Kufa Journal of Engineering
Vol. 15, No. 2, April 2024, P.P. 74-92
Article history: Received 19 November 2023, last revised 9 December 2023, accepted 9 December 2023

MODELLING THE FACTORS AFFECTING CRASH OCCURRENCE AND FREQUENCY RESULTED FROM MOBILE PHONE USE WHILE DRIVING: EVIDENCE FROM AL-NAJAF, IRAQ

Firas H. Asad[1] and Shaimaa N. Hadi[2]

[1] Department of Civil Engineering, Faculty of Engineering, University of Kufa, Najaf, Iraq. E-mail: firas.alwan@uokufa.edu.iq

[2] Department of Civil Engineering, Faculty of Engineering, University of Kufa, Najaf, Iraq. E-mail: Shiamaalhajjar@gmail.com

https://doi.org/10.30572/2018/kje/150206

ABSTRACT

The recent reports of road traffic accident statistics in Iraq have disclosed a rise in the number of crash injuries resulted from the use of mobile phone while driving. This paper aims to explore the factors contributing to the occurrence and prevalence of such crashes and near crashes in Al-Najaf governorate, Iraq. A representative sample of 417 drivers were interviewed as part of a questionnaire driving survey. Several frequency and modelling analyses were conducted using the IBM SPSS software. The frequency analysis revealed a high use of mobile phones for calling and texting activities while driving. Almost 20% and 55% of the interviewed drivers reported their involvement in a crash or in a near crash because of such use, respectively. Regarding the developed logistic models, the crash involvement sequential regression analysis revealed that factors such gender, education, handheld phoning, calling-answering frequency, and inadequate driving can affect the likelihood of crash occurrence. In contrast, the ordinal logistic near miss models revealed that age, gender, high phone use rate, and improper driving due to such use are influential factors in rising the likelihood of being in multiple near crashes. The analysis results confirm the influence of using phones in distracting the attention of drivers and hence threating their lives; as a result, these findings would be enlightening for agencies and policy makers interested in highway safety.

KEYWORDS: Distracted driving; questionnaire survey; Texting while driving; Phoning while driving; logistic regression; Crash analysis.

1. INTRODUCTION

Given the increasing integration of digital communications and modern technology devices into daily life, more research is necessary to fully understand the dynamics of driver distraction and how it relates to traffic safety outcomes. In 2021, about 8% of fatal crashes and 14% of injury crashes were mainly because of distracted driving (NCSA, 2023). Typical examples that have been shown to distract drivers comprise internal vehicle distractions such as conversing with passengers, smoking, and utilizing mobile phones, as well as external diversions like observing pedestrians or roadside signage (WHO, 2011). Most important is the growing prevalence of mobile phone use (MPU) while driving, which has been implicated as a notable factor contributing to the rising number of traffic accidents globally. Previous research has demonstrated how distracted driving resulting from increased mobile phone utilization during transport can undermine drivers' focus on the road and increase hazards (Cordellieri et al., 2022). Despite the government's interest worldwide in the issue of mobile-based distracted driving, legislative and law enforcement efforts to limit the use of cell phones by drivers have not been entirely successful (Rudisill et al., 2019; Chen et al., 2020; Truelove et al., 2021). The latest report of World Health Organization stated that all countries should have set national laws to restrict MPU during driving by 2030 (WHO, 2018). According to the European road safety observatory, traffic surveys conducted in United States and Europe specified that between 1% and 11% of vehicle drivers use their phones during driving. In addition, MPU while driving can escalate the risk of being involved in a road crash that led to serious injuries or property damage by 75%. The occurrence likelihood of a crash that involves a driver using a phone while driving is doubled for drivers who use them frequently (ERSO, 2018).

Recent studies have attempted to determine and measure the variables that may raise the chance of using a cell phone while operating a vehicle, as well as the ways in which this behavior may worsen road safety by raising the risk of collisions or impairing driving ability. The majority of these research' conceptual frameworks are generally shown in Fig. 1. Driving-related attributes and driver’s personal characteristics are among the risk factors that have been researched. While conventional examples of driving attributes include a driver's license, driving experience, type of vehicle, and quantity of daily travel, examples of driver’s traits include age, sex, academic qualification, and risk perception (Claveria et al., 2019; Fraschetti et al., 2021; Cordellieri et al., 2022). According to the research methods that have been commonly used by researchers, the studies investigating MPU while driving can be grouped into naturalistic (Bastos et al., 2020); driving surveys (Fraschetti et al., 2021; Claveria et al., 2019), and driving simulation

(Spyropoulou and Linardou, 2019; Chen, et al., 2020). Furthermore, there are primarily two categories of detrimental effects of mobile-based distracted driving that have been studied. The first is the unsafe driving behavior, which includes braking, steering, changing lanes, and speeding (Spyropoulou and Linardou, 2019; Wang et al., 2020). Getting into a car crash is the second (Bener et al., 2010; Asbridge et al., 2013). The majority of these studies conclude with the recommendation of many intervening safety measures that might help decision-makers create short and long-term strategies for transportation safety and accident-prevention plans.

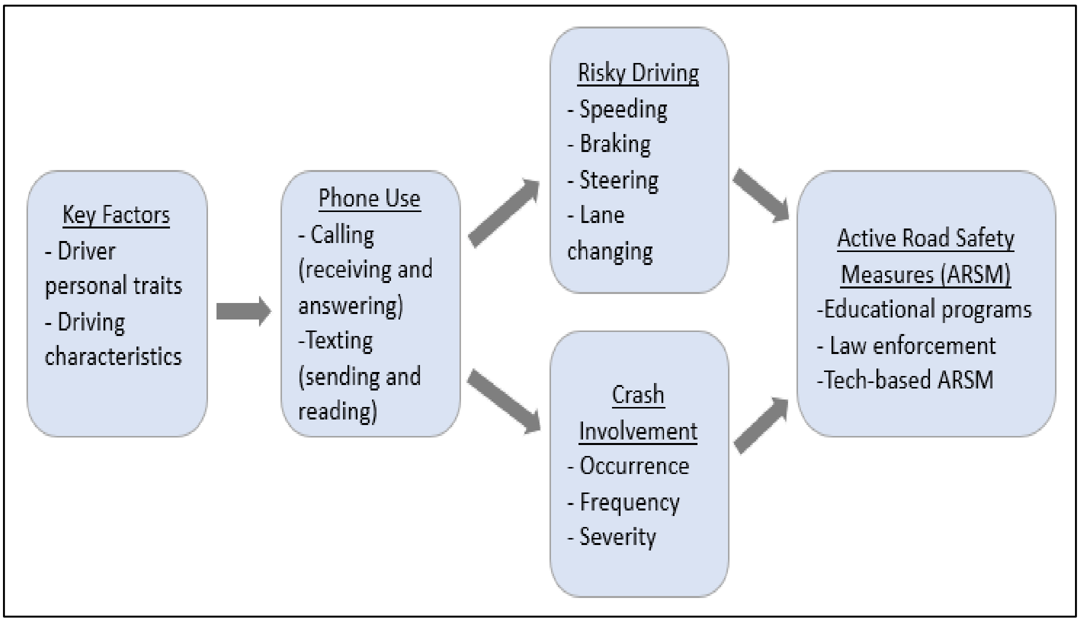

Fig. 1. Typical conceptual framework for most studies addressing the consequences of mobile phone use while driving (by the researchers).

The Iraqi central statistical organization (CSO) website reports that the ratio of mobile phones per 100 people climbed from 90.6 in 2015 to 98.8 in 2021 (CSO, 2023a). According to the national crash statistics, the number of injuries from crashes in Iraq caused by careless driving rose from slightly over 500 in 2018 to over 1660 in 2022 (CSO, 2019; 2023b). With respect to Al-Najaf governorate, the number of such injuries rose from 68 case in 2019 to 280 case in 2022 (CSO, 2020a; 2023b). Additionally, the total number of crash injuries in all Iraqi governorates because of MPU while driving rose from about 120 in 2020 to 275 in 2022 (CSO, 2021a; 2023b).

Lastly, aside from the statistics mentioned above, it is acknowledged that data about the MPU form of distracted driving are important to obtain because they are not regularly collected in many countries (WHO, 2011). This is also the case in Iraq, where based on the WHO safety report, there is no pertinent data regarding the use of mobile phones while driving (WHO, 2018). In light of this, the current study attempts to add to the body of knowledge about drivers who use their phones while driving by examining their distinct characteristics and exploring the effect of these characteristics in rising the odds of being in a crash.

2. DATA AND METHODS

2.1. Study area

The study area is Al-Najaf governorate which is with spatial coordinates of 32° 01' 33.38" N (latitude) and 44° 20' 46.50" E (longitude). Al-Najaf is located in the Iraq's central regions about 160 km to the south-west of Baghdad, the capital. Based on the annual statistical group report, the governorate has a total area of 28,824 km2 (CSO, 2021b), which represent about 6.6% of the total area of Iraq; and it is with 2021-based estimated population of about 1,500,000 (CSO, 2020b).

2.2. Survey Design

The survey's design and preparation are in accordance with the research's predetermined objectives. The current research involves carrying out a driving survey to gather self-reporting empirical data about the characteristics and safety consequences of mobile phone use (texting and calling) from a decent sample of 417 drivers. The interview process extended from December 2022 until May 2023.

2.2.1. Survey instrument and validity

A paper-based questionnaire was chosen as the survey instrument. The questionnaire was made up of three primary sections. In the first section, drivers' age, sex, educational background, and occupation are among the sociodemographic details that were questioned. Driving factors including vehicle type, driving experience, total daily driving distance, and total number of driving trips were included in the second section. The third section has three subsections because its goal is to gather enough information about texting and calling while operating a motor vehicle. Seven questions make up the first subsection. The first six are designed to collect information about how frequently drivers use their phones to make and receive calls, send and receive messages, and view other notifications. The purpose of the seventh question was to identify the mobile use mechanism (hands-free, hand-held, headset, and Bluetooth). Data about crashes resulted from MPU while driving were gathered in the second subsection. This subsection also looks at the unsafe driving practices connected to phone use-related driver attention. The third subsection focused on drivers' preferences and attitudes on topics including the impact of mobile phone use and the importance of banning phone use while driving. For questionnaire validity, valid survey produces accurate data. It is critical to ensure content validity, in that the questionnaire items (questions) can accurately represent the characteristics they are intended to measure. According to (Fink, 2017), conducting a pilot survey and using questions that have been theoretically accepted and common in literature are the two common mechanisms to ensure questionnaire's internal validity. In this work, 25 drivers were interviewed as part of the pilot survey and their responses, viewpoints, and comments were considered to amend some questions to make them clearer. In addition, several questions have been already used by previous researchers.

2.2.2. Participants recruiting, sampling technique, and sample size

Fink, (2017) stated that questionnaire participatory surveys can be classified into four groups based on the method of administration: these are, mail-out surveys, web-based surveys, phone-based surveys, and in-person interviews. Taking into account the typical advantages and disadvantages of these four approaches, in this research, the most effective method for conducting the survey was to perform in-person interviews. The researchers trained three university students to help in conducting the interviews and filling out the questionnaires. Face-to face interviews along with interviewer-administered questionnaires can greatly contribute to the accuracy of the collected responses and minimizing item nonresponse rate. Regarding sampling technique, the survey was designed to meet, as much as possible, the requirement of probability sampling method to ensure the representativeness of the sample the legitimacy of generalizing research findings from the sample to the general population. In so doing, procedures including volunteering and convenience sampling were excluded (Saunders et al., 2012; Rea and Parker, 2014). In contrast, a simple random sample was targeted by conducting a roadside interview survey. The survey locations were chosen to be safe, secure, and not to yield significant traffic disruption during the interview. A randomly-chosen drivers, that pass through randomly chosen survey locations along preselected segments of several urban streets within the study area, were recruited. However, as there was no traffic police officer to order oncoming vehicles to stop, only drivers that had the willing of participating were interviewed. For the sample size, minimum adequate sample size was determined to be 385 drivers based on recommendations mentioned in (Rea and Parker ,2014) for 5% margin of error and 95% level of confidence. A total sample of 417 drivers were interviewed which is generally appropriate and statistically sufficient for performing the subsequent statistical analysis and generalizing their results.

2.3. Data Analysis

2.3.1. Data processing

Before conducting the required statistical analysis, the questionnaires' responses were extracted, digitized, and then properly coded in order to create a computerized dataset. The initial data processing stage involve removing errors from the raw data that can negatively affect the accuracy of the output. This stage included data screening, cleaning, and imputing (Tabachnic et al., 2019). The IBM SPSS software (v. 26) was utilized to conduct the initial data analysis.

2.3.2. Data analysis and modelling

In contrast, the main quantitative analysis comprised two steps; the first includes a set of descriptive frequency analysis for the variables that will be utilized as explanatory or predicted variables in the crash modelling stage. The second step includes forming and appraising the developed crash models. These models are multivariate regression models with discrete (categorical) dependent variables (DVs); such models belong to a class of regression models that are commonly referred to as generalized linear models (GZLM). These models extend and generalize the traditional ordinary least square (OLS) regression into two aspects; first, the dependent variable is not necessary to be continuous; second, they do not require the errors to be normally distributed, independent, and with constant variance (Orme and Combs-Orme, 2009). In order to sustain the linearity between the DV (outcome) and the set of independent variables (IVs) (predictors), a specific link function should be used. For the case of binomial, multinomial, and ordinal regressions, this function is called "logit". Equation 1 shows the estimated logistic regression model based on a sample of population data (Orme and Combs-Orme, 2009):

$$\text{Logit}\ (Y) = Bo + B1X1 + B2X2 + \ldots + BkXk \qquad (1)$$

Where Y is the estimated mean value of the DV; Bo is the estimated intercept; X1 through Xk represents the included predictors (IVs); and B1 through Bk are the regression slopes.

The term logit (Y) is mathematically described as the natural logarithm of the odds of occurring Y. The odds of Y reflect the likelihood of happening of Y, p(Y) to the likelihood of not happening, 1 - p(Y). Equation 2 represent the mathematical expression of logit (Y) (Orme and Combs-Orme, 2009).

$$\text{Logit}\ (Y) = \ln\ (\text{Odds}\ (Y)) = \ln\ [p\ (Y) / 1 - P(Y)] \qquad (2)$$

Additionally, the odds ratio (OR) is defined as the ratio of the odds of the event of a certain IV value divided by the odds for a different IV value, typically a value one unit lower. The OR specifies the amount of change in the odds and the direction of the relationship between an IV and the DV. OR greater than 1 indicates that the odds of the event increase as values of the IV increase (a positive relationship); and vice versa. Three multivariate regression models were built in this research; the first is a sequential binary logistic regression for modelling crash occurrence, whereas the other two models are ordinal logistic regressions for modelling the frequency of near misses. The IBM SPSS software was used to run the analyses (IBM Corp., 2021).

2.3.3. Logistic regression model assumptions

There are some main assumptions that should be firstly examined before conducting the logistic regression. First, the sample size, it is advised that at least 10 events for each IV should be available, or a minimum sample size of 100 respondents (Orme and Combs-Orme, 2009). Both of these recommendations are valid in the analysis of the current paper. Second, linearity, logistic regression assumes a linear relationship between continuous predictors and the logit transform of the DV (see Eq. 1). The Box-Tidwell approach was used to prove the existence of such linearity, further details are listed in Tabachnick and Fidell (2019). Third, the test of parallel lines assumption, this assumption is necessary to be examined before running the ordinal regression; when it is not met, a multinomial regression should be alternatively used (Tabachnic and Fidell, 2019). In this paper, this assumption was examined and found satisfied.

2.3.4. Logistic regression model evaluation and goodness of fit

The Hosmer–Lemeshow test was adopted to explore whether the observed proportions of events are not significantly different from the predicted probabilities of occurrence in subgroups of the model population. The Pearson Chi-square was used to compare the predicted (E) to the observed (O) frequencies, see Eq. (3) (Park, 2013).

$$H = \sum \frac{(O-E)^2}{E} \tag{3}$$

In contrast, the Pseudo R-square statistic proposed by Nagelkerke (R_N), which is an adjusted version to the Cox and Snell R-square (R_{CS}), was adopted to measure how much the fit goodness of intercept model (base model) can be improved due to the inclusion of predictor variables (full model), see Eqs. 4 and 5. (Tabachnick and Fidell, 2019).

$$R_{CS}^2 = 1 - \exp[\,(-2LL(full) - (-2LL(base)))/n] \tag{4}$$

$$R_N^2 = R_{CS}^2/[1 - \exp[-(-2LL(base)\,/n)]] \tag{5}$$

where LL(base) and LL(full) signify the log-likelihood of the intercept and full models, respectively, and n is the sample size.

3. RESULTS AND DISCUSSION

3.1. Descriptive frequency analysis

A tabulated summary for the personal, driving, and phone use characteristics for the 417 interviewed drivers that have been used in building the subsequent logistic crash and near crash models are listed in Table 1. For drivers' socioeconomic characteristics, the majority of the interviewed drivers are male, employed, and with at least high school education. Regarding driving characteristics and patterns, 75% of vehicles were private cars, the remaining 25% comprised taxis, busses and trucks. For driving experience, nearly 30% of drivers are with less than 5 years of driving. The high standard deviation for the driving distance (61.13) is attributed to the high percentage of taxis, buses and trucks (25%) which usually significantly drive more kilometers than private cars. In specific, some commercial large trucks that transport goods between governorates were found to do only two daily trips but with long travelled distances that reach up to 600 km.

Regarding mobile use characteristics and patterns, the frequency analysis showed that the daily frequencies of MPU while driving for calling/answering activity and reading/texting activity are 7.85 and 3.83, respectively. The relatively high sd for reading-texting activity (4.444) is because whereas 48% of drivers are conservative in using their phones for this activity (two times or less daily), there are as low as 3% of the 417 drivers who reported the use of this activity more than 15 times a day. Further investigation showed that most of those drivers are traditional taxi drivers or work for ride-hailing companies such as Uber.

Other key point is that about 40% of drivers used handheld phones which is a serious challenge for safe driving. Nearly 39% of drivers reported that they occasionally or many times have conducted improper actions due to MPU while driving, such actions involved driving too fast or too slow, inadequate lane changing, and loss in attention. About 89% of drivers stated they believe that phone use can distract drivers at least sometimes. About 71% of drivers reported that they reduce their travel speed while using phones as a safety precaution. Finally, regarding crashes, almost 20% and 55% of drivers reported that they involved in a crash or in a near crash because the MPU while driving, respectively. These statistics should arise a concern and motivate a strict mobile phone use policy. It is worth mentioning that the number of drivers (n) varies from analysis to another because of the missing data of some questions.

Table 1. Descriptive statistics for the personal characteristics of truck drivers.

Characteristics	Statistics	Categories/statistics	Frequency	Percentage %
Socioeconomic Characteristics				
Age		n = 417 ; x = 40.12 ; sd = 10.513		
Gender	n = 417	Male	340	81.5
		Female	77	18.5
Employment status	n = 414	Employed	219	52.9
		Self-employed	146	35.3
		Students	34	8.2
		Others (including retirees)	15	3.6
Education level	n = 415	Higher studies	70	16.9
		University graduate	231	55.6
		High school grad.	62	14.9
		Intermediate school	26	6.3
		Primary school or less	26	6.3
Driving Characteristics and patterns				
Vehicle type	n = 417	Private car	313	75.0
		Taxi	39	9.4
		Others (buses and trucks)	65	15.6
Driving experience (years)	n = 417	0-5	129	30.9
		5-10	86	20.6
		10-15	76	18.2
		15-20	63	15.1
		20-25	32	7.7
		> 25	31	7.4
Driving distance (km)		n = 417 ; x = 49.31 ; sd = 61.13		
Daily trips		n = 415 ; x = 6.09 ; sd = 3.631		
Mobile use characteristics and pattern				
Calling-answering frequency		n = 417 ; x = 7.85 ; sd = 5.723		
Reading-texting frequency		n = 417 ; x = 3.83 ; sd = 4.444		
Mobile use method	n = 416	Handheld	167	40.1
		On speaker	153	36.8
		Bluetooth earpiece/Headset	96	23.1
Frequency of improper actions due to MPU	n = 417	Never	120	28.8
		Few times	135	32.4
		Sometimes	106	25.4
		Many times	56	13.4
MPU distraction (Mobile use distracts driver's attention?)	n = 417	Never or only Few times	46	11.0
		Sometimes	114	27.3
		Often	120	28.8
		Always	137	32.9
Safety measure during MPU	n = 416	Yes - Stopping the car	103	24.8
		Yes - Reducing the speed	295	70.9
		No	18	4.3

Crash frequency				
Crash occurrence	n = 417	Yes	82	19.7
		No	335	80.3
Near miss frequency	n = 415	Never	189	45.6
		Few times	108	26.0
		Several times	118	28.4

n = number of valid cases; x = arithmetic mean; sd = standard deviation.

3.2. Logistic Models

This section aims to conduct inferential statistical analysis, in specific, logistic and ordinal regression analyses, to identify the predictors that are influential on the probability of a driver to be in a crash or in a near miss category due to mobile phone usage. Three multivariate regression models were constructed – one for crash occurrence and two for near misses.

3.2.1. Crash models

The predicted categorical variable is the probability of a driver to be involved in a road accident due to mobile use. Different socioeconomic, driving, and attitudinal driver characteristics were used as predictors (explanatory variables). Due to the binary nature of the outcome variable (accident category: no / yes), the binary logistic regression technique has been performed. The sequential technique was used as a variable selection procedure to build the hierarchical regression with the aim of investigating to what extent the inclusion of driving and attitude driver attributes can enhance the prediction capacity (R-square) of the crash occurrence model. In so doing, the first block was assigned for the socioeconomic variable whereas the second one was assigned for the driving and attitudinal variables. Tables 2 and 3 depict the key details of the first and second variables blocks for the model.

Based on Table 2, two variables were found significant; gender at 5% level of significance (LOS) and education status at 10% LOS. Male drivers are more probable to be in a crash than female drivers; in numbers, they are with 3.149 times higher odds (exp(B)) of being in a crash than women drivers. Regarding education, drivers with university level as their highest education status are more likely to be involved in an accident than others. Table 2 also depicted that as the driver age increases, the possibility to be in a road crash due to phone use decreases. In contrast, drivers who are in employment or student and those who use mobile by handheld or on speaker are more likely to get into a crash than others who are retired or who use Bluetooth or headset devices, respectively. However, the three previous variables were not statistically significant.

Table 2. Sequential binary logistic regression crash model (Block 1)

Variables	B	S.E.	Sig.	Exp(B)
Age	-.014	.017	.420	.986
Gender (1=Male)	1.147	.564	**.042**	3.149
Gender (Ref. Cat.= Female)	-	-	-	-
Employment (1=Employed)	.486	1.090	.656	1.626
Employment (2=Self-employed)	.737	1.106	.505	2.090
Employment (3=Students)	.214	1.263	.865	1.239
Employment (Ref. Cat.= others)	-	-	-	-
Education (1=Higher studies)	1.688	1.162	.146	5.409
Education (2=University)	1.913	1.078	**.076**	6.772
Education (3=High school)	1.669	1.085	.124	5.307
Education (4=Intermediate school)	.741	1.261	.557	2.098
Education (Ref. Cat. = primary school or less)	-	-	-	-
Mobile use (1=Hand-held)	.039	.409	.924	1.040
Mobile use (2=On speaker)	-.362	.403	.370	.697
Mobile use (Ref. Cat. = Car Bluetooth/headset)	-	-	-	-
Constant	-4.459	1.719	.009	.012

B: regression coefficient (ln(OR)); S.E: standard error; Sig.: significance (p-value).
Predicted variable is the crash category (No, Yes); Reference category (No); Exp(B) = Odds ratio (OR).
Pseudo Nagelkerke R^2 = 0.065
Hosmer and Lemeshow Test: Chi-square = 4.432; Sig. = 0.817

Regarding Table 3, the analysis outputs revealed several predictors with significant ability to predict the outcome variable (probability of being in a crash). A one-level increase in calling-answering daily frequency can increase the odds of being in a crash by 1.080. Drivers who never did any improper driving action due to mobile use or did it only few times were with odds of involvement in a crash that are 0.127 and 0.301 times less than the odds of those who did it many times, respectively. Regarding the responses of drivers towards the phrase “MPU during driving causes distraction”, those who responded as “few times” were with odds of being in a crash that are 0.099 of the odds of those who responded as “always”. That is, they less probable to involve in a road collision. Finally, the drivers who stop their car or reduce their vehicle speed while using mobile are with odds of being in a crash that are 0.075 and 0.228 the odds of those who do not use any safety measure, respectively. This confirms the influence of using mobile in distracting the driver attention, and hence, threating their lives. Other factors that were found with higher relative odds but not statistically significant (at 5% or 10% LOS) are drivers with private cars or taxis, high reading-writing frequency, and daily driving distance.

Table 3. Sequential logistic regression crash model (Block 2)

Variables	B	S.E.	Sig.	Exp(B)
Age	.028	.032	.380	1.029
Gender (1=Male)	.685	.640	.285	1.983
Gender (Ref. Cat.= Female)				
Employment (1=Employed)	.254	1.439	.860	1.290
Employment (2=Self-employed)	.141	1.455	.923	1.151
Employment (3=Students)	.274	1.579	.862	1.315
Employment (Ref. Cat.= others)				
Education (1=Higher studies)	1.172	1.314	.373	3.228
Education (2=University)	1.260	1.229	.305	3.527
Education (3=High school)	1.254	1.210	.300	3.504
Education (4=Intermediate school)	.121	1.385	.930	1.129
Education (Ref. Cat. = primary school or less)				
Mobile use (1=Hand-held)	-.228	.493	.644	.797
Mobile use (2=On speaker)	-.350	.456	.442	.704
Mobile use (Ref. Cat. = Car Bluetooth/headset)				
Vehicle type (1=private cars)	.714	.713	.317	2.041
Vehicle type (2=Taxi)	.632	.832	.447	1.882
Vehicle type (Ref. Cat. = others)				
Daily driving distance	.002	.004	.680	1.002
Daily trips	-.013	.061	.827	.987
Driving experience	-.011	.039	.785	.989
Calling-Answering Frequency	.077	.035	**.030**	1.080
Reading-texting frequency	.043	.044	.334	1.043
Improper driving action frequency (1=never)	-2.064	.641	**.001**	.127
Improper driving action freq. (2=few times)	-1.202	.477	**.012**	.301
Improper driving action freq. (3=sometimes)	-.501	.442	.257	.606
Improper driving action freq. (Ref. Cat=many times)				
MPU distraction (1=few times)	-2.310	.961	**.016**	.099
MPU distraction (2= sometimes)	-.247	.484	.609	.781
MPU distraction (3=often)	.198	.440	.653	1.219
MPU distraction (Ref. Cat.=always)				
MPU safety measure (1=yes-stopping car)	-2.593	1.008	**.010**	.075
MPU safety measure (2=yes-reducing speed)	-1.478	.779	**.058**	.228
MPU safety measure (Ref. Cat. = No)				
Constant	-3.505	2.346	.135	.030

Predicted variable is the crash category (No, Yes); Reference category (No).
Pseudo Nagelkerke R^2 = 0.304
Hosmer and Lemeshow Test: Chi-square = 4.432; Sig. = 0.869

Regarding modelling statistics, Table 3 revealed how the inclusion of driving and attitude driver attributes has increased the Pseudo R-square from 0.06 to 0.30. The Hosmer and Lemeshow tests are insignificant in the two models; p-values equal 0.817 and 0.869 respectively. This is a good indicator as it implies that there is no significant variance between the observed and computed probabilities of the outcome; i.e., crash occurrence.

3.2.2. Near miss Models

The predicted variable (outcome) is the frequency that a driver being involved in near misses (near crashes) which has been coded into three categories (never, few times, several times), with the "several times" as the reference category. As the outcome variable is categorical with ordinal level, the ordinal regression is basically the most suitable modelling technique for prediction. However, the "test of parallel lines" assumption should be firstly tested as otherwise the multinomial logistic regression would be the proper alternative. Two models have been constructed; model No. 1 involves the driver's socioeconomic variables as the explanatory variables whereas model No.2 involves the driver's driving and attitudinal attributes. For the both models, the test of parallel lines assumption was found valid; that is the Chi-square significance values were higher than 0.05 (insignificant) which means accepting the null hypothesis that the location parameters (slope coefficients) are the same across response categories.

A) Near miss model No.1

Table 4 lists the main output details for the developed near miss model based on several socioeconomic driver attributes. According to the table, three factors are with significant contribution in estimating the outcome variable at 5% LOS; these are: driver's age, gender, and the highest education level. Young drivers are more probable to involve in near misses than elderly ones; for example, the odds of drivers in (18-25 years) age groups to be in a near miss is 6.713 the odds of those drivers elder than 55 years. Regarding gender, men drivers have odds to be in multiple near misses that are 2.02 higher than those for women drivers. Finally, drivers with university or high school degree are with higher odds to be in near crashes than those in primary school or less education level.

Finally, regarding modelling goodness and model evaluation statistics, according to Table 4 (footnotes), the final model Chi-Square is significant (sig. = 0.000) which implies the significant contribution of the added variables in comparison with the reduced model which only includes the constant. The model's goodness of fit based on Pearson's Chi-Square test statistic is not significant (sig. = 0.405); this implies there is no significant difference between the predicted and actual category probabilities. Finally, the pseudo Nagelkerke R^2 is 0.107 which is relatively implies weak to medium prediction capacity.

Table 4. Ordinal logistic regression near crash outputs (model No.1).

Variables	B	S.E.	Sig.	Exp (B)
Near miss freq. = 0 (never)	3.215	.856	.000	-
Near miss freq. = 1 (few times)	4.415	.869	.000	-
Near miss freq. = 2 (several times) *	-	-	-	-
Age Cat. (1= 18-25)	1.904	.647	**.003**	6.713
Age Cat. (2= >25-35)	1.421	.514	**.006**	4.141
Age Cat. (3= >35-45)	1.464	.504	**.004**	4.323
Age Cat. (4= >45-55)	.726	.512	.157	2.067
Age Cat. (Ref. Cat. = >55)	0^a	-	-	-
Gender (1= male)	.703	.273	**.010**	2.020
Gender (Ref. Cat. = female)	0^a	-	-	-
Education (1= Higher studies)	.972	.539	.072	2.643
Education (2= University)	1.158	.471	**.014**	3.184
Education (3= High school)	1.095	.477	**.022**	2.989
Education (4= Intermediate school)	.091	.571	.873	1.095
Education (Ref. Cat. = Primary Sch. or less)	0^a	-	-	-
Employment (1= employed)	.613	.620	.323	1.846
Employment (2= self-employed)	.932	.637	.143	2.540
Employment (3= student)	.095	.751	.900	1.100
Employment (Ref. Cat. = others)	0^a	-	-	-
Mobile use (1= hand-held)	-.044	.273	.871	0.957
Mobile use (2= on speakers)	-.040	.253	.874	0.961
Mobile use (Ref. Cat. = Bluetooth or headset)	0^a	-	-	-

- Predicted variable is the near miss frequency (never, few times, several times); Ref cat. (several times).
- Model fitting information: Final model Chi-Square is significant (sig. = 0.000).
- Goodness of fit based on Pearson's Chi-Square (sig. = 0.405).
- Pseudo Nagelkerke R^2 = 0.107
- Test of parallel lines: General Chi-Square (sig. = 0.750).

B) Near miss model No.2

Table 5 shows the output details for the developed near miss ordinal model based on several driving and attitudinal driver attributes. According to the table, three predictors are with significant contribution in estimating the outcome variable. Three of them are significant at 5% LOS (calling/answering frequency, driving experience, and improper action frequency) and two of them at 10% LOS (reading/texting frequency and the type of vehicle). For phoning use, a one-unit increase in the calling/answering or reading/texting daily frequency would increase the odds of being in several times near miss category by a factor of 1.103 and 1.049, respectively. Private car drivers have odds to be in multiple near miss that are 1.774 times higher the odds of truck and buses drivers; that is probably because of their relative speedy driving combatively to taxi, bus and truck drivers. For driving experience, in general, the odds that drivers with less than 20 years driving experience would be in a near crash several times are higher than those odds of those drivers with more than 25-year experience. This is probably because the later drivers are elderly and they are more conservative. Finally, drivers who never conducted improper action due to mobile use or who did that only few times were less likely to be in multiple near crashes. In numbers, they are with corresponding odds of 0.094 and 0.356 than the odds of those who did such actions occasionally.

Table 5. Ordinal logistic regression near crash outputs (model No. 2).

Variables	B	S.E.	Sig.	Exp(B)
Near miss freq. = 0 (never)	1.330	.856	.120	-
Near miss freq. = 1 (few times)	2.864	.864	.001	-
Near miss freq. = 2 (several times) *	-	-	-	-
Calling Answering frequency	.098	.023	.000	1.103
Reading Writing frequency	.048	.028	.092	1.049
Vehicle type (1= private car)	.573	.321	.074	1.774
Vehicle type (1= taxi)	.641	.453	.157	1.898
Vehicle type (Ref. Cat. = others)	0[a]	-	-	-
Driving experience (1= 0 – 5 years)	1.616	.592	.006	5.033
Driving experience (2= 5 – 10 years)	1.881	.595	.002	6.560
Driving experience (3= 10 – 15 years)	1.463	.596	.014	4.319
Driving experience (4= 15 – 20 years)	1.856	.597	.002	6.398
Driving experience (5= 20 – 25 years)	1.014	.667	.129	2.757
Driving experience (Ref. Cat. > 25 years)	0[a]	-	-	-
Improper driving action freq. (1= never)	-2.361	.369	.000	0.094
Improper driving action freq. (2= few times)	-1.034	.324	.001	0.356
Improper driving action freq. (3=sometimes)	-.277	.330	.401	0.758
Improper driving action freq. (Ref. Cat. = many times)	0[a]	-	-	-
MPU safety measure (1=yes-stopping car)	-.548	.596	.358	0.578
MPU safety measure (2=yes-reducing speed)	-.280	.559	.616	0.756
MPU safety measure (Ref. Cat. = No)	0a	-	-	-

- Predicted variable is the near crash frequency (never, few times, several times); Ref cat. (several times).
- Model fitting information: Final model Chi-Square is significant (sig. < 0.001).
- Goodness of fit based on Pearson's Chi-Square (sig. = 0.385).
- Pseudo Nagelkerke R^2 = 0.380
- Test of parallel lines: General Chi-Square (sig. = 0.590).

With respect of modelling goodness, the footnotes of Table 5 show that the final model Chi-Square is significant (sig. < 0.001) which implies the significant contribution of the added variables in comparison with the constant-only model. The model's goodness of fit based on Pearson's Chi-Square test statistic is not significant (sig. = 0.385); this means there is no significant difference between the predicted and actual category probabilities. Finally, the pseudo Nagelkerke R^2 is 0.38 which relatively implies sufficient model prediction capacity.

4. CONCLUSIONS

This paper inspected the distinct characteristics of drivers who use their phone while driving and to explored the effect of these characteristics in rising the likelihood of being in a crash or a near crash. The frequency analysis revealed a high use of mobile phones for calling and texting activities while driving. About 40% of drivers used handheld phones which represent a serious challenge for safe driving. Almost 20% and 55% of them disclosed that they involved in a crash or in a near crash because of such phone use, respectively. Regarding the developed logistic regression models, for the crash involvement model, the sequential regression analysis revealed that drivers who are males and with university education are more likely to be involved into a crash than others. The mobile use by handheld or on speaker and high calling-answering frequency could also increase the possibility of crash occurring. Drivers who are less frequently do inadequate driving behavior or those who stop their car or reduce travel speed when using mobile phones are with low probability in getting into a crash. This confirms the influence of using mobile in distracting the driver attention, and hence, threating their lives. For the near-miss models, the ordinal logistic models revealed that young and men drivers are more probable to involve in near misses than others. High frequency of MPU while driving and conducting improper driving behavior due to this use can promotes the probability of being in multiple near crash occasions. The findings of this paper can aid in developing effective policy and efficient action plan regarding mobile phone use and their challenging consequences.

5. REFERENCES

Asbridge, M., Brubacher, J.R. and Chan, H., (2013). Cell phone use and traffic crash risk: a culpability analysis. International journal of epidemiology, 42(1), pp.259-267.

Bastos, J.T., Santos, P.A.B.D., Amancio, E.C., Gadda, T.M.C., Ramalho, J.A., King, M.J. and Oviedo-Trespalacios, O., (2020). Naturalistic driving study in Brazil: an analysis of mobile phone use behavior while driving. International journal of environmental research and public health, 17(17), p.6412.

Bener, A., Crundall, D., Özkan, T. and Lajunen, T., (2010). Mobile phone use while driving: a major public health problem in an Arabian society, State of Qatar—mobile phone use and the risk of motor vehicle crashes. Journal of Public Health, 18, pp.123-129.

Central Statistical Organization (2023a). CSO website; Retrieved September, 2023, from: https://www.cosit.gov.iq/ar.

Central Statistical Organization. (2019). Traffic Accidents Statistics for 2018. Directorate of Transport and Communication Statistics, CSO, Baghdad, Iraq. Retrieved April, 2023, from: https://cosit.gov.iq/ar/2015-11-23-08-05-11.

Central Statistical Organization. (2020a). Traffic Accidents Statistics for 2019. Directorate of Transport and Communication Statistics, CSO, Baghdad, Iraq. Retrieved April, 2023, from: https://cosit.gov.iq/ar/2015-11-23-08-05-11.

Central Statistical Organization. (2020b). Population Estimation of Iraq 2019. Directorate of Population & Labour Force Statistics 2019. Baghdad, Iraq: CSO. Retrieved June 2023, from http://cosit.gov.iq/ar/pop-main/manpower.

Central Statistical Organization. (2021a). Traffic Accidents Statistics for 2020. Directorate of Transport and Communication Statistics, CSO, Baghdad, Iraq. Retrieved April, 2023, from: https://cosit.gov.iq/ar/2015-11-23-08-05-11.

Central Statistical Organization. (2021b). Annual Statistical Group Report. Baghdad, Iraq: CSO. Retrieved August 12, 2022, from https://cosit.gov.iq/ar/1129-aas.

Central Statistical Organization. (2023b). Traffic Accidents Statistics for 2022. Directorate of Transport and Communication Statistics, CSO, Baghdad, Iraq. Retrieved April, 2023, from: https://cosit.gov.iq/ar/2015-11-23-08-05-11.

Chen, Y., Fu, R., Xu, Q. and Yuan, W., (2020). Mobile phone use in a car-following situation: Impact on time headway and effectiveness of driver's rear-end risk compensation behavior via a driving simulator study. International journal of environmental research and public health, 17(4), p.1328.

Claveria, J.B., Hernandez, S., Anderson, J.C. and Jessup, E.L., (2019). Understanding truck driver behavior with respect to cell phone use and vehicle operation. Transportation research part F: traffic psychology and behaviour, 65, pp.389-401.

Cordellieri, P., Lausi, G., Fraschetti, A., Mari, E., Burrai, J., Quaglieri, A., Paoli, E. and Giannini, A.M., (2022). Young Drivers 'on the Phone: Personality Traits, Driving Attitudes, Beliefs, and Risk Perception on Using Mobile Phone While Driving. Psychological Studies, 67(4), pp.459-467.

European road safety observatory (ERSO) (2018). Cell phone use while driving 2018 Summary.

Fink, A. (2017). How to conduct surveys: A step-by-step guide. 6th ed. Sage Publications. ISBN: 9780203505991.

Fraschetti, A., Cordellieri, P., Lausi, G., Mari, E., Paoli, E., Burrai, J., Quaglieri, A., Baldi, M., Pizzo, A. and Giannini, A.M., (2021). Mobile phone use "on the road": A self-report study on young drivers. Frontiers in psychology, 12, p.620653.

IBM Corp. (2021). IBM SPSS Regression 28. International Business Machines Corporation.

National Center for Statistics and Analysis (NCSA). (2023). Distracted driving 2019 (Report No. DOT HS 813 111). National Highway Traffic Safety Administration.

Orme, J.G. and Combs-Orme, T., (2009). Multiple regression with discrete dependent variables. Oxford University Press.

Park, H. A. (2013). An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain. Journal of Korean Academy of Nursing, 43(2), 154-164.

Rea L. M. and Parker R. A (2014). Designing and conducting survey research: A Comprehensive Guide. 4th ed., Jossey-Bass. ISBN 978-1-118-76700-9 (ebk).

Rudisill, T. M., Baus, A. D., & Jarrett, T. (2019). Challenges of enforcing cell phone use while driving laws among police: a qualitative study. Injury prevention, 25(6), 494-500.

Saunders, M., Lewis, P. and Thornhill, A., (2012). Research methods for business students. 6th edition Pearson Education Limited, UK.

Spyropoulou, I. and Linardou, M., (2019). Modelling the effect of mobile phone use on driving behaviour considering different use modes. Journal of advanced transportation, 2019.

Tabachnick, B. G. and Fidell, L. S. (2019). Using multivariate statistics (7th ed.). Pearson Education, Inc. ISBN 9780134790541.

Truelove, V., Oviedo-Trespalacios, O., Freeman, J. and Davey, J., (2021). Sanctions or crashes? A mixed-method study of factors influencing general and concealed mobile phone use while driving. Safety science, 135, p.105119.

Wang, X., Xu, R., Asmelash, A., Xing, Y. and Lee, C., (2020). Characteristics of driver cell phone use and their influence on driving performance: A naturalistic driving study. Accident Analysis & Prevention, 148, p.105845.

WHO (2018). Global status report on road safety 2018. Geneva. World Health Organization. ISBN: 9789241565684. https://doi.org/10.1016/j.amar.2019.100102.

World Health Organization, (2011). Mobile phone use: a growing problem of driver distraction.