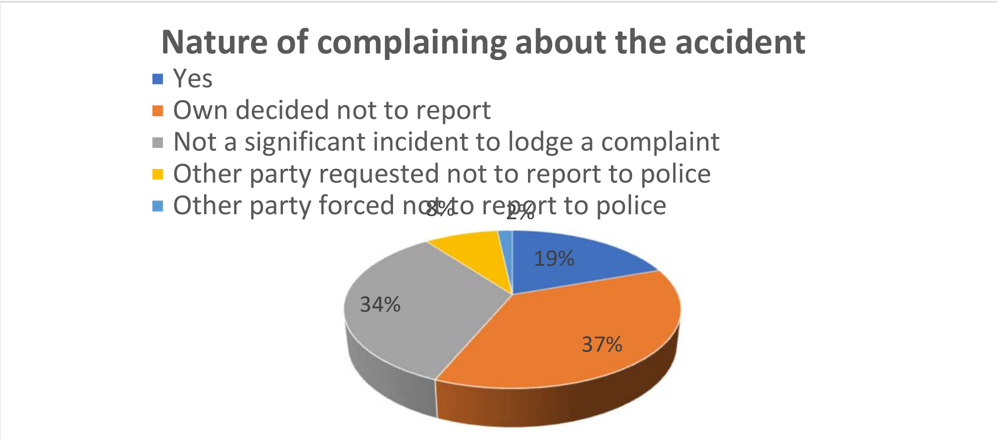

Figure 1(c): Percentage shares of Accidents by the Status of Complaining

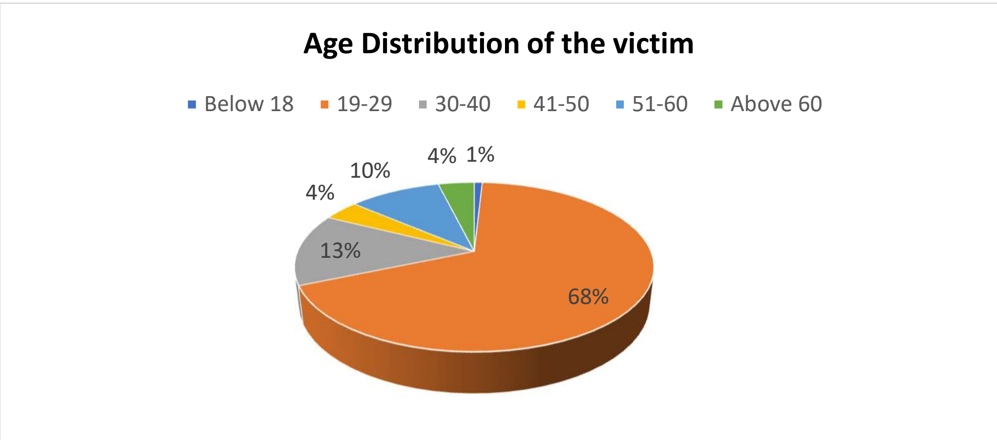

Figure 1(d): Percentage Share of Recorded Accidents by the Age of the Victims

According to Figure 1(a), it could be noticed that the highest percentage of accidents, recorded in the survey, occurred with a motor bicycle while the least was with a roadside structure/tree/pole. Accidents with a three-wheeler and due to slipping on the road occurred in equal ratios. Figure 1(b) shows that 27% of those facing accidents were admitted to a hospital for treatment while 25% were treated without being admitted to a hospital. One quarter of them was treated without visiting a hospital/ medical centre and nearly one quarter had escaped without injuries. In Figure 1(c), it can be seen that the percentage of complaints lodged was below 20% and more than one-third of the cases were underreported due to the respondents' own decision while another one-third decided not to lodge a complaint due to the insignificance of the incident. Figure1(d) makes it clear that most respondents were in the age group of 19-29 years followed by those in the age groups 30-40 and 51-60.

Bivariate analysis was performed on the independent variables of Gender, Profession, Age, Educational qualifications, Nature of the accident, its Severity, Loss of productivity, and Compensation, with the dependent variable of reporting behaviour. This analysis revealed the following major characteristics of the data. The highest percentage of females (39.6%) did not report accidents due to their insignificance followed by a similar percentage (37.10%) voluntarily deciding against reporting. Comparatively, the highest percentage of males (36.7%) stayed away from reporting on their own and 30.3% due to the insignificance of the incident. Higher percentage of females treated the injuries without being admitted to a hospital, but that of males got admitted to a hospital for treatment. In each employment category, the majority of underreporting of accidents were due to own decision and insignificance of the incident. In each complaint group, students were the highest percentage, which may be due to them belonging to the largest category among all respondents. They also accounted for the highest underreported cases.

Half of those in the age group below 18 years did not report accidents due to their insignificance while a majority of those in 30 - 40 year age group did so voluntarily. However, 51 - 60 age group was found more likely to report accidents compared to the other age groups, which could be due to their experience and concerns on the consequences of not reporting. If the accident involved a four-wheeled vehicle, a motor cycle, or a three-wheeler, a large proportion of respondents decided on their own not to lodge a complaint while the accidents due to slipping on the pavement / road or with a roadside structure / tree / pole were underreported due to their lower significance caused without involvement of another person.

Figure 2 depicts the distribution of the reporting behaviour within each category of accident. Accordingly, the largest proportion of reported cases was accidents with a four-wheeled vehicle and the least was slipping on the pavement. The largest share of respondents had faced accidents with a motor cycle, which were largely underreported.

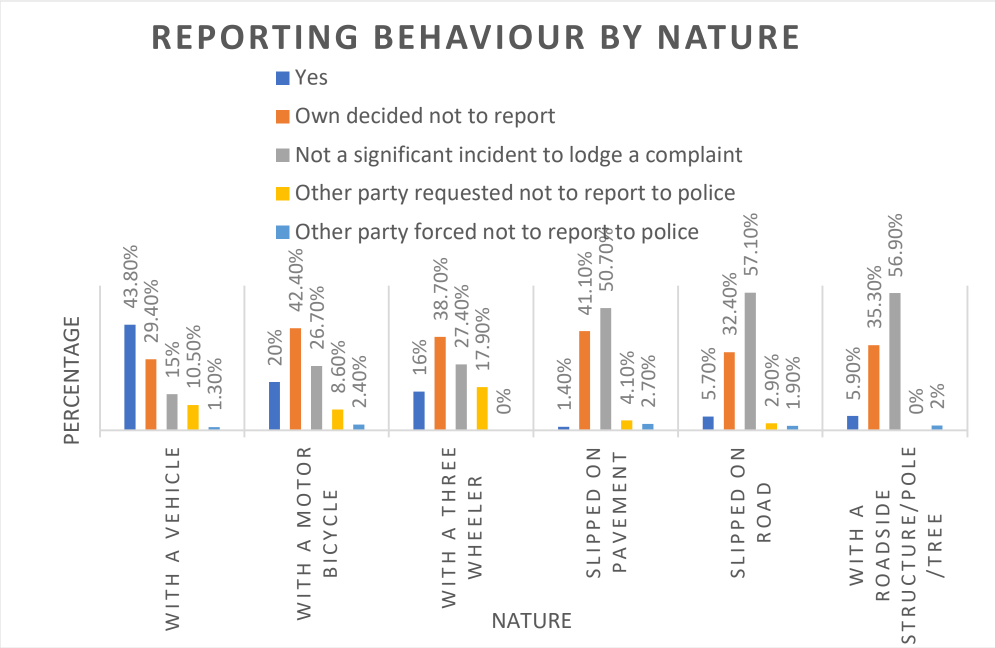

Figure 2: Reporting Behaviour by Nature

Furthermore, the following results could be obtained from the descriptive analysis.

A majority of those who got admitted to a hospital (51.1%) for treatment had lodged a complaint and vice versa, i.e., a majority of those who had reported the accident (71.3%) were admitted to a hospital.

Of the respondents who had lost one or two days due to an accident, a vast majority (80.8%) decided not to lodge a complaint voluntarily or due to insignificant impact of the accident. Moreover, even if the lost productivity was a few months, only 37.1% of them complained about it. These statistics clearly indicate the reluctance of the majority of people to report accidents, which may be due to their lack of confidence in getting any compensation or the hassle to be experienced in the reporting process.

More than half of those who received full cost/ adequate amount as compensation had lodged a complaint. A higher percentage of respondents who had reported the accident had received full/ adequate amount of compensation, whereas a large majority (72.4%) of those underreported voluntarily had not received any compensation.

A vast majority (82.1%) of those admitted to a hospital for treatment had lost two days to two weeks of productivity and a majority (54.8%) of the respondents, who experienced accidents causing a few months of productivity loss, were admitted to a hospital for treatment. In each accident category, one or two days of productivity loss was more frequent than other durations of loss of productivity, wherein the accidents due to slipping on pavements had the highest percentage (72.6%). The longest productivity loss (a few months) was higher among accidents with a four-wheeled vehicle than those with other vehicle categories considered.

Of all kinds of accidents, the highest percentage was encountered by the students. Among the government employees and retired persons, accidents with a four-wheeled vehicle were higher compared to other categories under the Nature of involvement, whereas among the private sector employees, self–employed persons, and students, accidents with a motor bicycle were higher. The largest representing group of all the respondents (46.8%) were students and they mostly had accidents with a motor cycle. Among all employment groups, retried people were much less likely to meet with accidents compared to government or private sector employees, self–employed persons and students.

3.1. Chi-square Tests for Variables

Having observed the above salient features of data through the bivariate analysis, the next step was to carry out Chi-square tests of independence to test the association between the reporting tendency and each independent variable summarised in Table 3 with the following hypotheses:

Null hypothesis [H_0]: No significant association between the two variables

Alternative hypothesis [H_1]: There is an association between the two variables.

Further, Cramer's V value was also calculated with the Chi-square and p-values to find the strength of the associations.

Table 3: Chi-square Value, p-value and Cramer's V Value between the Reporting Tendency and the Categorical Independent Variables.

	Gender	Profession	Education	Nature	Severity	Loss	Compensation
Chi-square	6.593	3.515	3.334	92.433	166.389	31.132	145.515
P-Value	0.010	0.621	0.649	0.000	0.000	0.000	0.000
Cramer's V	0.097	0.071	0.069	0.364	0.488	0.211	0.457

The Pearson's correlation was calculated between the reporting tendency and the continuous variable, age, which was 0.081 with the p-value of 0.032.

According to the statistics in Table 3, null hypothesis could not be rejected for the variables of profession and education as the p-values for each of them was greater than 0.05 at 5% level of significance. As the other p-values were less than 0.05, it could be concluded that each of the variables, Gender, Nature, Severity, Loss, Compensation and Age was associated with the reporting tendency. Based on the Cramer's V value, it could be further determined that the variables, namely, nature, severity, and compensation had strong pairwise relationships with the reporting behaviour while the variables of gender and age had weak pairwise relationship with the reporting behaviour. The pairwise associations between the lost productivity and the dependent variables of severity and nature were also examined; the results are summarised in Table 4.

Table 4: Chi-square value, p-value, and Cramer's V value between the Loss and the Independent Variables of Severity and Nature

	Severity	**Nature**
Chi-square	197.291	62.417
P Value	0.000	0.000
Cramer's V	0.307	0.173

Applying the same criteria, it could be concluded that both severity and nature were associated with loss of productivity such that the severity showing a strong relationship while nature having a weak relationship with the loss.

3.2. Binary Logistic Regression to Fit the Model

Although six factors were found to be associated with the accident reporting behaviour using the Chi-square test and correlation in the bivariate analysis, it was possible that one or more of the other variables could be significant in a complex multivariate analysis. A binary logistic model [21] was fitted to investigate all factors affecting the accident reporting behaviour considering all independent variables. The dummy variables and the reference category in respect of each categorical variable were coded by the SPSS package to carry out the binary logistic regression with the forward conditional method.

Table 5 presents the odd ratios, Exp(B), p-values, and the coefficients in respect of the dummy variables for each predictor retained in the following regression equation:

$$logit\ P(x) = \ln\frac{P(x)}{1-P(x)} = 1.553 - .507 * \text{Nature}(1) - .878 * \text{Nature}(2) - 2.320 * \text{Nature}(3) - 1.265 * \text{Nature}(4) - 1.632 * \text{Nature}(5) - 2.153 * \text{Severity}(1) - 1.849 * \text{Severity}(2) - 1.845 * \text{Severity}(3) - .536 * \text{Compensation}(1) - 1.397 * \text{Compensation}(2) - 2.037 * \text{Compensation}(3) \quad (4)$$

It could be noticed that only the variables Nature, Severity and Compensation have made their presence in the BLR model.

The following hypothesis was used to check the significance of each predictor in the regression equation based on the p-values in Table 5, produced by the SPSS software.

H_0: The variable did not make a significant contribution to the prediction

vs

H_1: The variable made a significant contribution to the prediction.

As the null hypothesis is rejected for the variables with p–values less than 0.05, the nature, severity and compensation could be highlighted as the factors making the most significant contribution to the reporting tendency.

Table 5: Odd ratios, Exp(B), p-values, and Coefficients for Each Predictor in the BLR Equation

Predictor	Label	Coefficient (B)	p-value	Exp(B)
Nature:			.006	
Nature(1)	With a motor bicycle	-.507	.084	.602
Nature(2)	With a three-wheeler	-.878	.018	.415
Nature(3)	Slipped on pavement	-2.320	.029	.098
Nature(4)	Slipped on road	-1.265	.011	.282
Nature(5)	With a roadside structure/pole/tree	-1.632	.013	.196
Severity:			.000	
Severity(1)	Treated without admitting to a hospital	-2.153	.000	.116
Severity(2)	Treated without going to a hospital/medical centre	-1.849	.000	.157
Severity(3)	No injuries to treat	-1.845	.000	.158
Compensation:			.000	
Compensation(1)	Partial cost	-.536	.120	.585
Compensation(2)	Pending	-1.397	.007	.247
Compensation(3)	Not at all	-2.037	.000	.130
Constant		1.553	.000	4.725

According to the statistics presented in Table 5, except for the coefficient for Nature(l), i.e., accidents involving a motor bicycle, all the other coefficients of Nature

(2-5) were significant (p – value < 0.05) and hence they were associated with the tendency of reporting. Odds ratios in respect of accidents involving a motor bicycle, a three-wheeler, slipping on pavement/ road, a roadside structure/pole /tree were less than 1, which means all such accidents were less likely to be reported compared to accidents with a vehicle. More specifically, the odds of reporting accidents with each of the above modes declined by 39.8%, 58.5%, 90.2%, 71.8%, and 80.4% respectively compared to reporting accidents with a vehicle. It could be noted that among all accident types, slipping on pavement was much less likely to be reported compared to the rest.

Looking at the results for severity in the Table 5, it could be noticed that the coefficients for all Severity (1-3) were significant and negative. The odds ratio of the accidents treated without being admitted to a hospital means that they were 0.116 times (or 88.4%) less likely to be reported than those admitted to a hospital for treatment (reference category). Further, the accident-caused injuries treated without visiting a hospital/ medical centre and the accidents without any injuries were equally likely to be reported but their reporting tendencies were 0.157 (84.3%) and 0.158 (84.2%) times respectively lower than reporting accidents that caused admission to the hospital (reference category).

The coefficients for all Compensation (1-3) categories were negative but only (2) and (3) showed a significant association with the reporting behaviour. Compensation (1), which distinguished 'Partial Cost' from 'Full Cost/ Adequate Amount', had an odd ratio of 0.585, which meant that an accident partially compensated was 0.585 (41.5%) times less likely to be reported than an accident that received full cost/ adequate amount of compensation. Moreover, when the compensation was either pending or not eligible, i.e., Compensation (2) or (3), the odds of reporting those were 75.3% and 87% lower with reference to those receiving full cost/ adequate amount.

Since the knowledge of victims about the options available for claiming compensation was not tested, a further study on the knowledge of accident victims on procedures would be needed to clarify the degree of independence of the variable, compensation received.

To assess the goodness of fit of the model, the Hosmer and Lemeshow Test was conducted with the following hypotheses:

H_0: model was a good fit for data

vs.

H_1: model was not a good fit for data.

The resultant test statistics are summarised in Table 6.

Table 6: Hosmer and Lemeshow Test for Goodness of Fit of the Model

Step	Chi-square	p-value
3	13.653	.091

The p-value of 0.091 implies that there is a probability of greater than 0.05 for incorrectly rejecting the null hypothesis when it is actually true. Thus, at 5% level of significance, the H_0 cannot be confidently rejected, which indicates the model's possible good fit for data. However, the resultant low p-value (less than 0.1) is noteworthy, and hence, further research, preferably with more data, may be required to confirm the suggestive evidence pertaining to the goodness of fit of the model yielded from this study.

Analogous to the R^2 (coefficient of determination) in Multiple Linear Regression, the Cox and Snell R Square value and Nagelkerke R Square value in Table 7 confirmed that 29.1% to 46.4% of the accident reporting behaviour was explained by the binary logistic regression model.

Table 7: Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
3	448.477	.291	.464

3.3. Discussion

Pedestrian involved accidents as well as under-reporting of road traffic collisions are two focus areas among research related to road traffic accidents. While reinforcing the findings of such research on the degree of under-reporting and underlying reasons for it, this study focused on the severity of under-reporting accidents involving pedestrians such as slipping on pedestrian walkways (pavements) and roads. The loss of productivity due to different types of traffic accidents is a major setback to the economy of a country unless it is minimised. Authorities could act if there is a proper database of accidents taking place on roads and related infrastructure facilities, since the pedestrians, who account for a significant portion of traffic accidents, play an important role in controlling traffic. The database and the results reported in this article could be used as an initiative for the development of a regularly updated databank as incomplete data cause a serious problem in introducing viable mechanisms to control the number of accidents on roads.

According to the findings of this study, the severity, parties involved, and the nature of the second party influenced the under-reporting. Creating awareness on the rights and responsibility of pedestrians and bringing the authorities responsible for road safety to limelight would be effective in overcoming this problem. Though 33% of accidents of pedestrians appeared happening due to slipping on road or pavement or due to roadside structures it was shown that the percentages of reporting such accidents were at dismal 1.4% (slipped on pavement), 5.7% (slipped on the road), and 5.9% (with a roadside structure/pole/tree). If there was a proper mechanism for road safety audit on regular basis, some of these issues could have been prevented. As most of the road infrastructure facilities are designed considering the safety and convenience of vehicle occupants, this study has provided compelling evidence to demand basic facilities required for the safety of pedestrians.

The above findings call for more attention paid on providing safe and appropriate facilities for the pedestrians when road reconstruction and rehabilitation projects are planned in the future. Specially, unsafe fixings on pedestrian paths and improper materials used on foot walks should be avoided as a considerable number of unreported accidents are related to such obstacles.

In order to make a comparison with studies relevant to the present research, the following findings could be traced from limited information in the literature [13]. In low-income countries, the under-reporting of fatal and non-fatal accidents varied within the ranges of 0% - 61% and 69% - 80% respectively while that in middle-income countries for slight, severe, non-fatal, and fatally injured accidents was reported to lie within 93% - 98%, 32.5% - 96%, 34% - 99%, and 0.5% - 89.5% respectively.

A limitation of this study would be that accident reporting percentages by pedestrians may vary depending on the nature of the locality; that is, whether they are metropolitan, suburban, and rural areas. In particular, reporting pattern may change closer to urban areas considering the variation in pedestrian categories. This may be an aspect that could be explored through further research.

4. CONCLUSIONS

The present study was conducted to find out the issues pertaining to the reporting behaviour of road accidents occurred in Sri Lanka based on a primary data collection covering the entire island. The data were collected during the period from January to July, 2020 through a detailed questionnaire administered to 1242 respondents. After considering the multiple accidents experienced by the same respondent as separate cases and removing the missing values, a sample of 698 cases was carried forward to the study by means of univariate and bivariate analyses; Chi-square testing for the association of eight independent variables with the reporting behaviour, and fitting a

Binary Logistic Regression Model to identify the most significant factors affecting the reporting tendency.

Findings of the analysis revealed that 34% of respondents experienced at least one accident and the most significant group causing pedestrian accidents was motor bicycles. This inference aligns with the revelations reflected from accident records maintained by the Police. Therefore, it could be recommended that strict regulations should be introduced to control reckless motor cyclists and to safeguard pedestrians.

The considerably high incidence of accident under-reporting, revealed through this research, is also a major concern. In order to address this issue, hospitals and all other injury treatment organisations could be directed to collect relevant data and to share those with relevant regulatory bodies, including the Police, so that the reliability of data available could be improved and made available for making appropriate policy decisions as well as conducting further research on the subject.

When the severity of accidents is considered, accidents causing injuries and requiring treatment without visiting a hospital/ medical centre, and those that did not inflict any injuries, had approximately the same effect on the reporting tendency. Further, the accidents with compensation pending or ineligible for compensation were nearly twice as unlikely to be reported compared to partially compensated accidents.

Though not investigated in the present research, it may be worth if relevant authorities could closely monitor and scrutinise the procedures involved in the payment of compensation through motor insurance in respect of third-party involvement of accidents. This is because the majority of victims appeared to be not receiving any compensation, possibly owing to complex procedures involved in claiming third-party insurance. Victims are either unaware of the procedures involved, or reluctant to apply because of the perception that the procedures are not properly implemented. Introducing less complicated and transparent procedures and creating public awareness in that regard might help victims of accidents in claiming their compensation from third-party insurance policies.

It could also be recommended that the relevant authorities pay attention to the regulations pertaining to maintenance of accident-free and safe facilities. The legislations should be strong enough to bring those authorities neglecting the safety of road users before courts of law, which will be conducive towards improving the safety of all road user categories.

REFERENCES

[1] World Health Organisation. "Global status report on road safety 2018", WHO, Geneva, 2018

[2] E.A. LaScala, D. Gerber & P.J. Gruenewald, "Demographic and environmental correlates of pedestrian injury collisions: a spatial analysis. *Accident Analysis & Prevention* **32**(5): 651–658.,2000

[3] A. Loukaitou-Sideris, R. Liggett, & H.G. Sung, "Death on the crosswalk: a study of pedestrian-automobile collisions in Los Angeles". *Journal of Planning Education and Research* **26**(3): 338–351,2007, DOI: https://doi.org/10.1177/0739456X06297008

[4] C. D. Cottrill, & P. Thakuriah, "Evaluating pedestrian crashes in areas with high low-income or minority populations", *Accident Analysis & Prevention* **42**(6),2010, 1718–1728. DOI: https://doi.org/10.1016/j.aap.2010.04.012

[5] S. Ukkusuri, S. Hasan, S. & H.M.A. Aziz, "Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency". *Transportation Research Record: Journal of the Transportation Research Board* **2237**(01): 98–106.,2011, DOI: https://doi.org/10.3141/2237-11

[6] C. Siddiqui, M. Abdel-Aty, & K. Choi, "Macroscopic spatial analysis of pedestrian and bicycle crashes". *Accident Analysis & Prevention* **45**: 382–391.,2012, DOI: https://doi.org/10.1016/j.aap.2011.08.003

[7] J. Lee, M. Abdel-Aty, K. Choi, & C. Siddiqui, "Analysis of residence characteristics of drivers, pedestrians, and bicyclists involved in traffic crashes". *Transportation Research Board 92nd Annual Meeting,* Transportation Research Board of the National Academics, Washington DC, United States, 13-17, January2013, pp 22

[8] J. Lee, M. Abdel-Aty, K. Choi, & H. Huang, "Multi-level hot zone identification for pedestrian safety". *Accident Analysis & Prevention* **76**: 64-73. ,2015, DOI: https://doi.org/10.1016/j.aap.2015.01.006

[9] E. Hauer, and A. Hakkert, "Extent and some implications of incomplete accident reporting". Transportation Research Record **1185**: 1-10.,1988

[10] R. Elvik, and A. Mysen, "Incomplete accident reporting: meta-analysis of studies made in 13 countries". *Transportation Research Record: Journal of the Transportation Research Board* **1665**(01): 133–140.,1999, DOI: https://doi.org/10.3141/1665-18

[11] D.L. Rosman, "The Western Australian Road injury database (1987–1996): ten years of linked police, hospital and death records of road crashes and injuries". *Accident Analysis & Prevention* **33**(01): 81–88.,2001, DOI: https://doi.org/10.1016/S0001-4575(00)00018-X

[12] K.H.Janstrup,T. Hels, S. Kaplan, H.M. Sommer, and J. Lauritsen, "Understanding traffic crash under-reporting: linking police and medical records to individual and crash characteristics". *Proceedings of 5th conference of Transport Research Arena (TRA) 2014*, Paris, France, 14-17, April, pp 11.

[13] A.Ahmed, A.F.M. Sadullah, and A.S. Yahya, "Errors in accident data, its types, causes and methods of rectification analysis of the literature", *Accident Analysis & Prevention* 130: 3-21,2010, DOI: https://doi.org/10.1016/j.aap.2017.07.018

[14] S. Sciortino, M. Vassar, M. Radetsky, and M.M. Knudson, "San Francisco pedestrian injury surveillance: mapping, under-reporting and injury severity in police and hospital records". *Accident Analysis & Prevention* **37**(06): 1102–1113.,2005, DOI: https://doi.org/10.1016/j.aap.2005.06.010

[15] E.Amoros,J.L. Martin, and B. Laumon, "Under-reporting of road crash casualties in France". *Accident Analysis & Prevention* **38,** 2006**,** (04): 627–635. DOI: https://doi.org/10.1016/j.aap.2005.11.006

[16] B.P.Y. Loo, and K.L. Tsui, "Factors affecting the likelihood of reporting road crashes resulting in medical treatment to the police". *Injury Prevention* **13**(03): 186–189. ,2007, DOI: 1 https://doi.org/10.1136/ip.2006.013458

[17] E. Mako, and P. Szakonyi, "Evaluation of human behaviour at pedestrian crossings". *Transportation Research Procedia* **14**: 2121 – 2128. 2016, DOI: https://doi.org/10.1016/j.trpro.2016.05.227

[18] S. Meeyai, "Logistic Regression with Missing Data: A Comparison of Handling Methods, and Effects of Percent Missing Values". *Journal of Traffic and Logistics Engineering* **4,** 2016 (2).

[19] D.W. Hosmer Jr., S. Lemeshow and R. X. Sturdivant, "Applied logistic regression", volume 398, 3rd edition, pp 155-168, John Wiley & Sons, Hoboken, New Jersey. 2013, DOI: https://doi.org/10.1002/9781118548387

[20] A.Agresti, "Introduction to categorical data analysis", 2nd edition, pp 72, John Wiley & Sons, Hoboken, New Jersey, 1996

[21] J.M. Hilbe, "Logistic Regression". *International encyclopedia of statistical science*, pp 15-32. ,2011