The Role of Interaction-Based Effects on Fatal Accidents Using Logic Regression

Marzieh Rohani-Rasaf, Yadollah Mehrabi[1], Seyed Saeed Hashemi-Nazari[1,2], Mehdi Azizmohammad Looha[3], Hamid Soori[2]

Student Research Committee, Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, [1]Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, [2]Safety Promotion and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, [3]Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ORCID:
Marzieh Rohani-Rasaf: https://orcid.org/0000-0002-7945-7542
Yadollah Mehrabi: https://orcid.org/0000-0001-9837-4956
Seyed Saeed Hashemi-Nazari: https://orcid.org/0000-0002-0883-3408
Mehdi Azizmohammad Looha: https://orcid.org/0000-0002-0700-1431
Hamid Soori: https://orcid.org/0000-0002-3775-1831

Abstract

Background and Objectives: Road traffic accidents (RTAs) were estimated to be the eighth major cause of death worldwide in 2016. Investigation of various factors alone can distort the results. Thus, it is important to consider interactions among the various factors associated with RTAs. Logic regression was used to investigate the important combinations among traffic accident variables. **Methods:** In this analytical study, the existing 1-year data from the police accident database in 2014 were examined. The Legal Medicine Organization database was also used to correct death after 30 days. Logic regression, a generalized regression model, was used to explore the interactions among different factors of the accident. **Results:** Cross-validation results showed the best model in the form of three trees and eight leaves. Being a professional driver and exposure to a heavy vehicle on sandy or earthy road double the chance of death. Operating an unsafe car on a road with curve increases the odds of a fatal crash by 1.65 times. Driver error on a nonresidential road without any shoulders adds 90% to the odds of having a deadly crash. **Conclusions:** The significance of the interactions between the road and driver factors shows that roads with poor design can cause a driver to make mistakes and increase fatal accidents. Therefore, politicians must consider constructing structures alongside nonresidential roads and proper shoulders, install signs at curves, and repair pavement in order to reduce the fatality of accidents. It is also recommended that manufacturers of commercial vehicles install proper safeguards in all heavy vehicles to reduce fatal accidents.

Keywords: Interactions, logic regression, road traffic accidents

INTRODUCTION

Despite numerous studies, road traffic accidents (RTAs) remain a serious public health concern. These were the eighth major cause of death worldwide in 2016.[1] Road traffic injury is among the five main global causes of disability-adjusted life years.[2] Most RTA deaths and disabilities occur in developing countries.[3-6] The global death rate of RTAs per 100,000 people and the RTA death rate of the Eastern Mediterranean region are 18.2 and 18, respectively. Traffic accidents are a serious problem in Iran with a mortality rate of 20.5.[1] However, because they are often predictable, deaths and serious injuries due to road accidents can be prevented. Analysis of RTAs is a complex process. As yet, the relationship between RTA risk factors and fatality has not been established. One of the

reasons for this is that many factors are at play, such as the characteristics of the people involved, the roads, the vehicles, and the environment.[7,8] Identification of these components and their interactions is a basic prerequisite for determining the causes of any accident. Although the most common cause of accidents is human error,[9-11] successful reductions in death rates in high-income societies have come about by the creation of safer infrastructure, increased vehicle safety,

Address for correspondence: Prof. Hamid Soori,
Department of Epidemiology, Safety Promotion and Injury
Prevention Research Center, School of Public Health and Safety,
Shahid Beheshti University of Medical Sciences, Tehran, Iran.
E-mail: hsoori@yahoo.com

Access this article online

Quick Response Code:

Website:
www.archtrauma.com

DOI:
10.4103/atr.atr_105_18

and the implementation of a number of well-established interventions.[12] However, there is no one single standard intervention that can address the needs of every country in reducing and controlling its RTAs. This is a complex issue which requires a proper analysis of the specific factors involved.

It has long been assumed that the occurrence of RTAs and fatalities will differ for different drivers, vehicles, and environments. Therefore, the study of any one of these factors alone can produce distorted conclusions about their actual roles, which have not been properly identified, and this ultimately leads to inappropriate policies.[13] The interaction effects between variables in regression models cannot exceed two or three ways due to the corresponding complexity.[14] When the number of predictor variables is high, especially when these variables are binary variables, higher interaction levels can affect the fit of the model and make analysis impractical.[15] Logic regression provides a solution to these problems. To determine the precision of an interaction and to identify and consider n-sided interactions in regression models, a combination of variables can be used rather than a set of single variables when fitting the model. They can then be introduced into the model as a new independent variable.[14] No study was found on the prediction of fatal accidents by logic regression. In this article, a model was presented using logic regression which analyzed the interactions of the variables to identify factors related to fatal accidents.

METHODS

This analytical study was conducted to examine the existing 1-year data on light vehicle crashes in Iran. The suburban crash criteria included that the collision types were among the following: vehicle-to-vehicle collisions, multivehicle collisions, vehicle–fixed object collisions, overturning of the vehicle, veering off the road, multiple collisions, passenger ejections, and collisions with a parked vehicle.

Data source

Available data were extracted from the Police Accident Database in 2014. The traffic accident data registry system in Iran was compiled using "Com114" forms filled out by the police at the accident site. In the police database, information on deaths is limited to those occurring at the scene of the accident. Adjustments are made to these data using Legal Medicine Organization database to record deaths that occurred up to 30 days after. The common variables in the two datasets included name, gender, age, accident date, and accident location. In total, outcomes from the two data sources were combined into a database containing 83,235 vehicles involved in accidents and 2821 fatal crashes. Where any individual (driver or occupant) has died, then the outcome is considered fatal. The data included driver attributes (e.g., age, profession, education, driver error, and gender), vehicle features (e.g., vehicle's safety rating by Euro New Car

Assessment Program (NCAP), exposure to heavy vehicle, and color), road characteristics (e.g., visibility barrier, area type, area used, road type, surface conditions, location geometry, pavement condition, and the presence of shoulders), and the environment including weather condition and level of daylight. Variable descriptions and different categories are illustrated in Table 1. The outcomes consist of accident fatality (death, injuries, and property damage).

Statistical methods

Logic regression with logistic link function and deviance score was used to explore the interactions among different factors of the accident. A generalized regression model was proposed by Ruczinski *et al*.[14] to create new and better predictors for outcomes through assembling Boolean compounds of binary variables.

Logit $P (Y = 1 \mid L_1, ..., L_p) = \beta_0 + \beta_1 L_1 + ... + \beta_p L_p$

Each L is a Boolean combination of binary variable X_j as "Logic Terms." The way to denote logic term by using the operators $\vee$ ("or") and $\wedge$ ("and") and the superscript c refers to the complement ("not"). An example for a Logic Term is:

$L = ([X_1 \vee X_2^c] \wedge X_3)$, where 1 equals "L is True," 0 equals "L is False."

The search algorithm used in this method to find the best candidates for the logic term L_p is the simulated annealing algorithm, a stochastic optimization algorithm that is preferred for searching all possible spaces of such variables relative to other methods and for the presentation of the score function, with the ability to compare the adequacy of certain models. The characteristics of this algorithm depend on the Markov chain theory. For each step in a Markov chain, a possible move is randomly proposed. This move is always accepted if the new logic tree attains a better score than the old one; otherwise, it is accepted with a probability based on deviance and the stage of the chain. The score or deviance for logistic function is $-2\log(\text{likelihood})$. The "logreg" package in R software defines a score as a function that reflects the quality of the model under consideration; the best score being the single lowest score seen in any iteration.

In order to assess whether there is any association between response and predictors, "null permutation test" can be used. This is a histogram in which the scores obtained from the randomization permutation procedure were compared with the overall best scoring model given in the real data [left bar in Figure 1] and the score for the null model (right bar) fitting the intercept without any predictors. The null hypothesis is that there is no association between the predictor X and the response Y. If the null hypotheses were to be true, then the randomly permuted histogram should yield about the same score as the overall best model.

Model selection tool

In the logic regression, larger models have a better score than smaller ones; hence, to avoid overfitting, a cross-validation

approach was used for the model selection. This approach divides the data into *m* groups of equal size and extract Group i, finding the best score for k-size model using the m-1 data group and scoring the data using group *i* under this model. The score is called εki. The cross-validation produces a single deviance score for each model with the average of εki. Finally, the model with the smallest average deviance is selected. It is possible to compare various scores for a model with different sizes. In the present study, ten-fold cross-validation and an alternative approach called the training/test set were used to select the model with the best score. Data were randomly divided into ten subsets. One subset was used as a test set and the remaining subset as training data. Training data were used to select the best model, and then the test data were used to estimate the deviance. This procedure was repeated ten times, and then deviance scores were averaged. At the end, the optimal tree size with the smallest average deviance was obtained.

Link of two datasets (Police Accident Database and Legal Medicine Organization) for corrections of outcome was performed through Microsoft Excel. The missing data were subjected to a multiple-imputation chained-equation (MICE) model, using "mice" package in R software (version 3.4.1). In addition, logic regression was implemented using "LogicReg" package in R software.

RESULTS

There were 2821 fatal crashes in the suburban of Iran. The frequency and percentage of the outcome on variables used in the analyses are illustrated in Table 1. It represents a significant relationship between fatality and all factors except gender, weather condition, and visibility barrier (Chi-square test, $P < 0.05$).

To investigate signal in the data, a null model randomization test was used. Null permutation test fitted one tree and up to four leaves. A small model size results in little noise because the algorithm performs a wide search of all models. The null model suggested the model of size 0 which had a score of 24,641. Using simulated annealing, the best scoring model for this size had a score of 24,358. The randomization procedure was repeated 500 times. The histogram of the scores obtained from the randomization procedure was compared to the score of the best model and the score for the null model. Since all those scores were higher than the best score on the original data, we safely concluded that there was information in the predictors for the prognosis of a fatal accident [Figure 1].

The model with a greater number of trees and leaves has a better deviation in training score. This may be due to overfitting data; hence, cross-validation in the test set should be considered. Model with the smallest test set deviance is the best model size. We initially chose the model of 1–5 trees. Adding tree 2 to the model had a great influence on the score in the two training and test scores of cross-validation; hence, tree 1 was removed. Furthermore, trees 4 and 5 had no major difference with tree 3. Finally 2–3 trees and 3–10 leaves were

selected as the complexity model. The logic regression model was fitted, and the best model was selected and evaluated. Boolean combinations were then determined based on the best model size by simulated annealing algorithm.

The test set scores of cross-validation offered three trees and five leaves. Thereafter, the scores were not significantly different, until three trees and eight leaves showed a lower score. Hence, the best model size was three trees and eight leaves. The plot shows the cross-validation test set deviance ("test score") for models with a specific number of logic trees (numbers in squares) and total number of leaves ("model size") [Figure 2].

Finally, after searching for the models with these sizes using simulated annealing algorithm, the best model with 24,081 deviance score, which had the lowest possible score, was fitted. Figure 3 displays the results of the best model in the form of Logic Tree. The first tree, L_1, was unpaved road (X_1) OR exposure to heavy vehicle (X_2) OR professional driver $(X_3)^c$. This tree (L_1) indicated that a professional driver exposed to a heavy vehicle on unpaved road doubles the chance of death.

The second tree, L_2, involved a combination of unsafe car (X_4) AND road with curve (X_5). This tree suggested that operating an unsafe car on a road with curve had 1.65 times odds of fatal crash.

The third tree, L3, is determined by a combination of driver error (X_6) AND nonresidential road $(X_7)^c$ OR road without a shoulder X_8. This tree indicated that driver error on a nonresidential road without a shoulder led to 90% increase in odds of death [Table 2].

$$\text{Logit}\left(P\left[\begin{array}{l}\text{fatal accident} \mid \text{road, environment,}\\ \text{vehicle and driver risk factors}\end{array}\right]\right) =$$

$$-2.58 + 0.707\left(x_1 \vee x_2\right) \vee \left(x_3\right)^c$$

$$+ 0.493\left(x_4 \wedge x_5\right) + 0.641\left(\left[x_6 \wedge \left(x_7\right)^c \vee x_8\right]\right)$$

DISCUSSION

The primary purpose of this study was to investigate the interaction of drivers, vehicles, and roads in relation to fatal vehicle accidents in Iran, using logic regression. The results revealed a logical combination of all the three factors, the interaction of vehicle and road, and a complementary pattern of driver and road. The present study is the first to use logical regression as a substitute for the conventional model in identifying road accident interactions. Using appropriate Boolean compositions in the model instead of the main variables allows for the identification of possible interactions. It also removes the concern arising from the higher levels of interactions in data with a large number of binary variables.

One limitation of this study is that some of the factors that might be related to the fatality of an accident are not included in the analysis. Another limitation is "unknown" data.

Many studies have identified human errors as the main cause (75%–90%) of accidents.[16] The most frequent human

Table 1: Variable categories and summary table of main variables and outcomes

Various factors	Variables	Categories*	Frequency** (%)	Outcome frequency (%)	P
Vehicle factors	Exposure to heavy vehicle	Exposed	10,995 (13.2)	626 (5.69)	<0.001
		Nonexposed	72,240 (86.8)	195 (3.04)	
	Safety features	<3 stars	71,645 (86.08)	2534 (3.54)	<0.001
		3 stars and more	11,590 (13.92)	287 (2.48)	
	Color	Dark color	29,978 (36.02)	1856 (3.22)	0.042
		Light color	53,257 (63.98)	965 (3.48)	
Driver factors	Driver error	With error	58,536 (70.33)	2244 (3.83)	<0.001
		Without error	23,400 (28.11)	574 (2.54)	
	Age	<40, >65 years	53,213 (63.93)	1706 (3.21)	0.008
		40-65 years	28,124 (33.79)	1001 (3.56)	
	Education	Diploma and less	64,298 (77.25)	1895 (2.95)	<0.001
		Higher than diploma	6973 (8.38)	145 (2.08)	
	Professional driver	Not being driver as an occupation	62,666 (75.25)	1447 (2.31)	<0.001
		Being driver as an occupation	239 (0.29)	16 (6.69)	
	Gender	Male	78,262 (94.03)	2672 (3.41)	0.058
		Female	4882 (5.87)	142 (2.91)	
Road factors	Area type	Hilly	9437 (11.34)	410 (4.34)	<0.001
		Flat	73,151 (87.88)	2376 (3.25)	
	Area use	Residential	28,601 (34.36)	762 (2.66)	<0.001
		Nonresidential	54,081 (64.97)	2035 (3.76)	
	Road type	Main road	65,402 (78.58)	2156 (3.3)	0.005
		Side road	17,764 (21.34)	662 (3.73)	
	Pavement condition	Paved	436 (0.52)	34 (7.8)	<0.001
		Not paved	82,514 (99.13)	2772 (3.36)	
	Visibility barrier	Barrier	1595 (1.92)	60 (3.76)	0.392
		Not barrier	79,015 (94.93)	2663 (3.37)	
	Surface conditions	Wet	6511 (7.82)	189 (2.9)	0.027
		Dry	74,091 (89.01)	2534 (3.42)	
	Location geometry	Curved	15,070 (18.11)	736 (4.88)	<0.001
		Straight	65,524 (78.72)	1987 (3.03)	
	Shoulder	Presence of shoulder	48,950 (58.81)	1318 (4.16)	<0.001
		Absence	1663 (38.04)	1405 (2.87)	
Environmental factors	Weather condition	Cloudy and foggy	8299 (9.97)	276 (3.33)	0.636
		Sunny weather	72,838 (87.51)	2495 (3.43)	
	Daylight	Darkness, twilight	25,577 (30.73)	972 (3.8)	<0.001
		Daylight	55,560 (66.75)	1799 (3.24)	

*The first category is considered as 1 and another as 0, **Values are based on original data (before imputation). Frequency <83,235 in each variable due to missing value in these data
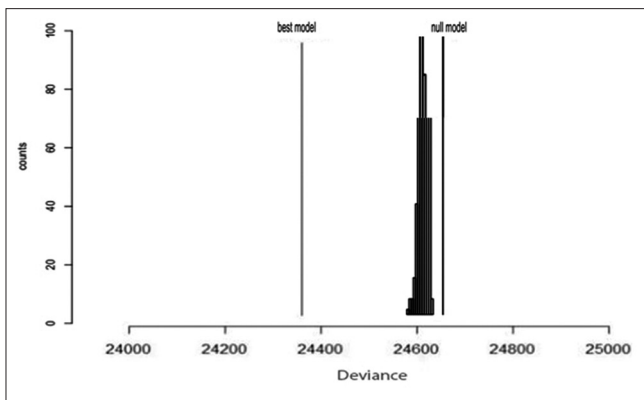

Figure 1: Null model randomization test based on one tree and up to four leaves

error in this study was excessive acceleration (62%). Speed was the most common parameter of accident risk, which was responsible for 30% of fatal accidents.[17] The studies have shown that speed increased the severity of an accident; a 10% increase in speed resulted in a 21% increase in crash severity and 46% of fatal accidents.[18,19] The interactions between the conditions of a road and drivers' behavior are complex. The conditions of cars and roads are highly influential on drivers' behavior before and during accidents, especially in developing countries, where roads and vehicles are not standardized in comparison with those in developed countries.[20] According to a study by Khalili and Pakgohar, 36% of accidents in Iran were related to unsafe roads; in Europe, this proportion was 24%. Road safety failures in

Iran were twice as likely as in Europe. In this study, road factors were present in three trees and had interactions with car or human. Research on roads in suburban areas in Iran showed that the most important factor in reducing road safety and increasing the severity of accidents was level difference between shoulder and road (odds ratio = 1.97). The absence of a shoulder increased the severity by 95% and was a cause of high accident rates.[21]

In the present study, interaction of driver error and nonresidential road increased the odds of a fatal accident, which is in line with a study in India.[22] Because the study is limited to suburbs and vehicle collision types, pedestrians have no role and the nonresidential road was introduced as a risk factor. According to most behavioral models, drivers adapt their behavior to the perception of danger. Risk perception is negatively related to risky driving; hence, risk perception is associated with safety.[23,24] The hypothesis is that, on a nonresidential road, the driver pays less attention to the road and because of absence of pedestrians, the driver relaxes and the speed accelerates. Another hypothesis is to increase the likelihood of fatigue on a nonresidential road. Fatigue reduces capacity and attention, and this reduces the likelihood of maintaining a safe speed.[25]

Insecure car on curve has increased the chance of fatal accidents by 65%. Studies show that the death risk in one-star vehicle doubles as a 5-star vehicle which has passed the Australian NCAP.[10] Studies have shown that the characteristics of the curves are related to accidents,[21,26] and in general, the crashes in the curves are three times more than the direct ones.[27]

Other factors that double the chance of fatal accidents are exposure to a heavy vehicle on unpaved surface and being a professional driver. The results of a study conducted in Iran showed that surface defect is a risk factor for severe accident (OR=1.43).[21] Heavy vehicles can be one of the most common causes of severe accidents.[28,29] Truck loading, long hours behind the wheel, and hazardous behaviors (e.g., driving under the influence of alcohol) are some of the known causes of accidents with heavy vehicles.[30] Heavy-vehicle drivers cannot act at the proper time to prevent an accident.[31] The role of heavy trucks is much more than other car in rear crashes.[16] Drivers seem to be more likely to crash due to longer journeys or fatigue.

Conclusions

The existence of a road factor in all the three interactions indicates the importance of this factor. The significance of the interaction of driver error in nonresidential road and without shoulder may indicate that roads with poor design increase driver error caused by factors such as fatigue and drowsiness and result in fatal accidents. Therefore, politicians must consider constructing structures alongside nonresidential roads, create amenities and special services for travelers along roads, create appropriate shoulders, install

signs at curves, and repair pavements to reduce the fatality of accidents.

The car's factor is a very serious issue, and many existing vehicles need to be equipped with improved safety features. Governments should necessitate that manufacturers of commercial vehicles install proper safeguards in the rear of the heavy vehicle to reduce the fatality of accidents.

Table 2: Results from the fitted logic model

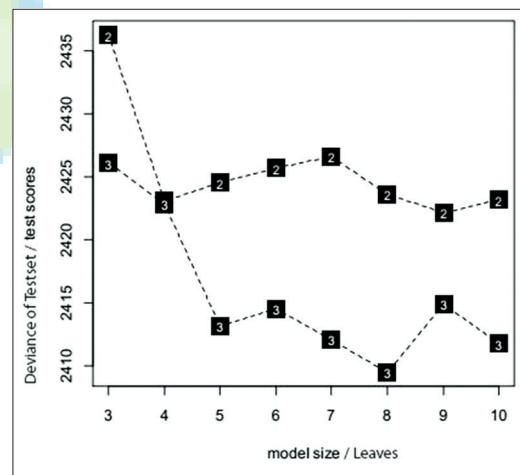| Tree	Boolean trees	OR*	95% CI	P
Tree 1	("Unpaved road" OR "exposure to heavy vehicle") OR "professional driver"	2.03	1.86-2.22	<0.001
Tree 2	"Unsafe car" AND "road with curve"	1.65	1.5-1.79	<0.001
Tree 3	("Driver error" AND "nonresidential road ") OR "road without a shoulder"	1.9	1.73-2.09	<0.001

CI: Confidence interval, *OR: Odds Ratio

Figure 2: The cross-validation test set deviance results for associations of road, vehicle, and driver with fatal accident. Numbers in squares are logic trees and the number of leaves indicates model size
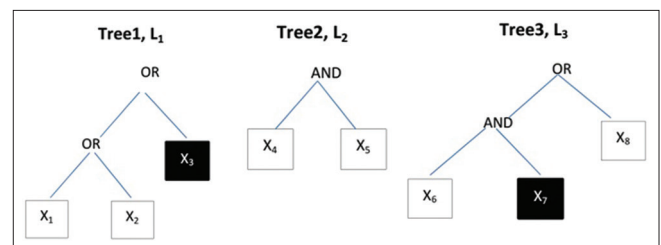

Figure 3: The best model (three trees and eight leaves) for odds of fatal accident

Rohani-Rasaf, *et al*.: Accident analysis

Financial support and sponsorship

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. World Health Organization. Global Status Report on Road Safety 2018. Geneva: World Health Organization; 2018.
2. World Health Organization. Disability-Adjusted Life Years (DALYs); 2018. Available from: http://www.who.int/gho/mortality_burden_disease/daly_rates/text/en/. [Last cited on 2019 Dec 02].
3. Nantulya VM, Reich MR. The neglected epidemic: Road traffic injuries in developing countries. BMJ 2002;324:1139-41.
4. Ameratunga S, Hijar M, Norton R. Road-traffic injuries: Confronting disparities to address a global-health problem. Lancet 2006;367:1533-40.
5. World Health Organisation. WHO Report 2015: Data Tables" (PDF) (Official Report). Geneva: World Health Organisation; 2015.
6. Peden MM, McGee K, Krug E. Injury: A Leading Cause of the Global Burden of Disease, 2000. Geneva: World Health Organization; 2002.
7. Moghaddam FR, Afandizadeh S, Ziyadi M. Prediction of accident severity using artificial neural networks. Int J Civil Eng 2011;9:41.
8. Niezgoda M, Kamiński T, Kruszewski M. Measuring driver behaviour-indicators for traffic safety. J KONES 2012;19:503-11.
9. World Health Organization. The Injury Chart Book: A Graphical Overview of the Global Burden of Injuries. Geneva: World Health Organization; 2002.
10. World Health Organization. Global Status Report on Road Safety 2015. Geneva: World Health Organization; 2015.
11. Naghavi M, Shahraz S, Bhalla K, Jafari N, Pourmalek F, Bartels D, *et al*. Adverse health outcomes of road traffic injuries in Iran after rapid motorization. Arch Iran Med 2009;12:284-94.
12. Peden M, Scurfield R, Sleet D, Mohan D, Hyder AA, Jarawan E, *et al*. World Report on Road Traffic Injury Prevention. World Health Organization: Geneva; 2004.
13. Schepers P, Hagenzieker M, Methorst R, van Wee B, Wegman F. A conceptual framework for road safety and mobility applied to cycling safety. Accid Anal Prev 2014;62:331-40.
14. Ruczinski I, Kooperberg C, LeBlanc M. Logic regression. J Comput Graphical Statistics 2003;12:475-511.
15. Lucek PR, Ott J. Neural network analysis of complex traits. Genet Epidemiol 1997;14:1101-6.
16. Kim YS. Effects of Driver, Vehicle, and Environment Characteristics on Collision Warning System Design: Institutionen för Teknik och Naturvetenskap; 2001.
17. Royal D. National Survey of Speeding and Unsafe Driving Attitudes and Behaviors: 2002. Findings. Vol. 2. United States: National Highway Traffic Safety Administration; 2004.
18. Vägverket PF. Swedish Road Administration. Borlänge, Sweden: Roads and Traffic, Publication. 2006. p. 23E-52.
19. Ratanavaraha V, Suangka S. Impacts of accident severity factors and loss values of crashes on expressways in Thailand. IATSS Res 2014;37:130-6.
20. Islam M, Tanaboriboon Y, editors. Crash Investigation and Reconstruction the New Experience in Developing Countries: Thailand Case Study. Proceedings of the Road Safety on Four Continents Conference; 2005.
21. Khalili M, Pakgohar A. Logistic regression approach in road defects impact on accident severity. J Emerg Technol Web Intell 2013;5:132-5.
22. Bandyopadhyaya R, Mitra S. Modelling severity level in multi-vehicle collision on Indian highways. Procedia Soc Behav Sci 2013;104:1011-9.
23. Plankermann K. Human Factors as Causes for Road Traffic Accidents in the Sultanate of Oman under Consideration of Road Construction Designs. PhD, Universität Regensburg: Philosophy; 2014.
24. Senders JW, Kristofferson A, Levison W, Dietrich C, Ward J. The attentional demand of automobile driving. Bolt, Beranek and Newman Incorporated. Highway Res Record 1967;195:15-33.
25. Williamson A. Fatigue and coping with driver distraction. In: Faulks IJ, Regan M, Stevenson M, Brown J, Porter A, Irwin JD, (Eds.). Distracted driving. Sydney, NSW: Australasian College of Road Safety. 2007. p. 611-22.
26. Al-Ghamdi AS. Using logistic regression to estimate the influence of accident factors on accident severity. Accid Anal Prev 2002;34:729-41.
27. Zegeer C, Twomey J, Heckman M, Hayward J. Safety Effectiveness of Highway Design Features. Washington, DC: FHWA; 1992.
28. Kanchan T, Kulkarni V, Bakkannavar SM, Kumar N, Unnikrishnan B. Analysis of fatal road traffic accidents in a coastal township of South India. J Forensic Legal Med 2012;19:448-51.
29. Mohamed MG, Saunier N, Miranda-Moreno LF, Ukkusuri SV. A clustering regression approach: A comprehensive injury severity analysis of pedestrian-vehicle crashes in New York, US and Montreal, Canada. Safety Sci 2013;54:27-37.
30. Chu HC. An investigation of the risk factors causing severe injuries in crashes involving gravel trucks. Traffic Inj Prev 2012;13:355-63.
31. Islam MB, Kanitpong K. Identification of factors in road accidents through in-depth accident analysis. IATSS Res 2008;32:58-67.