Case Studies on Transport Policy 11 (2023) 100935

Contents lists available at ScienceDirect

Case Studies on Transport Policy

journal homepage: www.elsevier.com/locate/cstp

Statistical approaches for assessing the effectiveness of safety devices use in preventing head injuries from motorcycle crashes

Mena Lao [a], T.M Lukusa [b,*]

[a] Department of Statistics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand
[b] Department of Statistics and Institute of Data Science, National Cheng Kung University, Taiwan

A R T I C L E I N F O

Keywords:
Binary response
Mosaic plots
Contingency table
Head injuries
Motorcycle crashes
Safety devices

A B S T R A C T

This study investigates the effectiveness of motorcyclist's helmet use in preventing head injuries using the binary regression model and the mosaic plots. The association between head injuries and helmet use was modeled by applying binary logistic regression. The fitted model to the Changhua city accidents dataset has revealed that wearing a helmet had a positive impact in preventing the motorcyclist from sustaining head injuries in road traffic crashes. In addition to the association between head injuries and non-wearing of helmet, we found that alcohol impairment risk factor was increasing the likelihood of head injuries during crashes. Besides, the results were supported by using the mosaic plots which provide a visualization of the joint effect of Alcohol, helmet and heading injury for motorcyclists involved in crashes.

1. Introduction

Statistics related to the motorcycle road traffic accidents (RTA) and road traffic injuries (RTI) in many Asian countries are worrisome, despite governments strategies and efforts. As far as the motorcycles RTA are concerned, Taiwan is not exempted (e.g., Hsu et al., 2003; Lin et al., 2022). The deaths and injuries tolls resulting from those crashes are believed to be highly correlated with head injuries (Nguyen, 2013). Head injuries are a leading cause of death and disability due to traffic accidents involving motorcycles (World Health Organization, 2015). Pai et al. (2018) state that with a total area of 36,197 km^2 and a population of 23,552,470, there were about 217 cars/km^2 and 378 motorcycles/km^2 registered, in Taiwan. Road traffic-related crashes are the leading cause of injuries requiring hospitalization. Bicyclists and motorcyclists substantially contribute to the morbidity and mortality rates of road crash casualties (Lin and Kraus, 2009). Motorcyclist fatalities in Taiwan contribute to the most traffic deaths (Pai et al., 2018). In 2013, official statistics showed that motorcyclists account for 51% of the total traffic fatalities, and the numbers of fatal injuries were 20 times higher among motorcyclists than among automobile drivers. On average, one motorcyclist death occurs every 24 h in Taiwan (Ministry of Transportation and Communication, 2015). Mayrose (2008) and Wiznia et al. (2016) have suggested that head injuries are the primary cause of deaths and hospitalization among motorcyclists and bicyclists. Previous studies have documented that wearing a motorcycle helmet can reduce the risk of death and severe injury, for instance, Jou et al. (2012), Lin et al. (2022) and Lin and Kraus (2009). Effective enforcement of motorcycle helmet laws can increase helmet-wearing rates and thereby reduce head injuries (World Health Organization, 2015). Studies conducted on motorcyclist injuries have reported that helmet use and related laws have successfully reduced head injuries, thus reducing fatalities among motorcyclists (e.g., Weiss, 1992; Keng, 2005). Ichikawa et al. (2003) reported a 41% reduction in head injuries in Thailand 2 years after the implementation of a mandatory helmet use law. Likewise, many countries in the world reported a decline in head injuries and fatalities, after the implementation of helmet use laws (Pai et al., 2017; Pai et al., 2018). Past studies of the effect of motorcycle and bicycle helmet on injured body regions were systematically reviewed (Liu et al., 2008). In general, these previous studies were concluded that helmet use was associated with reduced odds of death and head injury. The effectiveness of helmet use can be assessed by comparing head injuries among helmeted and unhelmeted motorcyclists. If the odds of a head injury for helmeted motorcyclists are significantly lower than for unhelmeted motorcyclists, it means that helmet use is effective in preventing head injuries in traffic accidents (Weiss, 1992; Liu et al., 2008, etc.).

The objective of this research is to scrutinize the effectiveness of the motorcyclist helmet use in preventing head injuries in RTI by using the logistic regression and mozaic plots. We collected the RTA data from

* Corresponding author.
E-mail address: z10902020@ncku.edu.tw (T.M Lukusa).

https://doi.org/10.1016/j.cstp.2022.100935
Received 3 May 2021; Received in revised form 2 November 2022; Accepted 30 November 2022
Available online 9 December 2022

Taiwan National Police Agency (NPA) in charge of traffic accidents. These accidents had occurred in Changhua city, a city, located in the central part of Taiwan from 2011 to 2013. We examine primarily the association between head injury event and motorcyclist behaviors including helmet use and alcohol impairment in RTA. Head injury is a dichotomous outcome variable for the logistic regression model. Besides the helmet use and alcohol impairment behaviors, demographic variables and other factors generally associated with motorcyclists's head injuries in RTA are included in logistic regression.

2. Materials and methods

2.1. Data

To evaluate the effectiveness of the helmet use in preventing head injuries in Taiwan RTI, the data was collected in Changhua, a city located in the central part of Taiwan, from 2011 to 2013 by the Taiwan NPA. The total number of motorcycle crashes was 24,351. The flowchart for sampling of data is presented in Fig. 1. For each crash, we collected age, gender, injury type (head/non-head injury), helmet use, alcohol impairment, time of a day (Day/night), traffic position, (Intersection/straight line) and road categories. The data description is summarized in Table 1. Among all 24,351 motorcycle related accidents, 6.56% of them were head injuries, 56.48% of them were male, 99.14% of them were helmeted, and 5.50% of them were identified as alcohol impaired riding. For convenience, factor Age was recorded into four age groups (levels) including extreme age, young adults (18–24) who are university students mostly, adults (25–44) and middle adults (45–64). This classification mimics the one used in Lin et al. (2022). Approximately 60% = (29.23% +31.53%) of them were between 18 and 44 Years old, and 25% were between 45 and 65 Years old. In the extreme age groups, there were about 15% = (3.03% +11.56%) of motorcyclists. Regarding road category, 60.23%, 20.25% and 19.52% occurred in rural, township, and urban roads, respectively. Roads types in Taiwan is classified into 5 main categories according to Jou et al. (2012) including Freeways, provincial roads, county roads, rural roads and urban roads. Changhua city has mostly county roads, rural roads and urban roads. County roads, rural roads and urban roads are positioned as intra-city with a speed limit lower of 40–60 km/h. In relation to traffic position, 63.43% of them were accidents at the intersection while the rest were accidents on a straight line. Moreover, 74.09% of them were accidents during day time.

The distribution of the motorcyclists head injury in connection to other categorical variables is shown in Table 2. It can be seen that the percentage of head injury among helmeted and unhelmeted motorcyclists are 6.34% and 32.06%, respectively. The percentage of head injury in male is slightly more than in female. The highest proportion of head injury per age group was among extreme age groups, (<18 and >65 Years old). These two groups can be seen as made of most vulnerable people. The head injury was the least among 18–24 years old.

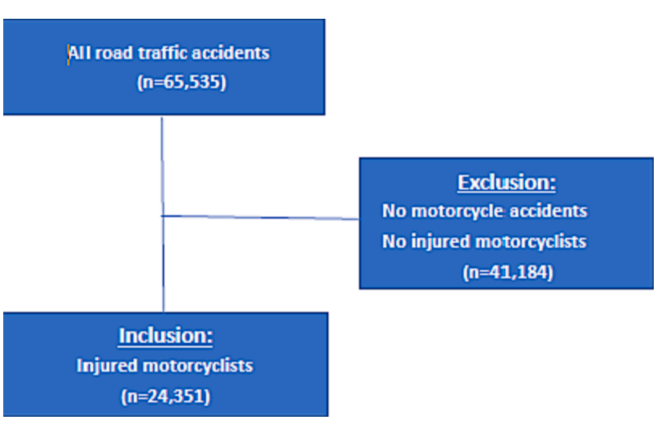

Fig. 1. Sampling Flowchart.

Table 1
Data Description

Variables	Categories	Frequencies	Percent
Injury Type	Head injury	1597	*(6.56 %)*
	Not head injury	22754	*(93.44 %)*
Gender	Male	13754	*(56.48 %)*
	Female	10597	*(43.52 %)*
Age-group (Years-old)	less 18	739	*(3.03 %)*
	18–24	7118	*(29.23 %)*
	25–44	7679	*(31.53 %)*
	45–64	5999	*(24.64 %)*
	over 65	2816	*(11.56 %)*
Safety device	Helmet	24142	*(99.00 %)*
	Not helmet	209	*(1.00 %)*
Alcohol impairment (mg/l)	No alcohol	23012	*(94.50 %)*
	Alcohol	1339	*(5.50 %)*
Time of the Day	Day	18041	*(74.00 %)*
	Night	6310	*(26.00 %)*
Traffic Position	At the intersection	14917	*(61.00 %)*
	Straight line	9434	*(39.00 %)*
Semaphore Type	Traffic Control	6011	*(24.68 %)*
	Flashing	3015	*(12.38 %)*
	None	15325	*(62.93 %)*
Road Types	Township	4924	*(20.22 %)*
	Urban	4755	*(19.53 %)*
	Rural	14672	*(60.25 %)*

Table 2
Distribution of Accidents by Injury and Regressors.

	Injury Type				
	Head Injury		Non-head injury		
Variables	**Categories**	**counts**	**Percent**	**counts**	**Percent**
Gender	Male	983	*(7.15 %)*	12,771	*(92.85 %)*
	Female	614	*(5.79 %)*	9,983	*(94.21 %)*
Age (Years old)	less 18	71	*(9.61 %)*	668	*(90.39 %)*
	18–24	330	*(4.64 %)*	6,788	*(95.36 %)*
	25–44	468	*(6.09 %)*	7,211	*(93.91 %)*
	45–64	467	*(7.78 %)*	5,532	*(92.22 %)*
	over 65	261	*(9.27 %)*	2,555	*(90.73 %)*
Safety device	Helmet	1530	*(6.34%)*	22,612	*(93.66%)*
	Not helmet	67	*(32.06%)*	142	*(67.94%)*
Alcohol consumption (mg/l)	No alcohol	1,269	*(5.51 %)*	21,743	*(94.49%)*
	Alcohol	328	*(24.50 %)*	328	*(75.50%)*
Time of the Day	Day	1,089	*(6.04 %)*	16,952	*(93.96 %)*
	Night	5802	*(8.05 %)*	5,802	*(91.95 %)*
Traffic Position	At the intersection	816	*(5.47%)*	14,101	*(94.53%)*
	Straight line	781	*(8.28%)*	8,653	*(91.72%)*
Semaphore Type	Traffic Control	290	*(4.82%)*	5,721	*(95.18%)*
	Flashing	178	*(5.90%)*	2,837	*(94.10%)*
	None	1,129	*(7.37%)*	14,196	*(92.63%)*
Road Types	Township	358	*(7.27%)*	4,566	*(92.73%)*
	Urban	148	*(3.11%)*	4,607	*(96.89%)*
	Rural	1,091	*(7.44%)*	13,581	*(92.56%)*

Counts refer to the cell frequencies.

The percentage of head injury in no alcohol consumption is less than in alcohol consumption. The percentage of head injury in daytime and night time are about 6% and 8% respectively. The percentage of head injury which are related to crashes at the intersection and the straight line are 5.47% and 8.05%, respectively. The percentage of head injury which are related to crashes to township, urban and rural road types are 7.27%, 3.11% and 7.44% respectively. There are other types of records in the NPA database such as crash types, motorcycle engine size, and crash partner types. In general, only relevant or specific records are used in a particular study.

2.2. Binary logistic regression model

Logistic regression model describes how a binary (or multinomial) response variable is associated with independent variables, which can be discrete or continuous (Dobson and Barnett, 2008). In this regression model, the probability or odds of the response having a certain value is modeled based on the values of independent variables (Agresti, 2013). Binary logistic regression is a type of regression where a binary response variable is associated with independent variables. Logistic regression models have been intensively used in investigating causes related to crashes. For instance Tay et al. (2008), Borucka (2020), Ghamdi (2002), and Karacasua et al. (2014) used binary logistic regression to explain the effects of the independent variables on the binary response.

Let Y be a binary response variable. For observation i, Y_i takes on the value 1 if an event of interest is present, otherwise Y_i takes on the values 0. Let $X_1, X_2, \cdots, X_k$ be k independent variables, and for the observation i, letX_{ij} be the observed value of the independent variable X_j for $j = 1, 2, \cdots, k$. The logistic regression estimates the probability that a specific event is present given the values of independent variable. Let π_i be the probability that an event is present for observation i depending on independent variables, that is, $\pi_i = P(Y_i = 1)$. The binary logistic regression model is in the form:

$$log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}, \tag{1}$$

where the term $\frac{\pi_i}{1-\pi_i}$ is the odd that the event is present. In practice, the log odds are modeled based on independent variables. Therefore, model (1) can also be written as:

$$\pi_i = \frac{e^{\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik}}}{1+e^{\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik}}}, \tag{2}$$

where $\beta_0, \beta_1, \cdots, \beta_k$ are the regression coefficients (parameters). Let $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_k)$ be the vector of parameters, such as $\boldsymbol{\beta} \in \Theta$, with Θ known as the space parameter. Hence, one short form for expression (2) can be written as follows

$$\boldsymbol{\pi} = \frac{e^{\boldsymbol{\beta}\mathbf{X}^T}}{1+e^{\boldsymbol{\beta}\mathbf{X}^T}} = \left[1+e^{-\left(\boldsymbol{\beta}\mathbf{X}^T\right)}\right]^{-1}, \tag{3}$$

where $\mathbf{X} = (1, X_1, X_2, \cdots, X_k)$ is known as the design matrix. For completely uncorrelated independent variables, the coefficient β_j represents the change in the log of the odds per unit change in the independent variable X_{ij}, that is, $e^{\beta_j} = \exp(\beta_j)$ is the multiplicative effect on the odds for one unit increase in X_j (Agresti, 2013). If β_j is greater than zero, then e^{β_j} is greater than 1, and the odds increase; on the other hand, If β_j is less than zero, then e^{β_j} is less than 1, and the odds decrease. It is known that e^{β_j} is the odds ratio for a unit change in X_j. For parameter estimation, the maximum likelihood (ML) method can be applied (Agresti, 2013). The ML method is a likelihood-based method. For n observations,

$$L\left(\beta_0, \beta_1, \beta_2, \cdots, \beta_k\right) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i} \tag{4}$$

By applying the natural logarithm on the expression (4), we obtain the corresponding log-likelihood function, denoted by $\ell(\boldsymbol{\theta}) = \log L(\beta_0, \beta_1, \cdots, \beta_k)$, and expressed as

$$\begin{aligned}\ell\left(\boldsymbol{\theta}\right) &= \log\left[\prod_{i=1}^{n}\pi_i^{y_i}(1-\pi_i)^{1-y_i}\right] = \sum_{i=1}^{n}\log\left(1-\pi_i\right) + \sum_{i=1}^{n} y_i \log\frac{\pi_i}{(1-\pi_i)} \\ &= \sum_{i=1}^{n} -\log\left(1+e^{\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik}}\right) + \sum_{i=1}^{n} y_i \log\left(e^{\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik}}\right).\end{aligned} \tag{5}$$

The maximum likelihood estimate (MLE) of $\boldsymbol{\beta}$, denoted by $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_0, \widehat{\beta}_1, \cdots, \widehat{\beta}_k)$, is derived by taking partial derivatives of $\ell(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ and setting them equal to zero, then solving. For $j = 1, 2, \cdots, k$, the partial derivatives of $\ell(\boldsymbol{\beta})$ with respect to β_j is

$$\frac{\ell(\boldsymbol{\beta})}{\partial\beta_j} = \sum_{i=1}^{n} \frac{e^{\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik}}}{1+e^{\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik}}} x_{ij} + \sum_{i=1}^{n} y_i x_{ij} = \left(y_i - \pi_i\right) x_{ij} = 0. \tag{6}$$

The term $\sum_{i=1}^{n} y_i x_{ij}$ in (6) is not relevant in deriving the MLE of $\boldsymbol{\beta}$, since it does not have information about $\boldsymbol{\beta}$. Unfortunately, expression (6) does not have a closed-form solution leading to $\widehat{\boldsymbol{\beta}}$ smoothly. Thus, the MLE of $\boldsymbol{\beta}$ is obtained using the iterative reweighing least squares (IWLS) algorithms.

To assess the goodness-of-fit of the logistic regression model, we use the Hosmer–Lemeshow statistic (Hosmer and Lemeshow, 1980). This test statistic divides the data set into g groups according to the estimated probabilities that the binary outcome is one, and then compares the number of observed and expected frequencies. The overall fit of the model can be tested under the null hypothesis H_0 that the model fits the data adequately. The Hosmer–Lemeshow test statistic is calculated by

$$H = \sum_{k=1}^{g} \frac{\left(O_k - n_k'\overline{\pi}_k\right)^2}{n_k'\overline{\pi}_k(1-\overline{\pi}_k)}, \tag{7}$$

where n_k' is the total number of subjects in the k^{th} group, c_k denotes the number of covariate patterns in the k^{th} decile, $O_k = \sum_{j=1}^{c_j} y_j$ is the number of responses among the c_k covariate patterns, and g is the number of groups. The term covariate pattern refers to a single set of values for the covariates in a model. Suppose the fitted model contain p independent variables, $\boldsymbol{x} = (x_1, x_2, \cdots, x_p)$, and let J denote the number of distinct values of $\boldsymbol{x}$ observed. Let $m_j, j = 1, 2, 3, \cdots, J$, be the number of subjects with $\boldsymbol{x}' = \boldsymbol{x}$. Then, $\overline{\pi}_k = \sum_{j=1}^{c_k} \frac{m_j \widehat{\pi_j}}{n_k'}$ is the average estimated probability, where $\overline{\pi}_k$ is the estimated probability. The distribution of the H is approximated by the chi-square distribution with $g-2$ degrees of freedom.

3. Results

To obtain our numerical and graphical results, we used the statistical software R. Let Y_i be the injury position on victim i, and Y_i is a binary outcome which takes on the values 1 for head injuries and 0 for non-head injuries. Let $\pi = P(Y_i = 1)$ be the probability of head injury. The number of motorcyclists related to head and non-head injuries were 1, 597 and 22, 754, respectively, as shown in Table 1. The potential risk factors to explain the head injuries event in Changhua city RTI are age, helmet use, alcohol impairment, time of a day, traffic position, traffic control semaphore and crashes site. Thus, these factors are the predictors in the model, controlling for possible confounding. The logistic regression model (*Model* 1) has the following form:

$$\begin{aligned}log\left(\frac{\pi_i}{1-\pi_i}\right) = {} & \beta_0 + \beta_1 \mathrm{AGELEVEL1}_i + \beta_2 \mathrm{AGELEVEL2}_i + \beta_3 \mathrm{AGELEVEL3}_i \\ & + \beta_4 \mathrm{Day2}_i + \beta_5 \mathrm{SITE2}_i + \beta_6 \mathrm{POSITION2}_i + \beta_7 \mathrm{SEMAPHORE2}_i \\ & + \beta_8 \mathrm{HELMET}_i + \beta_9 \mathrm{ALCOHOL}_i,\end{aligned} \tag{8}$$

where age risk factor is defined by four levels: AGELEVEL1 (baseline) refers to motorcyclists aged less than 18 or above 65 old (extreme groups), AGELEVEL2, AGELEVEL3, and AGELEVEL4 refer to riders with age 18–24, 25–44, 45–65, respectively. The Day risk factor has by two levels defined as DayTime and NightTime (baseline). The road types included Urban, Township and Rural roads, as they are defined in Section 2.1. HELMET is a categorical variable representing helmet used and defined as helmet used and helmet not used; POSITION is a factor referring to at the intersection and straight lane (baseline); SIGNALIZATION (Semaphore & flashing) is a factor referring to presence of signalization and absence of signalization (baseline); ALCOHOL is a categorical variable with two levels defined as alcohol impairment and no alcohol impairment. The coefficient β_8 is the log odds ratio of head injury comparing helmeted versus unhelmeted motorcyclists, thus it indicates the effectiveness of helmet use in preventing head injury. To account for the joint effect between HELMET and ALCOHOL, an interaction term is added in the Eq. (8). The extended binary regression model with interaction (*Model* 2) is expressed as follows:

$$log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1\text{AGELEVEL1}_i + \beta_2\text{AGELEVEL2}_i + \beta_3\text{AGELEVEL3}_i + \beta_4\text{Day2}_i + \beta_5\text{SITE2}_i + \beta_6\text{POSITION2}_i + \beta_7\text{SEMAPHORE2}_i + \beta_8\text{HELMET}_i + \beta_9\text{ALCOHOL}_i + \beta_{10}\text{HELMET}_i \times \text{ALCOHOL}_i \quad (9)$$

Both regression models given in (8) and (9), respectively, are investigated and discussed. The result of the fitted logistic regression model is shown in Table 3.

The overall fit of the model was assessed by testing the null hypothesis that the model fits the data adequately. The value of the Hosmer–Lemeshow goodness-of-fit statistic is $H = 9.0578$ using $g = 10$, and has chi-square distribution with $g - 2 = 8$ degrees of freedom with the corresponding p-value of 0.3374. Clearly, the H statistic was not significant at the 0.05 level with the p-value > 0.05. The null hypothesis H_0 is rejected. Hence, Model 1 fits the data adequately (Table 4).

Based on the results in Table 3, the coefficient estimate of HELMET, $\widehat{\beta_8} = 1.5193$, was statistically significant at 0.05 level. It showed that for unhelmeted motorcyclists, the odds of a head injury were $e^{\widehat{\beta_8}} = 4.5688$ times higher than helmeted motorcyclists, given that the other variables were held constant. In other words, the helmeted motorcyclists were 4.5688 times less likely than unhelmeted motorcyclists to suffer a head injury. It can be concluded that using a helmet tends to prevent head injuries when compared with not using a helmet.

The coefficient estimate of ALCOHOL, $\widehat{\beta_9} = 1.5398$, was statistically significant at 0.05 level. This result indicated that for Alcohol motorcyclists, the odds of a head injury were $e^{\widehat{\beta_9}} = 4.6638$ times higher than No alcohol motorcyclists, given that the other variables are held constant.

Table 3
Logistic regression analysis.

Param	**Estimate**	**Std. Error**	exp(**coef**)	**2.5 %**	**97.5 %**
Intercept	−2.8981***	0.0886	0.0551	0.0462	0.0655
Age.18–24	−0.6807***	0.0828	0.5063	0.4304	0.5955
Age.25–44	−0.5505***	0.0776	0.5766	0.4955	0.6717
Age.45–64	−0.3239***	0.0779	0.7233	0.6212	0.8430
Night.time	0.1482*	0.0603	1.1598	1.0297	1.3044
Urban.County. Road	0.3075***	0.0574	1.3601	1.2161	1.5229
Not.at.Intersection	0.2684***	0.0577	1.3078	1.1680	1.4645
No.Signalization	0.1551*	0.0732	1.1678	1.0128	1.3498
Helmet.not.used	1.5193***	0.1616	4.5688	3.3107	6.2438
Drink.Alcohol	1.5398***	0.0758	4.6638	4.0162	5.4059

head injury is used as outcome variable.
p-value: < 0.001 is '***', < 0.01 is '**', < 0.05 is '*', and < 0.1 is.
Odd ratio (OR) is given by exp(**coef**).

Table 4
2 by 2 by 2 Contingency Table.

Alcohol Consumption	Helmet		
	Helmet used	Helmet not-used	Total
No alcohol	22,851 ***99.30%***	161 ***0.70%***	23,012
Alcohol	1,291 ***96.42%***	48 ***3.58%***	1,339
Total 2	24,142	209	24,351

Since Helmet and Alcohol risk factors were both very significant in the binary regression model (8), it was natural to examine the contribution of their joint effect on head injuries. Therefore, a binary logistic regression with the interaction term HELMET × ALCOHOL was investigated (9). The table results of logistic regression with interaction did not show any substantial improvements from the results of model 1. Showing this table result seems unnecessary tables. However, it can be provided upon request. Regarding model 2, the regression coefficients and the standard errors (in brackets) of the intercept term, factor HELMET, factor ALCOHOL and the interaction term HELMET × ALCOHOL are −2.90197(0.08877); $\widehat{\beta}_8 = 1.61269(0.18636)$; $\widehat{\beta}_9 = 1.55345(0.07697)$; and $\widehat{\beta}_{10} = -0.33507(0.35339)$, respectively. Overall, these values are very similar to those in Table 3.

From the analysis results, it happens that the joint effect HELMET × ALCOHOL was not statistically significant at all. Their odds ratios are 0.055, 5.02, 4.7278 and 0.7152. Except the intercept term which is not significant, the interpretation of the HELMET and ALCOHOL factors agree with binary regression without interaction term. If the interaction term in model 2 $\widehat{\beta}_{10} = 0.7152$ was significant, $\exp(\beta_{10}) = 0.7152$ would be the odds ratio of head injury comparing helmeted versus unhelmeted motorcyclists amongst motorcyclists with alcohol impairment.

The coefficient β_{10} measures the difference between the log odds ratio of head injury comparing helmeted versus unhelmeted sober motorcyclists; and the log odds ratio of head injury comparing helmeted versus unhelmeted impaired motorcyclists. Thus, β_{10} reveals how the effectiveness of wearing a helmet in preventing head injuries for impaired motorcyclists differs from the effectiveness of helmet use for the sober motorcyclists.

4. Supplementary analysis

In regression *Model* 2, we found that the joint effect between helmet wearing and alcohol impairment factors does not significantly predict the head injury event. To this end, contingency table and mosaic plots are performed to examine the conditional associations between head injury, helmet wearing and alcohol impairment numerically and graphically. From, helmet wearing rates for motorcyclists with no alcohol and with alcohol impairment are 99.30% and 96.43% respectively. Although these two helmet wearing rates are quite close, but it can be seen that the effectiveness of helmet use in preventing head injuries for those with no alcohol impairment is slightly higher than the one for those riding under the effect of alcohol impairment. Overall, the effectiveness of helmet use to prevent motorcyclists head injuries for those riding not under the influence of alcohol impairment is not statistically significantly different from those under the effect of alcohol consumption.

*Mosaic Plots [*Friendly, 1994*;* Hartigan and Kleiner, 1981*]*. Mosaic plots is used for visualizing the patterns of association between two or more categorical variables. This plot is made up of rectangular tiles whose sizes are commensurate to the frequencies of the table cells. The width of the box is equal to the total count in that column. The height of the box is the proportion of individuals in the column which fall into that cell. For color reference, deep blue means there are more observations in a particular cell than would be expected under the null model, whereas deep red represents fewer observations than would have been expected. The inference of mosaic plots uses a chi-square statistics (χ^2). Reject the null hypothesis means the chi-square (χ^2) with K degree of freedoms indicates a substantial departure from independence. The overall Pearson χ^2 statistic is just the sum of squares of the residuals. The residual cells are denoted by r_{ij}.

Note that small, medium, and large residuals are characterized by $|r_{ij}| < 2, 2 < |r_{ij}| < 4$, and $|r_{ij}| > 4$, respectively. The intense highlighted cells are those with residuals individually significant at approximately the $\alpha = 0.05$ level, whereas less (or no) highlighted cells are believed to have a nonsignificant contribution to the total χ^2. Weak significance lies between the two extreme color shadings. From Fig. 2, different patterns on the mosaic plot are not similar. It indicates that conditional associations exist between safety device (HELMET), alcohol impairment (ALCOHOL) and head injury. Moreover, the residual Pearson χ^2 has p-value less than 0.001, which shows that there is a substantial departure from the independence. Nevertheless, the 2 by 2 contingency table for assessing the association between HELMET and Alcohol has $\chi^2 = 743.9514$ with degree of freedom $K = 1$ and is p-value $p < .001$. This agrees with the fact that the interaction term was not significant in *Model 2*. Hence, helmet wearing and alcohol impairment seem to be conditionally associated with head injury as shown in the mozaic plots shown in Fig. 2.

5. Discussion

Overall, our results agree with many previous findings on different specific aspects. We discuss our results with regards to some related works in the literature. Jou et al. (2012) investigated high-risk factors for motorcyclist fatality in Taiwan. They found that rider age above 60 years, not wearing a motorcycle helmet, riding after drinking, and driving without a valid license as high-risk factors of motorcyclists deaths and severe injuries. We found that the age groups were positively associated in increasing order with RTI head injury. Analyzing the Malaysian RTI data, Ramli et al. (2011) concluded that the use of motorcycle helmets decreases the incidence and severity of lethal head injury in motorcycle crashes when compared with unhelmeted riders. This Ramli et al. (2011)'s finding agrees with our results.

We should mention one important aspect that the frequency of impaired and unhelmeted motorcyclists tends to be small when people care more for their safety (see, Table 2). Eventually, countries where the compulsory helmet law is still challenged, the frequency is likely to be large. Lin et al. (2022) found that in Taiwan there was high compliance with the compulsory helmet law. Thus, investigating unhelmeted behavior in some cases may not work due to lack of observations. In the same way, Chiu et al. (2000) found that the Taiwan motorcycle helmet use law was effective in reducing the number and severity of head injuries related to motorcycle crashes. However, many of these studies did examine the joint effects between wearing helmet and level of alcohol impairment while riding the motorcycle. It is probable that group has very few observations. Consequently, many statistical analyses did not provide meaningful results. In our study, we found that the interaction term HELMET and ALCOHOL was not significant under the logistic regression. This is the reason why the graphical analysis by the mozaic plot has become very important in detecting a conditional and positive association between wearing helmet and level of drinking alcohol in RTI.

Iamtrakul and Moinul (2007)'s Table 2 revealed that the severity in Thailand RTA could be as high as 2.26 times for the impaired and unhelmeted motorcyclists if compared with the victims wearing helmet as well as not drunk. In our study, unhelmeted motorcyclists were 4.6

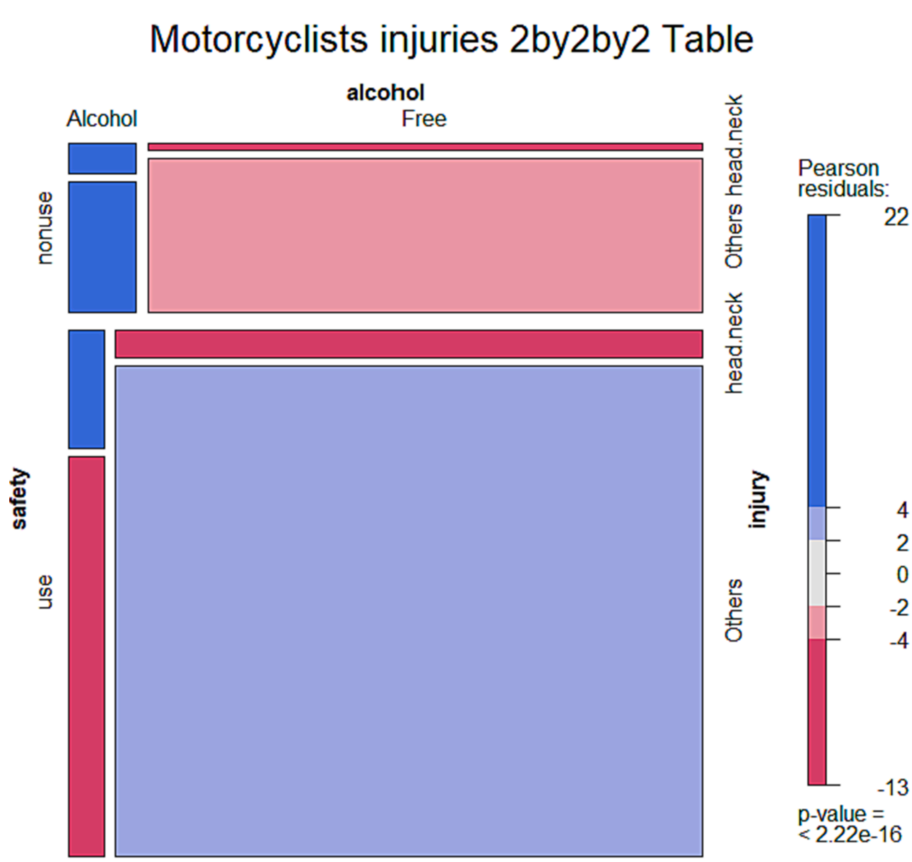

Fig. 2. Mozaic Plots.

times at risk than the helmeted ones to sustain head injury. This finding agrees with results from Kasantikul et al. (2005). However, the association of unhelmeted and impaired riding behaviours was positive but not significant. Contrarily to Taiwan, in Thailand, Malaysia and Vietnam, the proportions of unhelmeted and impaired motorcyclists is much larger. In this respect, it is possible to estimate the contribution of this joint risk factor on head injury event. Dos Santos et al. (2019) in their study related to the north eastern part of Brazil showed that riding a motorcycle under the effects of alcohol was associated with other risky forms of behaviors in traffic including speeding, not using a helmet, and not having a motorcycle driver's license. They pointed out that the interaction between those behaviors might lead to greater occurrence and severity of accidents. Contrarily to our study, they included the factor speed in analyzing their RTI data. Besides, they found that alcohol and not using a helmet risk factors were associated with greater occurrence and severity of accidents. As for the risk factors associated with sustaining head injuries among Taiwanese motorcyclists, our findings can be supported by Chen and Pai (2019), Pai et al. (2018) and Pai et al. (2017)'s findings. However, the magnitudes of our estimates differ from Chen and Pai (2019), Pai et al. (2018) and Pai et al. (2017)'s estimates because they linked NPA database and National heath insurance records (NHIR). In this matter, Yeh (2011) revealed many scenarios with the ratios for the accuracies between police-reported data and the data base obtained by linking the NPA report to the NHIR. In the case of motorcyclist's death, the ratio is larger (1.6–1.8) than for motorcyclist's injury (1.1–1.6). Again, it shows the benefits of using linked NPA records to the NHIR. This shows the benefits of using linked NPA records to the NHIR.

6. Conclusion

In this study, our interest was primarily to investigate the effectiveness of wearing helmet in predicting the motorcyclist's head injury in connection with alcohol impairment. To this end, we used a binary logistic regression model and found that the unhelmeted motorcyclists were 4.6 times more likely than helmeted motorcyclists to sustain a head injury in RTA. Regarding the effect of alcohol impairment, we found that impaired motorcyclists were 4.7 times more likely than those who were sober to sustain a head injury in RTA. We examined the effect of the joint risk factor of alcohol impairment and wearing/not wearing an helmet by a motorcyclist in RTA. The joint effects of riding motorcycle under alcohol effect and helmeted versus being sober but unhelmeted motorcyclists were found to be nonsignificant, contrarily to our expectations. Due to a small sample size for this specific category of victims, logistic regression could not reveal the expected results.

Besides, the contingency table and the mosaic plots were conducted to deeply examine the joint effect of factors Helmet wearing and alcohol impairment. We found that the effectiveness of wearing the helmet in preventing head injuries in RTA for sober motorcyclists was slightly more crucial than the effectiveness of those wearing helmet but alcohol impaired. However, this difference was not statistically significant. At least, mosaic plots revealed a conditional association between helmet use and alcohol impairment. Further investigation is needed to find the effect of an association between the helmet use, the alcohol impairment and head injury in motorcycles RTA.

About the mosaic plots, the future direction should include risk factors such as speed, crash types, motorcycle engine size, and crash partner types, since they are potential factors in RTA. Importantly, the frequency cells of the various categories in the contingency table need to have enough observations.

About the TRA database, Yeh (2011), Chen and Pai (2019), Pai et al. (2018) and Pai et al. (2017) recommend the use of the linked NPA database and NHIR in order to avoid underestimation due to using NPA database only or NHIR only. The study of traumas in RTA would require follow-ups of victims which is beyond police-recorded data. Consequently, the use of a NPA-recorded data can be seen as a limitation in this study. To have reliable results that are based on the linked NPA-NHI RTI records, Taiwan Ministry of Health and Welfare should relax procedures for obtaining the NHIR.

Since, our aim was primarily to investigate existence of any associations between wearing helmet, alcohol impairment and head injury. We believe that as far as motorcyclist's head injured is concerned in RTA, Taiwan road traffic police is enough professional to record crashes much accurate details. To Taiwan and specifically the city of Changhua, we recommend that signalization signs should be placed where they do not exist or work properly. Motorcyclists should be aware about dangers and safety measures at intersections of rural, urban and county roads. Furthermore, impaired motorcyclists should be discouraged even punished if they intend to go the road.

CRediT authorship contribution statement

Mena Lao: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **T.M Lukusa:** Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Ministry of Science and Technology (Taiwan) Grant No. 109-2118-M-006-012-MY2.

References

Agresti, A., 2013. An introduction to categorical data analysis, third ed. Wieley, New York.
Borucka, A., 2020. Logistic regression in modeling and assessment of transport services. Open Eng. 10, 26–34.
Chen, P.L., Pai, C.W., 2019. Evaluation of injuries sustained by motorcyclists in approach-turn crashes in Taiwan. Accid. Anal. Prevent. 124, 33–39.
Chiu, W.T., Kuo, C.Y., Hung, C.C., Chen, M., 2000. The effect of the Taiwan motorcycle helmet use law on head injuries. Am. J. Public Health 90 (5).
Dobson, A.J., Barnett, A.G., 2008. An Introduction to generalized Linear Models, third ed. Chapman and Hall-CRC, Boca Raton.
Dos Santos, W.J., Da Silva Coêlho, V.M., Do Bonfim, C.V., Costa, Da., de Ceballos, A.G., 2019. Alcohol and risky behavior in traffic among motorcyclists involved in accidents in a city in northeastern Brazil. Traffic Inj. Prev. 20.
Friendly, M., 1994. Mosaic displays for multi-way contingency tables. J. Am. Stat. Assoc. 89190–89200.
Ghamdi, A.S., 2002. Using logistics regression to estimate the influence of accident factors on severity. Accid. Anal. Prev. 34, 729–741.
Yeh, T.H., 2011. Changes in traffic safety policies and regulations in 7 countries (1950-2010), Taiwan. International Association of Traffic and Safety. Sciences 73–114.
Hartigan, J.A., Kleiner, B., 1981. Mosaics for contingency tables. In Eddy, W.F., editor, Comp. Sci and Stat.: Proceedings of the 13th Symposium on the Interface, 268, 273. Springer-Verlag, New York, NY.
Hsu, T.P., Nguyen, X.D., Sadullah, M.A.F., 2003. A Comparative study on motorcycle traffic development of Taiwan, Malaysia and Vietnam. J. East. Asia Soc. Transp. Studies (EASTS).
Ichikawa, M., Chadbunchachai, W., Marui, E., 2003. Effect of the helmet act for motorcyclists in Thailand. Accid. Anal. Prev. 35, 183–189.
Iamtrakul, P., Moinul, H., 2007. Motorcycle accidents, alcohol intoxication and futile helmet legislation in Thailand: How long must we tolelare? J. East. Asia Soc. Transp. Stud. 7.
Jou, R.C., Yeh, T.H., Chen, R.S., 2012. The Effects of helmet use on the severity of head injuries in motorcycle accidents. Traffic Inj. Prev. 13, 155–162.
Kasantikul, V., Ouellet, J.V., Smith, T., Sirathranont, J., Panichabhongse, V., 2005. Studied the role of alcohol in Thailand motorcycles crashes. Accid. Anal. Prev. 37, 357–366.
Karacasua, M., Ergül, B., Yavuz, A.A., 2014. Estimating the causes of traffic accidents using logistic regression and discriminant analysis. Int. J. Inj. Contr. Saf. Promot. 21, 305–312.
Keng, S.H., 2005. Helmet use and motorcycle fatalities in Taiwan. Accid. Anal. Prev. 37, 349–355.
Hosmer, D.W., Lemeshow, S., 1980. A Goodness-of-fit tests for the multiple logistic regression model. Commun. Stat. 10, 1043–1069.

Lin, M.R., Kraus, J.F., 2009. A review of factors and patterns of motorcycle injuries. Accid. Anal. Prev. 41, 710–722.

Lin, H.Y., Li, J.S., Pai, C.W., Chien, W.C., Huang, W.C., Hsu, C.W., Wu, C.C., Yu, S.H., Chiu, W.T., Lam, C., 2022. Environmental factors associated with severe motorcycle crash injury in university neighborhoods: a Multicenter study in Taiwan. Int. J. Environ. Res. Public Health 19, 10274.

Liu, B.C., Ivers, R., Norton, R., Boufous, S., Blows, S., Lo, S.K., 2008. Helmets for preventing injury in motorcycle riders. Cochrane. Db. Syst. Rev. 23, CD004333.

Mayrose, J., 2008. The effects of a mandatory motorcycle helmet law on helmet use and injury patterns among motorcyclist fatalities. J. Safety Res. 39, 429–432.

Ministry of transportation and communication, Taiwan, R.O.C. Traffic statistics of the year. 2015. Annual Transportation Repot.

Nguyen, H.H., 2013. A comprehensive review of motorcycle safety situation in Asian countries. J. Soc. Transp. Traffic Stud. 4, 20–29.

Pai, C.W., Lin, H.Y., Tsai, S.H., Chen, P.L., 2018. Comparison of traffic-injury related hospitalisation between bicyclists and motorcyclists in Taiwan. PLoS ONE 13, e0191221.

Pai, C.W., Chen, Y.C., Lin, H.Y., Chen, P.L., 2017. A population-based case-control study of hospitalisation due to head injuries among bicyclists and motorcyclists in Taiwan. BMJ Open 7 (11), 1–11.

Ramli, R., Siang, N.L., Chi, N.F., Rahman, N.A., Oxley, J., 2011. Helmet use amongst injured and non-injured motorcyclists in Malaysia. J. Australas. Coll. Road Saf, 22 No.2.

Tay, R., Shakil, R., Hoong, C., 2008. A logistic model of the effects of roadway, environmental vehicle, crash and driver characteristics on hit and run crashes. Accid. Anal. Prev. 40, 1330–1336.

Weiss, A.A., 1992. The Effects of helmet use on the severity of head injuries in motorcycle accidents. J. Am. Stat. Assoc. 87, 48–56.

Wiznia, D.H., Kim, C.Y., Dai, F., Goel, A., Leslie, M.P., 2016. The effect of helmets on motorcycle outcomes in a Level I trauma center in Connecticut. Traffic Inj. Prev. 17, 633–637.

World Health Organization, 2015. Road safety in the south-east Asia region. [cited 2020 May 16]. Available from: http://www.who.int/violence_injury_prevention/road_safety_status/2015/Road_Safety_SEAR_3_for_web.pdf.