Digital Health and Informatics Innovations for Sustainable Health Care Systems J. Mantas et al. (Eds.) © 2024 The Authors. This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).

doi:10.3233/SHTI240725

Multiple Logistic Regression Model for Assessing the Risk Factors of Traffic Accidents: Khon Kaen Model

Patcharawan SUJAYANONT^{a,b}, Wutjanun MUTTITANON ^c, Yann CHEMIN ^d, Jaturong SOM-ARD ^e and Nantawan TIPPAYANATE ^{a,f,1}
^a Faculty of Medicine, Mahasarakham University, Thailand
^b Tropical Health Innovation Research Unit, Faculty of Medicine, Mahasarakham University, Thailand
^eDepartment of Civil and Environmental Engineering, Mahidol University, Thailand
^dDepartment of Algorithmic Sciences, Seilio Douar, Plumergat, France
^eDepartment of Geography, Mahasarakham University, Thailand
^fPublic Health and Environmental Policy in Southeast Asia Research Cluster (PHEP-SEA), Thailand
ORCiD ID: Patcharawan Sujayanont https://orcid.org/0009-0005-2502-5807, Wutjanun Muttitanon https://orcid.org/0009-0005-2502-5807, Wutjanun

https://orcid.org/0000-0001-9232-5512, Jaturong Som-Ard https://orcid.org/0000-0002-5868-1860, Nantawan Tippayanate https://orcid.org/0000-0003-2376-5666

Abstract. Background: Thailand has consistently held the highest global ranking in traffic accidents since 2017, with Khon Kaen displaying the highest mortality rate in the Department of Disease Control Region 7. Objectives: This study aims to utilize Injury Surveillance (IS) data to identify risk factors associated with emergency room (ER) outcomes at the Emergency Department of Khon Kaen hospital in Khon Kaen Municipality. Methods: Data from the Injury Surveillance system's (IS system) database were collected, focusing on severity outcomes, time of events, and risk behaviors from January 1, 2008, to December 31, 2021. Data analysis was conducted using the R program, employing the Chi-square or independent T test to compare results and analyze associations between potential risk factors and ER outcomes. Multiple logistic regression (MLR) was used for classification analysis, and a confusion matrix was applied to evaluate the performance of the models. Results: MLR analysis revealed that being male, age, alcohol consumption, and nighttime driving were more likely to increase the probability of severity outcomes. Conclusion: Being male, age, alcohol consumption, and nighttime driving are identified as potential risk factors contributing to the development of severity outcomes following traffic accidents.

Keywords. Traffic accidents, Emergency room visits, Risk factors, Logistic models

1. Introduction

Traffic accidents present a significant global health challenge, leading to a substantial number of deaths and injuries. According to the World Health Organization (WHO), 1.3

¹ Corresponding Author: Nantawan Tippayanate; E-mail: nantawan.t@msu.ac.th.

million people worldwide succumb to traffic accidents annually, while an additional 20-50 million individuals suffer injuries or disabilities [1]. In 2017, Thailand ranked first globally in road accidents, with a rate of 36.2 accidents per 100,000 population [2]. This figure notably exceeds the WHO's goal of limiting road accident deaths to 10 per 100,000 population [3]. Despite continuous efforts, combining laws and various community education initiatives to promote traffic safety, it is not evident that the problem's trend has seen significant success.

In the Northeastern region, data from the Department of Disease Control Region 7, covering the provinces of Khon Kaen, Roi Et, Maha Sarakham, and Kalasin, reveals that during 2012-2016, Khon Kaen province exhibited the highest mortality rate at 18.21 per 10,000 population [4]. Serving as a trauma and critical care hospital and a WHO collaborating center for injury prevention and safety promotion, Khon Kaen Hospital provides a database of Injury Surveillance (IS system) indicating the severity of injuries, risk behaviors, time of injuries, and type of vehicle involved. So far, the information has been utilized individually, but the combined influence of multiple risk factors on severity outcomes has not yet been thoroughly investigated. To maximize the potential benefits of the data and, if feasible, provide insights on how to reduce the likelihood of severe outcomes, this research aims to examine IS data to identify risk factors associated with emergency room (ER) outcomes at the Emergency Department.

2. Methods

2.1. Data collection

This study design is a cross-sectional analytical study that focuses on individuals injured in traffic accidents for 14 years in Khon Kaen Municipality (from January 1, 2008, to December 31, 2021). The data were collected from the IS system, including the severity of injuries, risk behaviors, time of injuries, and the type of vehicle involved in the accident. Injuries characterized by hospital admission and death at the ER were classified as the severe group, while those with less severe outcomes were classified as the nonsevere group. The data collection process adhered to the principles of the Declaration of Helsinki and obtained approval from the Ethical Review Committee for Human Research at Mahasarakham University (No. 207-234/2566) and Khon Kaen Hospital (No. 66045).

2.2. Data analysis

Following data collection, the dataset's completeness was verified, and any item(s) with unidentified data were excluded from the study. The analysis was performed using R version 4.1.0. Descriptive statistics were used to portray the distribution of variables. The association between potential risk variables and the Emergency Room (ER) outcome was assessed using the chi-square test or independent T test. Variables with a p-value below 0.05 in the bivariate analysis were considered statistically associated with the ER outcome and subsequently included in the multivariable analysis, utilizing the multiple logistic regression (MLR) model and adjusted odds ratio. Model performance was measured for accuracy, precision, recall, F1 score, area under the curve (AUC), sensitivity and specificity.

3. Results

3.1. Risk associated traffic accidents at intersections

Total number of traffic accident victims collected from the IS database in Khon Kaen Municipality amounted to 19,877 after filtering out incomplete data. To assess the correlation between possible accident associated variables and ER outcomes, statistical differences were analyzed using the chi-square test or independent T test (Table 1). The results suggested that sex, age group, time interval, type of injury person, type of vehicle and alcohol consumption of the victims were identified as variables that could influence the severity outcomes.

Description		Frequency (%)		
		Discharge	Admit/death	OR (95% CI)
		(14,744)	(5,133)	
Age (mean \pm SD)		27.58 ± 13.49	30.79 ± 15.13	p value < 2.2e-16
Sex	1: Male	8065 (54.70%)	3582 (69.78%)	- 0.52 (0.48,0.55)
	2: Female	6679 (45.3%)	1531 (29.83%)	
Time interval	T1: 6.00-11.59 AM	3505 (23.77%)	926 (18.04%)	-
	T2: 12.00-5.59 PM	4310 (29.23%)	932 (18.16%)	0.82 (0.74,0.91)
	T3: 6.00-11.59 PM	4098 (27.79%)	1413 (27.53%)	1.31 (1.19,1.44)
	T4: 12.00-5.59 AM	2831 (19.2%)	1842 (35.89%)	2.46 (2.24,2.71)
Type of injury person	1: Pedestrian	607 (4.12%)	255 (4.97%)	-
	2: Driver	11684 (79.25%)	4313 (84.02%)	0.88 (0.75,1.03)
	3: Passenger	2453 (16.64%)	545 (10.62%)	0.53 (0.44,0.63)
Type of vehicle	0: Pedestrian	607 (4.12%)	255 (4.97%)	-
	1: Helmet used type	13454 (91.25%)	4655 (90.69%)	0.82 (0.71,0.96)
	2: Seatbelt used type	453 (3.07%)	132 (2.57%)	0.69 (0.54,0.89)
	3: Other	230 (1.56%)	71 (1.38%)	0.74 (0.53,1)
Alcohol use	0: Not used	12127 (82.25%)	2856 (55.64%)	- 3.66 (3.42,3.93)
	1: Used	2617 (17.75%)	2257 (43.97%)	
Accident	1: Road section	10134 (68.73%)	3502 (68.23%)	-
spot	2: Signalized intersection	3040 (20.62%)	1025 (19.97%)	0.98 (0.9,1.06)
	3: Unsignalized intersection	1570 (10.65%)	586 (11.42%)	1.08 (0.97,1.2)

Table 1. The association between potential risk factors and ER outcome in Khon Kaen Municipality from January 1, 2008, to December 31, 2021

3.2. MLR model of traffic accidents at intersections

In this study, MLR in R was employed to model the differences in severity outcomes at the ER. The independent variables with statistically significant differences, as collected from the results in Table 1, were incorporated using the generalized linear model (glm) function. Equation (1) and table 2 present the results of the multivariate logistic regression model and the corresponding adjusted odds ratios for traffic accidents during the period from January 1, 2008, to December 31, 2021.

 $logit(severity) = -1.742994 - 0.298946 \times female + 0.021263 \times age - 0.453202 \times driver - 0.710173 \times passenger + 0.242581 \times helmet used type - 0.189538 \times seatbelt used type + 1.026858 \times alcohol - 0.160785 \times T2 + 0.133110 \times T3 + 0.438858 \times T4$ (1)

Description	intercept	sex (female)	age	injp (driver)
OR (95% CI)	0.17 (0.14, 0.21)	0.74 (0.69, 0.80)	1.02 (1.02, 1.02)	0.64 (0.46, 0.87)
Description	injp (passenger)	injt (helmet used)	injt (seatbelt used)	alcohol
OR (95% CI)	0.49 (0.35, 0.68)	1.27 (0.96, 1.69)	0.83 (0.58, 1.17)	2.79 (2.56, 3.04)
Description	T2: 12.00-5.59 PM	T3: 6.00-11.59 PM	T4: 12.00-5.59 AM	
OR (95% CI)	0.85 (0.77, 0.94)	1.14 (1.03, 1.26)	1.55 (1.39, 1.73)	

Table 2. Adjusted odds ratio of potential risk factors and ER outcome in Khon Kaen Municipality from January 1, 2008, to December 31, 2021

3.3. Accuracy, precision, recall, F1 score and AUC of the model

To evaluate the logistic model's performance, a confusion matrix was calculated using the caret package in R. The results indicate that the accuracy, precision, recall, F1 score, AUC, sensitivity, and specificity are 73.69%, 80.47%, 85.27%, 82.80%, 69.31%, 85.27%, and 40.31%, respectively.

4. Discussion

4.1. Risk associated traffic injuries

The investigation of traffic accident injuries in Khon Kaen Municipality was conducted to observe risky behaviors during traffic accidents, which emerged as a significant factor. The results found that risky behaviors accounted for 25.82% of severity grading (Table 1). The findings revealed that females tend to experience less severe outcomes compared to male victims, as also found in another study [5]. Similarly, in line with other reports [6,7], alcohol consumption emerged as a major risk factor in traffic crashes. The results indicated that the highest incidence occurred during time interval 4 (12:00-5:59 AM). This differs from peak hours in Istanbul, where accidents usually happen between 07:00-09:00 and 17:00-19:00, coinciding with increased traffic volume and accident rates [8], suggesting that accidents in Khon Kaen Municipality tend to occur during nighttime later than during the daytime. Regarding the age of the victims, this study found that the average age of those with more serious injuries was higher than those with minor injuries. As age increases, there is a higher probability of the outcome from a road accident being more serious, in parallel with other studies [5,9]. The results also showed that pedestrians are more likely to have serious outcomes compared to car occupants, as also found in another study [10].

4.2. The association of multiple variates in ER outcome

The results in Table 2 revealed that while being female, being a car occupant, and daytime accidents reduced the probability of a serious outcome, factors such as being male, age, alcohol consumption, and nighttime driving might increase this probability, consistent with findings from other studies [6-9]. To evaluate the validity of the logistic model, a confusion matrix was created, showing that the accuracy, precision, recall, F1 score, AUC, sensitivity, and specificity are 73.69%, 80.47%, 85.27%, 82.80%, 69.31%, 85.27%, and 40.31%, respectively. It might be inferred that this model has acceptable performance, as it provides an acceptable range of accuracy, precision, recall, F1 score, AUC, and sensitivity, but low specificity, resulting in poor differentiation of victims with

more serious outcomes. This finding suggests limitations in analyses involving imbalanced and missing datasets.

5. Conclusions

In conclusion, the MLR model identified being male, age, alcohol consumption, and nighttime driving as potential risk factors influencing an increase in severity outcomes in the ER. However, future studies should consider incorporating additional variables or conducting a prospective study to avoid overlooking important confounding factors. Addressing dataset imbalances and missing data to enhance the model's predictive capabilities. This will lead to more effective interventions aimed at preventing serious outcomes in car accidents in the future.

Acknowledgement

The authors would like to thank the Public Health and Environmental Policy in Southeast Asia Research Cluster (PHEP-SEA), the Faculty of Medicine at Mahasarakham University, and Mahasarakham University, Thailand, for providing funding for this research. Additionally, sincere appreciation is expressed to Ms. Natthira Daengpruan and The Trauma and Critical Care Center at Khon Kaen Hospital for their invaluable assistance in making the data accessible.

References

- Global status report on road safety 2018 [Internet]. [cited 2023 Sep 7]. Available from: https://www.who.int/publications-detail-redirect/9789241565684
- [2] Countries With the Most Car Accidents. WorldAtlas. 2018 [Internet]. [cited 2023 Dec 20]. Available from: <u>https://www.worldatlas.com/articles/the-countries-with-the-most-car-accidents.html</u>
- [3] World Health Organization. Global Status Report on Road Safety 2015, Supporting a decade of action. Switzerland; 2015.
- [4] Injury Data Collaboration Center Division of Injury Prevention. (2022). Situation of fatal road traffic accident in Thailand [Internet]. [cited 2023 Mar 7]. https://dip.ddc.moph.go.th/new/%E0%B8%9A%E0%B8%A3%E0%B8%B4%E0%B8%81%E0%B8% B2%E0%B8%A3/3base status new (In Thai).
- [5] Razi-Ardakani H, Ariannezhad A, Kermanshah M. A Study of Sex Differences on Road Crash Severity. In Proceedings of the 3rd International Conference on Civil, Structural and Transportation Engineering (ICCSTE'18); 2018 Jun 10-12; Niagara Falls, Canada, doi: 10.11159/iccste18.113
- [6] Valen A, Bogstrand ST, Vindenes V, Frost J, Larsson M, Holtan A, et al. Driver-related risk factors of fatal road traffic crashes associated with alcohol or drug impairment. Accid Anal Prev. 2019 Oct;131:191-199.
- [7] Tokko T, Eensoo D, Vaht M, Lesch KP, Reif A, Harro J. Relapse of drunk driving and association with traffic accidents, alcohol-related problems and biomarkers of impulsivity. Acta Neuropsychiatr. 2019 Apr;31(2):84-92.
- [8] Fatal motorcycle accidents in Fars Province, Iran: a community-based survey. Chin J Traumatol. 2012;15(4):222-7.
- [9] Casado-Sanz N, Guirao B, Gálvez-Pérez D. Population ageing and rural road accidents: Analysis of accident severity in traffic crashes with older pedestrians on Spanish crosstown roads. Res Transp Bus Manag. 2019; 30: 100377.
- [10] Shu CC, Dinh M, Mitchell R, Balogh ZJ, Curtis K, Sarrami P, Singh H, Levesque JF, Brown J. Impact of comorbidities on survival following major injury across different types of road users. Injury. 2022 Oct;53(10):3178-3185.