Injury Severity among Young Drivers: A combination of Parametric and Non-parametric Analysis

Rusdi bin Rusli
Department of Civil Engineering,
Politeknik Sultan Mizan Zainal Abidin
Km 08, Jalan Paka,
23000 Dungun, Terengganu
diri65@yahoo.com

ABSTRACT

Young drivers among the group recorded highest fatalities index in road traffic crashes. The objective of this study was to identify factors affecting injury severity of young drivers using 5-years crash data (2008-2012) in Sabah, Malaysia. This study used a combination of parametric and non-parametric analysis to allows the specification of nonlinearities and interactions in addition to main effects. The results indicate that crashes on nighttime, federal roads and involved with a single-vehicle are positively associated with injury among young drivers. Interestingly, municipal roads, female drivers, crashes on the roundabout and T/Y junction are less likely to involve injury. A higher-order interaction suggests that not-at-fault young drivers involved with out-of-control or hit the object are more likely to be severe. On the other hand, young passenger car drivers involved in overturn and sideswipe collisions are negatively associated with the injury. It was also found that young drivers with driving too close behaviour are less likely to injure when involved in rear-end collisions for passenger car and four-wheel drive. Findings of this study will help relate authorities to design well-targeted restrictive measures in reducing the severity level of young drivers in traffic crashes.

Keywords : Young Drivers, Malaysia, Injury severity, Parametric and Non-Parametric analysis.

INTRODUCTION

Road traffic crashes are a leading cause of death especially to people aged 5-29 years. In 2018, the World Health Organization (WHO) has declared road traffic crashes are a number one killers to this group (WHO, 2018). In Malaysia, this group represents 30% of total fatalities in a road traffic crash in the year 2013 (JKJR, 2014). Specifically, in Sabah, drivers aged less than 25 years recorded the highest proportion involved in road traffic crashes and contributed to more than half of fatalities (see Table 1). This issue needs to be explored further in avoiding the loss of valuable human resource for future development.

Research on young drivers is not something new in road safety (e.g. Dissanayake and Lu, 2002; Huang and Winston, 2011; Ismail *et al.*, 2016; Dissanayake, 2004). Young drivers have been identified poor in safety performance. In addition, compared to adult drivers, young drivers have a little technical ability to control a vehicle in safe. Dissanayake and Lu, (2002) reveal that young drivers are inexperienced, risk-taking behaviour and immaturity and at greater risk of exposure. Ismail *et al.*, (2016) surveyed on 127 students with mean age 22 to identify the relationship between personality traits and aggressiveness among young Malaysian drivers. They found young drivers tend to underestimate their driving ability. They suggest that the implementation of a personality test to reduce the road rage phenomena and aggressive driving should be considered in the future. Huang and Winston (2011) recommended developing driving skills, expertise and competencies including psychomotor, cognitive and perceptual proficiencies to fulfil the gap between safety limit and the freedom of explore in driving among young drivers.

Previous research identified few of factors that influence the injury severity of young drivers. Dissanayake, (2004) developed models to predict crash severity of young and older drivers in single-vehicle crashes. They identified that frontal impact increase severity of older drivers than young drivers. In addition, night time crashes, speeding behaviour, the existence of a grade or curve also influenced of crash severity. Clarke *et al.* (2010) analyse 1184 fatal vehicle occupants in the UK. They found that most of the fatal crashes along bend segments are the majority involved with young drivers. Although there is some research on young drivers, much more remains to be unknown about how crashes among young drivers are different from middle and older aged groups especially in middle-income countries like Malaysia. The objective of this study was to identify the characteristics of crashes, road, driver and vehicle factors associated with injury-producing crashes among young drivers.

Table 1. Crash involvement by driver age group (MIROS, 2012)

Age Group (Years)	Injury severity		Total
	Non-severe	Severe	
<25	45,585	2,495	**48,080**
25-44	9,829	689	**10,518**
45-64	20,415	1,142	**21,557**
>64	1,618	83	**1,701**
Overall			81,856

RESEARCH METHODS

a. Data description

Five years crash data from 2008 to 2012 obtained from Malaysian Institute of Road Safety Research – Road Accident Analysis and Database System (M-ROADS) was used in this study. In Malaysia, all crash victims are needed to make a police report for claim purpose using POL27 form. However, not all the variables are recorded in the system especially for property damage only crashes. The variables that have been used in this study are time of day, day of the week, season of the year, school season, collision type, crash type, road type, road geometry, driver gender, driver errors and vehicle type.

A total of 48,080 crashes involved young drivers occurred within the study period. In Malaysia, driver severity was categorized into four; fatal, serious injury, slight injury and property damage only (PDO). Due to large differences between PDO and other categories of injury, two-level of injury was created in this study; severe (fatal, serious and slight injury) and non-severe (PDO). Afterwards, crash severity is referring to the injury severity of young drivers.

b. Methodology

Two-step modelling technique has been used with a combination of decision tree and logistics regression. This method has been applied in road safety research (e.g. Washington, 2000; Haque *et al.*, 2016; Rusli *et al.*, 2018). A decision tree is a nonparametric method, which identifies the possible relationships among explanatory factors and driver injury severity. The Chi-Squared Automatic Interaction Detection (CHAID) data mining algorithm has been used in this study. However, the non-parametric method is suffering from Type I errors. The identified group from decision tree are converted to indicator variables to develop a statistical interference.

Binary logistics was used to examine the relationship between explanatory variables and response variable. The logit was the natural logarithm of the odds that the response variable Y was severe (Y=1) versus non-severe (Y=0) as shown in Eq. 1:

$$Logit\ (P) = \ln\left(\frac{1-p}{p}\right)\beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_i x_i \qquad (1)$$

where P is the probability of driver injury severity involved in crashes, x_i is the independent variable and β_i is the model coefficient directly determining the odds ratio.

RESULTS AND DISCUSSION

Table 2 presents the summary statistics of the variables included in this study. The severe crashes involved with young drivers was higher for crashes that occur during daytime (61.8%) compared to night time (38.4%). In the week, weekdays represent a higher proportion of severe crashes involved young drivers. The highest proportion of severe crashes also recorded on crashes during dry seasons and school days.

Out of 8 collision types, a head-on collision was a dominant recorded higher proportion of severe crashes (39.3%). It is followed by angle and right-angle side collision (28.8%) and rear-end collision (13.5%). 'Out-of-control', sideswipe, overturn, vehicle-animal/pedestrian and vehicle-object collisions respectively recorded 11.7%, 3.1%, 1.6%, 1.1% and 1.0% severe crashes among young drivers. 62.1% of severe crashes involved with more than one vehicle.

In this crash data set, three road types were included in the analysis; state road, federal road and municipal road. The highest proportion of severe crashes occurred on the federal road (55.9%). State road recorded 997 severe crashes represents 40.0% of total severe crashes during the study period. In term of road geometry, crashes on straight road section recorded a higher proportion of severe (63.8%). Bend section recorded 21.7% of severe crashes and followed by T/Y junction with 11.1%.

Male young drivers represent the highest group involved in severe (85.3%) and non-severe (76.1%) crashes. Six driver behaviours were included in this analysis with one variable represent not-at-fault drivers. Out of these, not-at-fault drivers represent the highest group involved with severe crashes with a proportion of 71.1%. Crashes involved in speeding behaviour has been identified recorded a higher proportion of severe crashes (12.6%) and followed by dangerous turning behaviour (6.4%). Passenger car made up most crash-involved vehicle types (62.3% of non-severe crashes and 48.6% of severe crashes). Four-wheel drive represents the second type of vehicle was recorded higher severe and non-severe crashes, with corresponding proportion respectively 20.6% and 21.2%.

Table 2. Summary statistics of variables

Variables	Non-Severe (%)	Severe (%)	Total (%)
Time of day (TD)			
Day time (DT)	31159(68.4)	1536(61.6)	32695(68.0)
Night time (NT)	14426(31.6)	959(38.4)	15385(32.0)
Day of week (DW)			
Weekdays (WD)	31843(69.9)	1714(68.7)	33557(69.8)
Weekend (WK)	13742(30.1)	781(31.3)	14523(30.2)
Season of year (SY)			
Dry season (DS)	34154(74.9)	1850(74.1)	36004(74.9)
Wet season (WS)	11431(25.1)	645(25.9)	12076(25.1)
School season (SS)			
School days (SD)	36263(79.6)	1970(79.0)	38233(79.5)
School holidays (SH)	9322(20.4)	525(21.0)	9847(20.5)
Collision type (CT)			
Rear-end (RE)	20923(45.9)	336(13.5)	21259(44.2)

Variables	Non-Severe (%)	Severe (%)	Total (%)
Angle and right-angle side (AS)	10503(23.0)	718(28.8)	11221(23.3)
Out-of-control (OC)	8650(19.0)	291(11.7)	8941(18.6)
Head-on (HO)	1615(3.5)	980(39.3)	2595(5.4)
Sideswipe (SW)	1746(3.8)	78(3.1)	1824(3.8)
Overturn (OT)	789(1.7)	40(1.6)	829(1.7)
Vehicle-animal/pedestrian (VAP)	664(1.5)	28(1.1)	692(1.4)
Vehicle-object (VO)	695(1.5)	24(1.0)	719(1.5)
Crash type (CP)			
Multi-vehicle (MV)	32962(72.3)	1549(62.1)	34511(71.8)
Single vehicle (SV)	12623(27.7)	946(37.9)	12569(28.2)
Road type (RT)			
State road (SR)	21202(46.5)	997(40.0)	22199(46.2)
Federal road (FR)	21554(47.3)	1394(55.9)	22948(47.7)
Municipal road (MR)	2829(6.2)	104(4.2)	2933(6.1)
Road geometry (RG)			
Straight (ST)	29542(64.8)	1592(63.8)	31134(64.8)
Bend (BD)	6965(15.3)	542(21.7)	7507(15.6)
Roundabout (RA)	2748(6.0)	41(1.6)	2789(5.8)
Cross junction (CJ)	912(2.0)	44(1.8)	956(2.0)
T/Y junction (TY)	5418(11.9)	276(11.1)	5694(11.8)
Gender (GD)			
Male (ML)	34689(76.1)	2127(85.3)	36816(76.6)
Female (FM)	10896(23.9)	368(14.7)	11264(23.4)
Driver errors (DE)			
Not-at-fault (NF)	22226(48.8)	1773(71.1)	23999(49.9)
Speeding (SP)	9476(20.8)	314(12.6)	9790(20.4)
Driving too close (DC)	8913(19.6)	50(2.0)	8963(18.6)
Dangerous turning (DT)	2640(5.8)	159(6.4)	2799(5.8)
Dangerous overtaking (DO)	1355(3.0)	91(3.6)	1446(3.0)
Careless driving (CD)	822(1.8)	81(3.2)	903(1.9)
Traffic light violation (TV)	153(0.3)	27(1.1)	180(0.4)
Vehicle type (VT)			
Passenger car (PC)	28379(62.3)	1197(48.0)	29576(61.5)
Four-wheel drive (4WD)	9664(21.2)	515(20.6)	10179(21.2)
Van (VN)	2095(4.6)	167(6.7)	2262(4.7)
Bus (BS)	2291(5.0)	248(9.9)	2539(5.3)
Lorry (LR)	3156(6.9)	368(14.7)	3524(7.3)

Figure 1 shows the decision tree for the crash's severity among young drivers in Sabah Malaysia. The decision was validated using 10-fold stratified cross-validation in which on each cycle nine-tenths of the data were used to train the decision tree and the remained one-tenth was used to measure the fitness. The estimated decision tree correctly classified 94% of instances, using 33 leaves for a total tree size of 53 nodes. At the top of the tree, the collision type represents the highest information gain. The estimated decision tree classified the driver severity by segmenting the dataset into 33 smaller and more homogenous groups. The group statistics that indicate the classification rules for crash severity are reported in Figure 1. For example, the statistics of branch 1 suggest that 14.6% of overturn or sideswipe collisions involved lorry or bus in daytime resulted in injury and 85.4% did not result in reported injury. Likewise, branch 33 identifies out of total 48,080 crashes involved young drivers, 0.6% crashes involved collision between passenger car with animal or pedestrian during the daytime. Out of these, 54.2% were crashes that resulted in injury and 45.8% were crashes that did not result in reported injury. Following the branches of the estimated decision tree, 33 higher-order interaction terms were

created and numbered as shown in the brackets at the bottom of each tree branch. These 33 higher-order interaction terms coupled with main effect variables were tested in the binary logistics model.

Table 3 shows the estimation of logistics regression for young driver's injury severity. Out of eleven selected factors, eight of them found statistically significant influencing injury severity of young drivers including time of the day, crash type, collision type, road geometry, gender, driver errors and vehicle type. In addition to these main effect variables, the logistics model also identified 16 higher-order interaction variables affecting the severity of young drivers including interaction variable 4, 6, 7, 12, 15, 16, 17, 19, 20, 22, 26, 28, 29, 30 and 31.

The odds of severe crashes during night time were 1.4 times *(OR1.393, 95% CI: 1.259-1.541)* than daytime crashes. Crashes involved with single vehicle had higher being severe *(OR1.429, 95% CI:1.155-1.769)* compared to multi-vehicle crashes.

In terms of collision type, the odds of angle and right-angle collision *(OR2.474, 95% CI:1.015-6.031)*, head-on *(OR33.656, 95% CI:15.033-75.348)*, side-swipe *(OR3.052, 95% CI:1.477-6.306)*, overturn *(OR2.571, 95% CI:1.178-5.614)* and vehicle-animal/pedestrian *(OR87.370, 95% CI:43.943-173.711)* higher than rear-end collisions. Young drivers involved in crashes along federal roads resulted in higher severe odds than state roads. However, young drivers involved in crashes along municipal roads have lower odds *(OR0.639, 95% CI:0.503-0.812)* compared to state roads.

Crashes on the roundabout and T/Y junction had a lower odds being severe than straight section, with corresponding odds respectively about 60% lower *(OR0.40, 95% CI: 0.256-0.627)* and 21% lower *(OR0.79, 95% CI:0.643-0.986)*. Young female drivers had lower of being involved in severe crashes than male drivers *(OR0.673, 95% CI: 0.592-0.764)*. Compared to not-at-fault drivers, drivers with bad driving behaviours like speeding *(OR0.639, 95% CI:0.514-0.793)*, driving too close *(OR0.466, 95% CI:0.253-0.859)* and dangerous overtaking *(OR0.659, 95% CI:0.511-0.850)* had lower odds being injury. However, crashes from traffic light violation behaviours resulted in higher odds injury than crashes on not-at-fault drivers.

Four-wheel drive, van, bus and lorry are positively associated with injury-producing crashes, with the corresponding the likelihood that the crash leads to reported injury respectively about 1.2 times *(OR1.229, 95% CI:1.088-1.388)*, 1.4 times (*OR1.398, 95% CI:1.126-1.729)*, 1.4 times *(OR1.375, 95% CI:1.141-1.657)* and 1.8 times *(OR1.831, 95% CI:1.553-2.160)* higher than passenger car.

Among 33 higher-order interaction variables found in decision tree analysis, respectively eight of them are found positively and another eight are negatively associated with injury-producing crashes among young drivers. Interaction variable 6 is found positively associated with injury-producing crashes, with corresponding odds about 8.9 times *(OR8.925, 95% CI:1.936-41.154)* higher. The same observations also found for interaction variable 7, interaction variable 12 and interaction variable 15, with corresponding odds respectively 5.0 times *(OR4.996, 95% CI:1.102-22.637)*, 3.8 times *(OR3.848, 95% CI:1.963-7.542)* and 3.2 times *(OR3.232, 95% CI:1.422-7.346)* higher. In addition, interaction variable 17, 19, 20 and 22 also found have a higher probability of severe, with corresponding odds respectively about 3.3 times *(OR3.264, 95% CI:1.506-7.074)*, 6.4 times *(OR6.402, 95% CI:3.135-13.075)*, 2.2 times *(OR2.242, 95% CI:1.065-4.721)* and 2.4 times *(OR2.454, 95% CI:1.240-4.858)* higher. Besides that, there are eight interaction variables found negatively associated with injury-severity producing crashes including interaction variable 4 *(OR0.266, 95% CI:0.121-0.583)*, interaction variable 16 *(OR0.327, 95% CI:0.136-0.788)*, interaction variable 26 *(OR0.484, 95% CI:0.289-0.811)*, interaction variable 28 *(OR0.498, 95% CI:0.292-0.811)*, interaction variable 29 *(OR0.265, 95% CI:0.106-0.664)*, interaction variable 30 *(OR0.138, 95% CI:0.101-0.190)*, interaction variable 31 *(OR0.031, 95% CI:0.016-0.060)* and interaction variable 33 *(OR0.387, 95% CI:201-0.745)*.
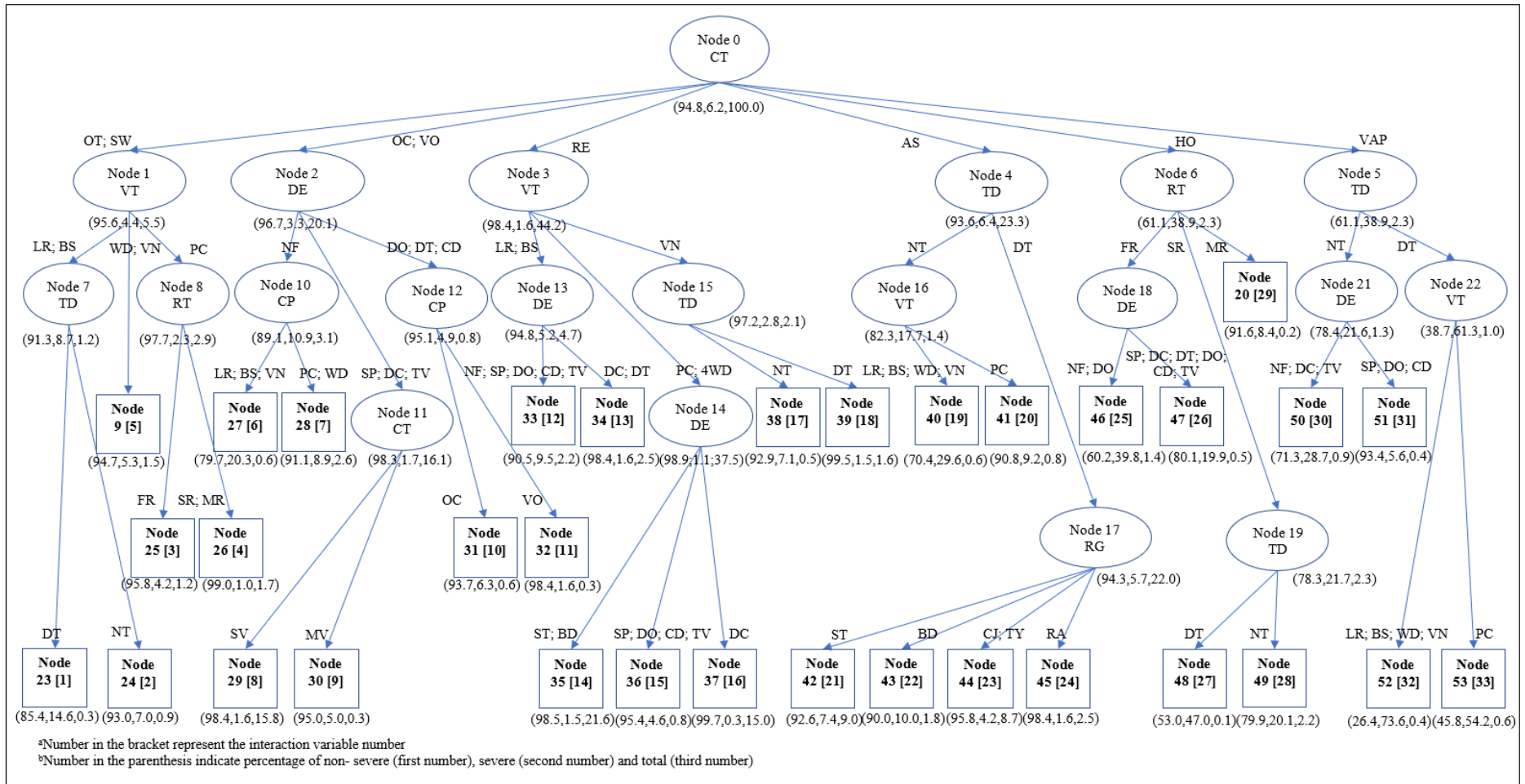
Figure 1. Injury severity for young driver's injury severity model

Table 3. Model estimation for injury severity of young drivers

Variable	Reference category	Odds Ratio (OR)	Confidence Interval (95%)	*p*-value
Time of the day				
Night-time	Daytime	1.393	1.259-1.541	0.000***
Crash type				
Single-vehicle	Multi-vehicle	1.429	1.155-1.769	0.001**
Collision type				
Angle and right-angle side		2.474	1.015-6.031	0.046*
Out-of-control		0.801	0.154-4.157	0.792
Head-on		33.656	15.033-75.348	0.000***
Sideswipe	Rear-end	3.052	1.477-6.306	0.003**
Overturn		2.571	1.178-5.614	0.018*
Vehicle-animal/pedestrian		0.370	0.343-3.711	0.780
Vehicle-object		0.702	0.145-3.401	0.660
Road type				
Federal road	State road	1.158	1.040-1.290	0.007**
Municipal road		0.639	0.503-0.812	0.000***
Road geometry				
Bend		0.962	0.843-1.097	0.561
Roundabout	Straight	0.401	0.256-0.627	0.000***
Cross junction		0.718	0.504-1.025	0.068
T/Y junction		0.796	0.643-0.986	0.037*
Gender				
Female	Male	0.673	0.592-0.764	0.000***
Driver errors				
Speeding		0.639	0.514-0.793	0.000***
Driving too close		0.466	0.253-0.859	0.014*
Dangerous turning		0.932	0.767-1.132	0.479
Dangerous overtaking	Not-at-fault	0.659	0.511-0.850	0.001**
Careless driving		0.978	0.694-1.380	0.901
Traffic light violation		1.900	1.230-2.933	0.004**
Vehicle type				
Four-wheel drive		1.229	1.088-1.388	0.001**
Van	Passenger car	1.395	1.126-1.729	0.002**
Bus		1.375	1.141-1.657	0.001**
Lorry		1.831	1.553-2.160	0.000***
Interaction variables				
Interaction variable 4 *(1 if interaction 4 was involved, 0 otherwise)*		0.266	0.121-0.583	0.001**
Interaction variable 6 *(1 if interaction 6 was involved, 0 otherwise)*		8.925	1.936-41.154	0.005**
Interaction variable 7 *(1 if interaction 7 was involved, 0 otherwise)*		4.995	1.102-22.637	0.037*
Interaction variable 12 *(1 if interaction 12 was involved, 0 otherwise)*		3.848	1.963-7.542	0.000***
Interaction variable 15 *(1 if interaction 15 was involved, 0 otherwise)*		3.232	1.422-7.346	0.005**

Variable	Reference category	Odds Ratio (OR)	Confidence Interval (95%)	*p*-value
Interaction variable 16 *(1 if interaction 16 was involved, 0 otherwise)*		0.327	0.136-0.788	0.013*
Interaction variable 17 *(1 if interaction 17 was involved, 0 otherwise)*		3.264	1.506-7.074	0.003**
Interaction variable 19 *(1 if interaction 19 was involved, 0 otherwise)*		6.402	3.135-13.075	0.000***
Interaction variable 20 *(1 if interaction 20 was involved, 0 otherwise)*		2.242	1.065-4.721	0.034*
Interaction variable 22 *(1 if interaction 22 was involved, 0 otherwise)*		2.454	1.240-4.858	0.010*
Interaction variable 26 *(1 if interaction 26 was involved, 0 otherwise)*		0.484	0.289-0.811	0.006**
Interaction variable 28 *(1 if interaction 28 was involved, 0 otherwise)*		0.498	0.292-0.848	0.010*
Interaction variable 29 *(1 if interaction 29 was involved, 0 otherwise)*		0.265	0.106-0.664	0.005**
Interaction variable 30 *(1 if interaction 30 was involved, 0 otherwise)*		0.138	0.101-0.190	0.000*
Interaction variable 31 *(1 if interaction 31 was involved, 0 otherwise)*		0.031	0.016-0.060	0000***
Interaction variable 33 *(1 if interaction 33 was involved, 0 otherwise)*		0.387	0.201-0.745	0.005**

*significant at 0.05 level
**significant at 0.01 level
***significant at 0.001 level

Discussion

a. Crash characteristics

The results show that crashes during night time are found to be more severe. This finding is in line with previous studies (e.g. Dissanayake & Lu, 2002; Clarke et al., 2010). Single-vehicle crashes found to be more to produce injury when involved with young drivers. Kockelman & Kweon, (2002) suggest that reckless driving behaviour increase the probability of severe single-vehicle crashes involvement among young drivers compared to middle-age drivers. Dissanayake & Lu, (2002) explain that single-vehicle crashes are more severe because most of the case, the vehicle leaves the road and overturns of hits a roadside object such as a tree or a pole. Among seven collision types included in this study, angle and right-angle side, head-on, sideswipe and overturn are found statistically significant and positively associated with injury severity crashes. Previous studies also identified head-on collision increase the probability of severe because of this collision involved with the vehicle from the opposite direction (Hosseinpour, Shukri Yahaya, Farhan Sadullah, Ismail, & Reza Ghadiri, 2016). The impact of this collision is higher because of

the different speed between two opposite vehicles. Time for a reaction is also very limited in this collision. However, Dissanayake, (2004) identified that frontal impact only significantly influencing the severity of older drivers and not to the young drivers.

b. Road factors

Crashes that occurred on federal roads found a higher probability of severe than state roads. On the other hand, crashes on municipal roads found not reported to injury. The explanation behind this is a different standard on these three types of roads. In addition, different agencies are involved in the management and maintenance of these roads under different sources of budget. Roundabout and T/Y junction were found significant and negatively associated with crashes producing injury. These two types of intersection have less severe conflict point compared to other types of intersections.

c. Driver factors

The results show that the odds of female drivers involved with injury-producing crashes were lower than male drivers. This finding is in line with previous research on young drivers in Finland (Laapotti & Keskinen, 1998). They suggested that male is more apt to engage in risky driving behaviour. On the other hand, female drivers more cautious in driving and follow the traffic rules. Driver 'not-at-fault' is the highest proportion group involved with crashes producing injury. Traffic light violation among the driver errors that increase the probability to involve in severe crashes. This problem is very serious in Malaysia. Government has been taken many actions to account this behaviour including installation of a red-light camera. The effectiveness of the red-light camera in Malaysia has been identified by previous research (e.g. Jamil, Shabadin, & Rahim, 2014; Kabit, Sabihin, & WH, 2016).

d. Vehicle factors

The involvement of heavy vehicle in a crash involved young drivers is positively associated with crash severity. The main reason for this is the size and weight of heavy vehicle increase the probability of severe crashes. In addition to heavy vehicle, crashes involved van and four-wheel drive also found increasing the probability of severe among young drivers. The design of van is one of the reasons behind this situation. Van was designed with less protection from frontal impact. Four-wheel drive has a higher gravity center and less stable than passenger car (Keall, Newstead, & Watson, 2006).

e. Higher-order interaction

Interaction variable 4 includes higher-order interaction between passenger car, overturn and sideswipe and road type. Results suggest that the crash severity is lower if passenger car involved in overturn or sideswipe collisions on state and municipal roads. On the other hand, not-at-fault drivers involved with out-of-control or hit an object found positively associated with injury-producing crashes and represents by interaction variable 6 and 7. Interaction variable 12 represents young drivers of a heavy vehicle with behaviours of speeding, dangerous overtaking, careless driving, traffic light violations or not-at-fault involved in rear-end collision resulted in more severe in crashes. The same observation also found when drivers of a passenger car and four-wheel drive with speeding, dangerous overtaking, careless driving and traffic light violation behaviours involved with rear-end collision (interaction variable 15). However, if a careless driver

on passenger car or four-wheel drive involved in rear-end collision, the crashes do not result in reported injury. Van drivers only positively associated with injury-producing crashes in rear-end collision during night time (interaction variable 17). For angle and right-angle side collision, night time crashes involved passenger car found positively associated with injury-producing crashes (interaction variable 20). Interaction variable 22 includes higher-order interaction between collision type, time of the day and road geometry. Results suggest that the crash severity is higher if the angle and right-angle side collision happen along the bend section during the daytime.

Although head-on collision found positively associated with injury producing crashes, head-on collision on federal roads involved bad driver behaviours (e.g. speeding, driving too close, dangerous turning, dangerous overtaking, careless driving and traffic light violation) found negatively associated with injury-producing crashes (interaction variable 26). Head-on collision on state roads during night time also found negatively associated with injury-producing crashes (interaction variable 28). Head-on collisions on municipal roads are associated with crashes that do not result in reported injury.

Vehicle-animal/pedestrian collision on night time are found negatively associated with injury-producing crashes. In addition, during the daytime, vehicle-animal collision also found negatively associated with injury-producing crashes for a passenger car.

CONCLUSION

Five years of crash data in Sabah, Malaysia was used to identify factors influencing injury severity for crashes involved young drivers. A combination of parametric analysis (decision tree) and non-parametric analysis (binary logistics) was applied in this study offers more insights into crash severity through higher-order interaction variables. Results from the decision tree identified 33 high-order interactions between main effect variables. The binary logistics was performed to examine these interaction variables together with main effect variables in influencing injury severity of young drivers.

Out of eleven main effect variables, eight of them found statistically significant to explain injury severity of young drivers including time of the week, crash type, collision type, road type, road geometry, gender, driver errors and vehicle type. In addition to the main effect variables, 16 interaction variables found statistically significant influencing injury severity of young drivers.

Crash during night time, single-vehicle, federal roads and traffic light violation found were found more severe when involved young drivers. In terms of collision type, angle and right-angle side, head-on, sideswipe and overturn found increase the probability of injuries. The crash involved heavy vehicles (lorry and bus) was more severe than a passenger car. The same observation was also found on crashes involved with four-wheel drive and van.

On the other side, few of the variables found contributed to the less severe crash including municipal road, roundabout and T/Y junction and female drivers. Out of 6 driver errors contributed to the crashes, three of them found reported less severe crashes including speeding, driving too close and dangerous overtaking.

In terms of high-order interaction variables, eight of them found significantly associated with crashes producing injury and another eight were found reported contradictory. Findings of this study will help related authorities to design well-targeted restrictive measures in reducing the severity level of young drivers in traffic crashes.

REFERENCES

Clarke, D. D., Ward, P., Bartle, C., & Truman, W. Killer crashes: fatal road traffic accidents in the UK. *Accident Analysis & Prevention*, vol. 42, no. 2, pp. 764–770, 2010.

Dissanayake, S. Comparison of severity affecting factors between young and older drivers involved in single vehicle crashes. *IATSS Research*, vol. 28, no. 2, pp. 48–54, 2004.

Dissanayake, S., & Lu, J. Analysis of severity of young driver crashes: Sequential binary logistic regression modeling. *Transportation Research Record*, vol. 1784, no.1, pp. 108–114, 2002.

Haque, M. M., Ohlhauser, A. D., Washington, S., & Boyle, L. N. Decisions and actions of distracted drivers at the onset of yellow lights. *Accident Analysis & Prevention*, vol. 96, pp. 290–299, 2016.

Hosseinpour, M., Shukri Yahaya, A., Farhan Sadullah, A., Ismail, N., & Reza Ghadiri, S. M. Evaluating the effects of road geometry, environment, and traffic volume on rollover crashes. *Transport*, vol. 31, no. 2, pp.221–232, 2016.

Huang, P., & Winston, F. K. Chapter 23 - Young Drivers. In B. E. B. T.-H. of T. P. Porter (Ed.) (pp. 315–338). San Diego: Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-381984-0.10023-2, 2011.

Ismail, R., Voon, N. L., Saad, M. H. M., Saleem, M., & Ibrahim, N. Aberrant driving among young malaysian drivers. *Jurnal Teknologi*, vol. 78, pp.6–10, 2016.

Jamil, H. M., Shabadin, A., & Rahim, S. A. The effectiveness of automated enforcement system in reducing red light running violations in Malaysia: pilot locations, 2014.

JKJR. *Pelan Keselamatan Jalan Raya Malaysia 2014 - 2020*. Kuala Lumpur. Retrieved from http://www.mot.gov.my/SiteCollectionDocuments/Darat/Pelan+KJRM+2014-2020.pdf, 2014.

Kabit, M. R., Sabihin, N. A., & WH, W. I. Effectiveness of Automated Enforcement System (AES) in Reducing Red Light Violation (RLV) Behaviours: A Case Study In Kuala Lumpur. *Journal of Civil Engineering, Science and Technology*, vol. 7, no. 1, pp. 39–44, 2016.

Keall, M. D., Newstead, S. V, & Watson, L. *Four-wheel drive vehicle crash involvement patterns*, 2006.

Kockelman, K. M., & Kweon, Y.-J. Driver injury severity: an application of ordered probit models. *Accident Analysis & Prevention*, vol. 34, no. 3, pp. 313–321, 2002.

Laapotti, S., & Keskinen, E. Differences in fatal loss-of-control accidents between young male and female drivers. *Accident Analysis & Prevention*, vol. 30, no. 4, pp.435–442, 1998.

MIROS. *Crash Data (2008 - 2012) Sabah, Malaysia*. Kajang, 2012.

Rusli, R., Haque, M. M., Saifuzzaman, M., & King, M. Crash severity along rural mountainous highways in Malaysia: An application of a combined decision tree and logistic regression model. *Traffic Injury Prevention*, pp. 1–8. https://doi.org/10.1080/15389588.2018.1482537, 2018.

WHO. *Global status report on road safety 2018*. Geneva: World Health Organization, 2018.