Accident Analysis and Prevention 206 (2024) 107713

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Analyzing fatal crash patterns of recidivist drivers across genders and age Groups: A hazard-based duration approach

Richard Dzinyela [a], Emmanuel Kofi Adanu [b,*], Hardik Gupta [a], Pranik Koirala [a], Nawaf Alnawmasi [c], Subasish Das [d], Dominique Lord [a]

[a] Zachary Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX 77843-3136, USA
[b] Department of Civil, Construction and Environmental Engineering, The University of Alabama, Tuscaloosa, AL, USA
[c] Civil Engineering Department, College of Engineering, University of Hail, 55474, Saudi Arabia
[d] Civil Engineering Program, Texas State University, 601 University Dr, San Marcos, TX, 78666, USA

ARTICLE INFO

Keywords:
Recidivist Drivers
Hazard Duration
DUI
Repeated Offenders
Speeding Violations

ABSTRACT

Identifying factors that significantly affect drivers that are repeatedly involved in traffic violations or non-fatal crashes (defined here as recidivist drivers) is very important in highway safety studies. This study sought to understand the relationship between a set of variables related to previous driving violations and the duration between a previous non-fatal crash and a subsequent fatal crash, taking into account the age and gender of the driver. By identifying the characteristics of this unique driver population and the factors that influence the duration between their crash events strategies can be put in place to prevent the occurrence of future and potentially fatal crashes. To do this, a five-year (2015–2019) historical fatal crash data from the United States was used for this study. Out of 15,956 fatal crashes involving recidivist drivers obtained, preliminary analysis revealed an overrepresentation of males (about 75%). It was also found that the average duration between the two crash events was about a year and a half, with only an average of one month difference between male and female drivers. Using hazard-based duration models, factors such as number of previous crashes, previous traffic violations, primary contributing factors and some driver demographic characteristics were found to significantly be associated with the duration between the two crash events. The duration between the two events increased with driver's age for drivers who were involved in only one previous crash and the duration was shorter for those that were previously involved in multiple crashes. Previous DUI violations, license suspensions, and previous speeding violations were found to be associated with shorter durations, at varying degrees depending on the driver's age and gender. The duration was also observed to be longer if the fatal crash involved alcohol or drug use among younger drivers but shorter among middle-aged male drivers. These findings reveal interesting dynamics that may be linked to recidivist tendencies among some drivers involved in fatal crashes. The factors identified from this study could help identify crash countermeasures and programs that will help to reform such driver behaviors.

1. Introduction

Risky human behaviors are one of the numerous factors that contribute to crashes. For example, studies by (Tao et al., 2017; Dingus et al., 2016; Petridou and Moustaki, 2000; Bucsuházy et al., 2020; Shams and Rahimi-Movaghar, 2009; Ivers et al., 2009; Zhao et al., 2013; Adanu et al., 2019; Adanu and Jones, 2019) have shown that vast majority of road crashes could be linked with risky human driving behaviors. Reason et al. (1990) revealed that crash occurrences due to traffic violations and error/lapses are mainly attributed to risky driver behaviors. Similar findings by Tillmann and Hobbs (1949) and Treat et al. (1979) also attribute traffic violations and errors made by drivers to traffic crashes. Indeed, driving behaviors and styles are mainly influenced by both external factors and individual attributes. These factors may be peculiar to some groups of drivers or reflect a broader societal characteristic. One group of drivers that exhibit risky or crash prone driving

* Corresponding author.
E-mail addresses: dzinyela_1@tamu.edu (R. Dzinyela), ekadanu@crimson.ua.edu (E. Kofi Adanu), hardik_gupta@tamu.edu (H. Gupta), pranik@tamu.edu (P. Koirala), n.alnawmasi@uoh.edu.sa (N. Alnawmasi), subasish@txstate.edu (S. Das), dlord@civil.tamu.edu (D. Lord).

https://doi.org/10.1016/j.aap.2024.107713
Received 4 February 2024; Received in revised form 10 June 2024; Accepted 8 July 2024

behaviors are recidivist or repeat offender drivers. Repeated traffic law offenders pose a significant challenge to road safety, often exhibiting patterns of risky behavior that can lead to severe injury crashes. These individuals, who frequently accumulate multiple violations, demonstrate a disregard for traffic regulations and the safety of others. Previous studies have revealed that a high number of recidivist drivers often get involved in severe injury crashes resulting from driving under the influence of alcohol or drugs. For example, a study by Fell (1993) used 1988 data from the United States of America's Department of Justice and found that about 3.3 % of all licensed drivers have been re-arrested for driving while intoxicated (DWI) in the past three years. In addition, using data from the National Highway Traffic Safety Administration (NHTSA)'s Fatality Reporting System (FARS) revealed that 5.7 percent of all drivers in fatal crashes were previously arrested and charged with DWI in the past three years. In another study, Simpson and Mayhew (1991) separated first offenders from repeated offenders using FARS data and found that the average blood alcohol concentration (BAC) of recidivist drivers was 0.21 as compared to 0.17 for first time offenders. This finding is consistent with Gould and Gould (1992) study result that found the average BAC of the repeated offenders to be 0.18 compared to 0.15 in Louisianna. Furthermore, repeat DWI offenders were 50 % more involved in both alcohol and non-alcohol related crashes as compared to first offenders.

In some instances, these subsequent crashes can be fatal and may occur relatively soon after previous incidents. While recent studies have shifted their focus towards adopting a safe systems approach (Roberts et al., 2014; Hiekmann, 2012; Khan and Das, 2024), there remains a need to address and minimize behavioral risks of crash prone drivers. This study seeks to identify driver behaviors and crash variables that influence the time it takes between a previous non-fatal crash and a subsequent fatal crash among such crash prone drivers in the US and assesses whether these factors significantly differ by gender and age. The study further investigates how their previous driving records may have contributed to the length of time between the last previous crash and the eventual fatal crash. Utilizing fatal crash records of male and female recidivist drivers obtained from the Fatality Analysis Reporting System (FARS), the study applies a hazard-based duration modeling technique to understand how factors like previous speeding violation, DWI, number of previous crashes and demographic characteristics impact the duration between the most recent non-fatal crash and a subsequent fatal crash for male and female recidivist drivers. The findings of this research will aid in creating technical, socio-cultural, and psychological intervention programs designed to influence drivers involved in crashes, helping them to avoid behaviors that could lead to future fatal crashes. To the best of our knowledge, this is the first study to analyze crashes involving recidivist drivers by both age and gender, utilizing a hazard duration model on a nationally representative dataset.

2. Literature review

Considerable research has been carried out in the past to explore the factors that are associated with recidivist drivers. Crash proneness theory was considered as an accident theory during the period covering 1920–1950. The studies conducted in recent years explored several dynamic and evolving areas on this topic. The motivations for recidivism can vary widely among individuals and may encompass factors such as peer influence, lenient law enforcement, absence of consequences, risky behavior, and psychological factors. Numerous studies have identified correlations between repeat offenders and issues like drunk driving or an elevated crash risk across diverse age groups (Bogstrand et al., 2015; Summala et al., 2014; Ayuso et al., 2010; Factor, 2014). Driving under the influence (DUI) or driving while intoxicated (DWI) emerged as a notable factor contributing to fatal crashes among recidivist drivers, with observed recidivism rates ranging from 21 % to 47 % (Yu and Williford, 1995). Demographic characteristics such as age, gender, and occupation have shown significant influences on DUI recidivism (Ryan et al., 1996). A study tracking teenage drivers who were arrested for serious offenses found that 56 % of the drivers committed similar offenses within six months of their initial arrest, and 14 % repeated offenses more than once. Additionally, male teen drivers were reported to be 8 to 21 times more likely to re-offend (Manno et al., 2012). Findings from other previous studies associated with recidivism are summarized in Table 1.

3. Data description

Recidivist drivers killed in crashes between 2015 and 2019 were collected from the Fatality Analysis Reporting System (FARS) database. Out of 15,956 fatal crashes recorded in this period, an overwhelming 75.3 % (12,015) were males whiles the remaining 24.7 % (3,941) were females. Analysis of the crash data revealed notable differences in the involvement of males and females across various crash categories. Among these, seatbelt use exhibited the most significant disparity, with 53.44 % of males not wearing seat belts while involved in the fatal crash compared to 42.16 % of females (see Fig. 2). The Chi-Square Test result showing the difference between seatbelt use across gender was significant, $X^2\,(1, N = 15956) = 151.205, p < 0.001$.

Regarding previous traffic violations before the fatal crash, about 28.50 % of male drivers previously had their license suspended compared to 22.50 % of female recidivist drivers [$X^2\,(1, N = 15956) = 54.089, p < 0.001$]. About 8.58 % of male recidivist were previously charged with driving under the influence of alcohol as compared to 5.48 % of female recidivist drivers [$X^2\,(1, N = 15956) = 39.618, p < 0.001$]. Regarding speeding, about 31.94 % of male recidivist drivers were charged with speeding violations compared to 29.07 % of female recidivist drivers [$X^2\,(1, N = 15956) = 11.377, p < 0.001$]. The percentage of previous offense by recidivist drivers is shown in Fig. 1.

Similar distribution of crashes by gender was observed in traffic violations during the time of fatal crash. As shown in Fig. 2, about 15.25 % of male drivers were driving with invalid license before the fatal crash compared to 11.44 % of female drivers[$X^2\,(1, N = 15956) = 35.202, p < 0.001$]. About 14.14 % male drivers compared to 12.73 % of female drivers were also observed to be driving under the influence of drugs at the time of the fatal crash [$X^2\,(1, N = 15956) = 4.927, p < 0.001$]. Similarly, about 21.07 % of males and 15.83 % of females were driving while drunk before being killed in a fatal crash [$X^2\,(1, N = 15956) = 51.416, p < 0.001$]. Fig. 2 further showed that fatal crashes involving non-seatbelt use [$X^2\,(1, N = 15956) = 151.205, p < 0.001$] and overspeeding [$X^2\,(1, N = 15956) = 103.459, p < 0.001$] had a significant proportional difference between male and female drivers.

In terms of demographics, a higher percentage of female drivers between 16–24 years and more than 64 years (20.45 % for drivers between 16–24 years and 17.55 % for drivers older than 64 years) were observed to be involved in fatal crashes compared to male drivers (19.19 % for drivers between 16–24 years and 16.33 % for drivers older than 64 year).The Chi-square test result revealed a 90 % confidence interval in difference between younger drivers [$X^2\,(1, N = 15956) = 2.980, p < 0.01$] and older drivers [$X^2\,(1, N = 15956) = 3.219, p < 0.01$] based on gender. Conversely, about 64.47 % of male drivers were involved in fatal crashes as compared to 61.97 % of female drivers (see Fig. 3). Considering race, about 78.77 % of Caucasian female drivers were involved in fatal crash compared to 75.32 % for male drivers [$X^2\,(1, N = 15956) = 19.383, p < 0.001$]. However, the percentage of Hispanic and African- American drivers were higher for males than females [$X^2\,(1, N = 15956) = 17.490, p < 0.001$].

In comparing the number of previous crashes before fatal crash, 77.07 % of females were only involved in a non-fatal crash once compared to 75.8 % of males. Males on the other hand were involved in more than one crash before being involved in a fatal crash as shown in Fig. 4.

On average, male recidivist drivers take a shorter time (18.06 months) after their most recent non-fatal crash to be involved in a fatal

Table 1
Summary of previous literature related to recidivist drivers.

Author	Location	Method Used	Findings
Ferrante et al., (2001).	Australia	Cox Proportional Hazards Model	Young male drivers (<25 years) were more likely to experience repeat arrests for drink-driving offenses. Individuals with a background of prior criminal history were more prone to a repeated alcohol-related crash
Padilla et al., (2018).	Spain	Two-way ANOVA and Logistic Regression	Higher alcohol consumption and incautious driving were significantly associated with a higher likelihood of reoffending. In addition, there were overrepresentation of male drivers in connection between alcohol consumption and accidents records.
Tassoni et al. (2016)	Italy	Mantel-Haenszel Statistic	There was no statistically significant difference between the two genders regarding repeated offenses related to alcohol. Cocaine consumption was significantly associated with an increased probability of recidivism.
Møller et al., (2015)	Denmark	Logistic Regression	Female drunk drivers and individuals above 60 years were 57 % and 58 %, respectively, less likely to become repeat offenders.
Fell (1995)	United States	Statistical Analysis (p-test)	Approximately one-third of all annually arrested drivers for DUI are repeated offenders
Lawpoolsri et al. (2007)	United States	Cox Proportional Hazards Model	Motorists charged with a speeding offense were more likely to get additional speeding tickets in the future. The study also found that the less severe penalty lowered the likelihood of repeated offenses more effectively than harsher penalties.
Watson et al. (2015)	Australia	Chi-squared and Cramer's V	Men, young individuals, and individuals with previous traffic offense history were more likely to repeatedly engage in high-range speeding. The limitation of the study involved the use of data collected primarily for administrative purposes, not research, which can potentially impact the accuracy of the findings.
Watson et al. (2017)	Australia	Cochran-Mantel-Haenzel and Chi-squared	The study revealed that individuals exhibited a higher re-offense rate both before and after license disqualification compared to the period during the disqualification. The repeated drunk driving offenses were highest before license disqualification, consistent across genders and age group. Limitation: The study only includes individuals caught for an offense. Individuals who managed to evade detection are not included which might affect the results.
McKnight & Tippetts (1997)	United States	Chi-squared	The study examined the effectiveness of two driver improvement approaches on recidivism. It was found that the participants assigned to the recidivism prevention course (RPC) had significantly lower crashes (18 %) or violations (8 %) in the following years compared to participants assigned to the accident prevention course (APC).
McMillen et al. (1992)	United States	Multi-variate ANOVA and ANOVA	DUI recidivists reported consuming more alcohol and had significantly higher blood alcohol concentration (BAC) at the time of arrest compared to single offenders. DUI recidivists had significantly more non-traffic arrests, with a frequency three times higher than single offenders.
C'de Baca et al. (2001) (a)	United States	Sequential Logistic Regression	The study identified factors that predicted DWI recidivism. It was found that factors like age (<29 years), years of education (<12 years), arrest blood alcohol concentration, score on the receptive area of the Alcohol Use Inventory and MacAndrews Alcoholism Scale (MAC) of the MMPI-2 were significant variables to predict the same.
Schmitz et al. (2014)	Brazil	Statistical Analysis and Poisson Regression	The study found that the majority of the recidivists were between 41 to 50 years of age, had driving license for more than 12 years, and had low education. Limitation: The study focused on only one state in Brazil which limits the applicability of results to other contexts.
C'de Baca et al. (2001) (b)	United States	Cox Proportional Hazards Model	Referring first-time offenders (male or female) to Victim Impact Panels (VIPs) did not increase the likelihood of DUI recidivism. However, the study found that VIPs have a negative impact on female recidivists as female recidivists referred to VIPs were significantly more likely to be re-arrested than those not referred.
Lijarcio et al. (2022)	Spain	Multi-Group Structural Equation Modeling	It was found that recidivist traffic offenders who had more traffic offenses had lower risk perceptions regardless of their gender. Similarly, older recidivist traffic offenders had higher risk perceptions regardless of their gender. The study also found that, unlike females, more driving time had a positive relation with risk perception in males. Limitations: The study used anonymous interviews which may not fully eliminate biases on sensitive topics. The variables used were limited and there could be additional variables affecting risk perception of the driver.
Chen & Jou (2018)	China	Logistic Regression and Multilevel random effects logistic model	The study found that the increase in divorce rates in an area significantly increased the risk of DUI recidivism. The study model also predicted that the DUI recidivists are more likely to be male, motorcycle riders, and drive on Friday or Saturday. Limitations include the study using only the detected DUI drivers for analysis which can potentially create a bias between the recidivists and first-time offenders.
Das et al.,(2015)	United States	Logistic regression	The study found that about 34 % of crashes are committed by recidivist drivers who formed only 5 % of total licensed drivers in Louisiana.

crash compared to female drivers (19.37 months). However, the difference is marginal, about 1.3 months as shown in Fig. 5. The likelihood ratio test revealed that the data could be split and modeled both by gender and age groups. Table 2 present a descriptive statistic of all the variables considered in this study. Fig. 6 shows the frequency distribution for the time until the fatal crash after the latest non-fatal crash. The figure shows that a general downward trend in the risk of involving in a fatal crash after a few months of having a non-fatal crash for young and middle-aged male and middle-age female drivers. However, figure shows that a generally constant trend with time in the involvement risk in fatal crash after being a non-fatal crash for young and old female and old male drivers.

4. Methodology

Hazard-based duration models have been used in studies to analyze incident duration, representing the time elapsed from the start of an event until its end (Garib et al., 1997; Nam & Mannering, 2000). This method has been used extensively in different fields like biomedical, social sciences and engineering (Nam & Mannering, 2000). In transportation engineering, hazard-based duration modeling has been applied in incident clearance times on highways (Islam et al., 2022), the duration of shopping activities (Bhat, 1996), and the analysis of urban travel time (Anastasopoulos, 2012), among other many studies. Hazard duration models have also been used in crash data analysis to examine duration data (Islam et al., 2021; Balusu et al., 2020; Alzaffin et al., 2023; Ali et al., 2019; Lord et al., 2021). Sohrabi et al. (2024) used a

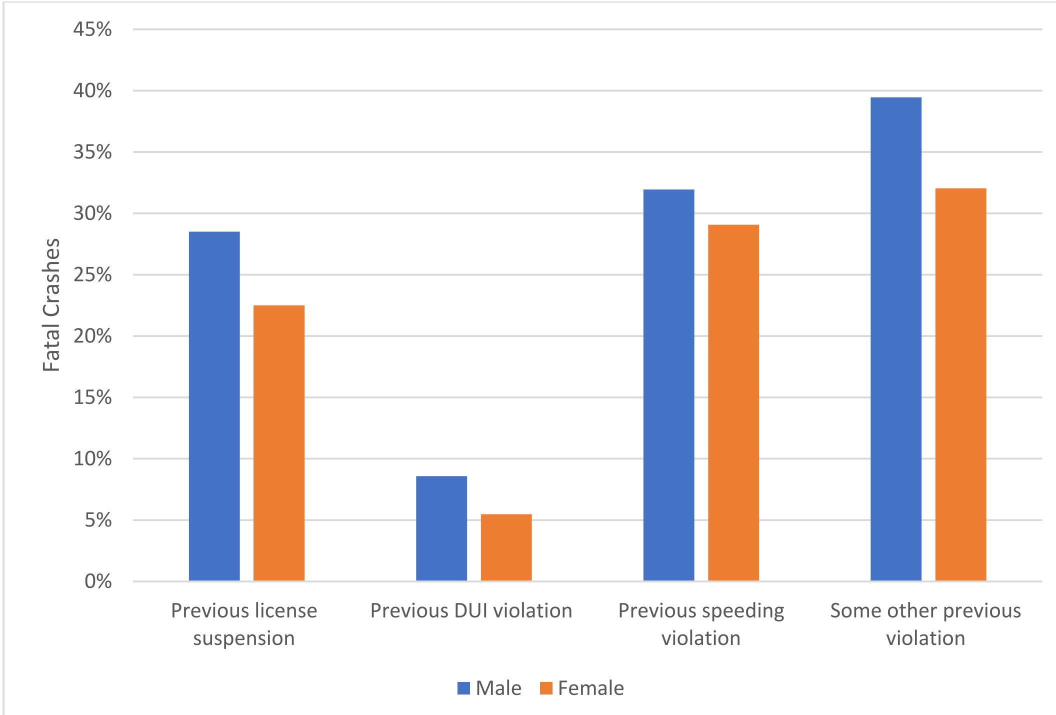

Fig. 1. Percentage of the number of previous traffic violations by males and females.

hazard function for estimating crash risk for automated and human-driven vehicles but used mileage/cumulative distances between crashes rather than time. For this study, the incident duration is defined as the time between a previous non-fatal crash involving a driver and a subsequent crash in which the same driver was killed. When examining incident duration data, a fully parametric hazard-based duration model is utilized to explore the likelihood that a time duration concludes at a specific time t, considering that the duration has persisted until time, t (Washington et al., 2020). In this study, the duration of time from the recent previous non-fatal crash until the fatal crash is represented as a continuous random variable denoted by T, with a cumulative distribution function F(t) which is defined as (Washington et al., 2020):

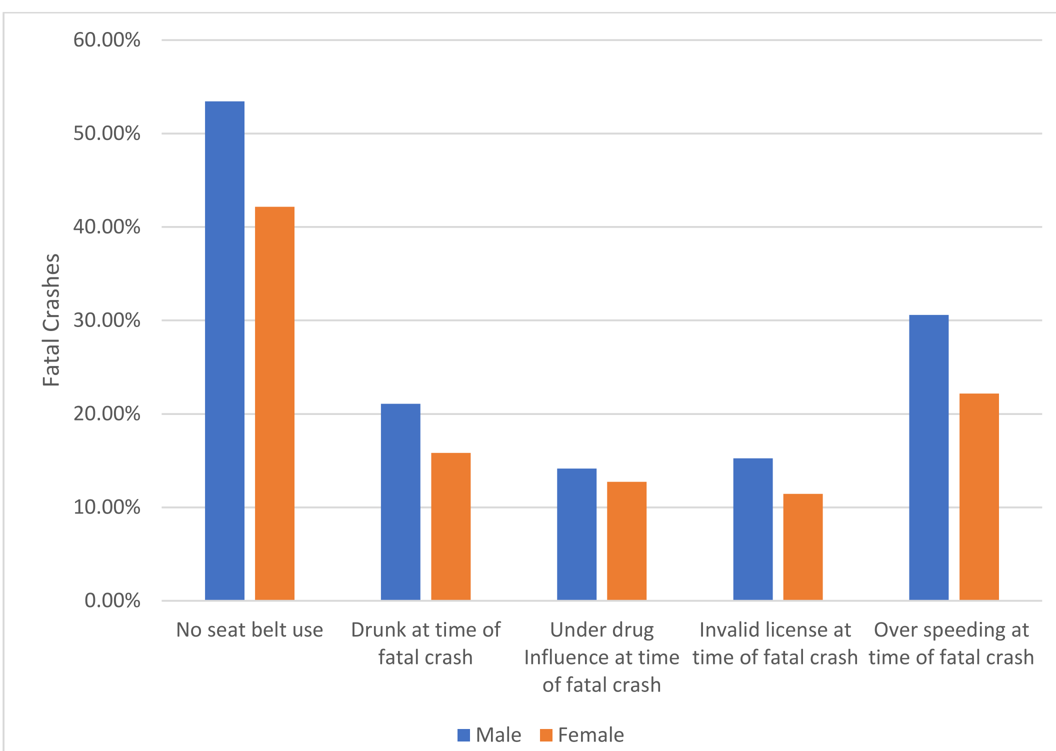

Fig. 2. Percentage of number of traffic violations at time of fatal crashes by gender.

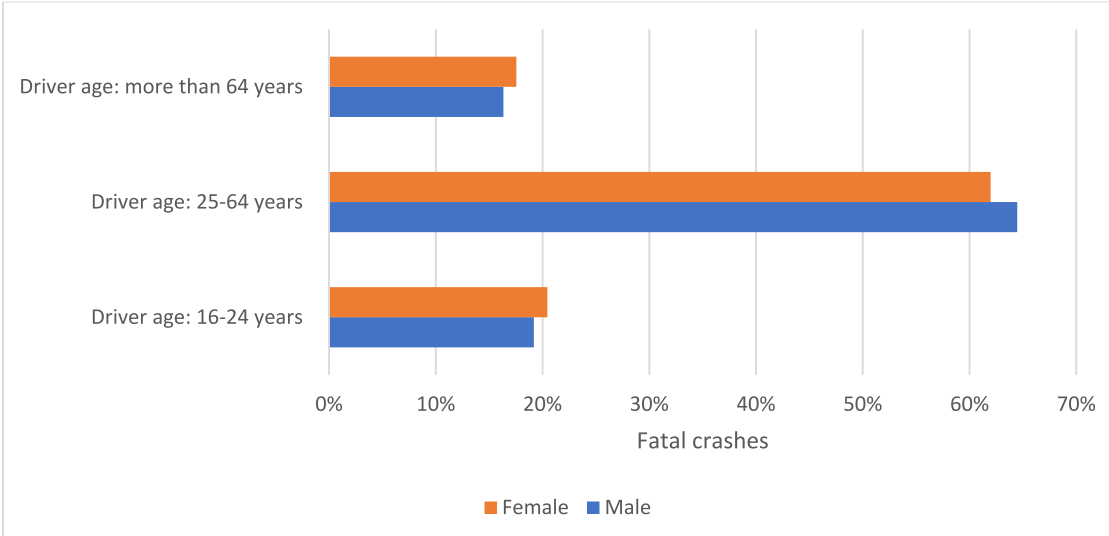

Fig. 3. Percentage of number of recidivist drivers by age groups for males and females.

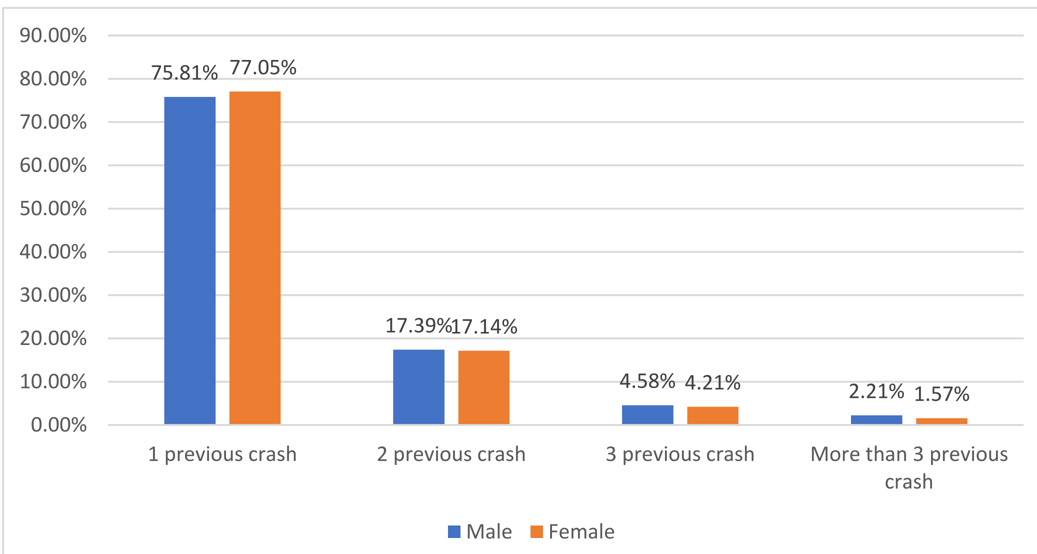

Fig. 4. Number of previous non-fatal crash by gender.

$$\mathrm{F(t)} = \mathrm{P(T \ < \ t)} \tag{1}$$

Alternatively, the survival function, denoted as S(t), represents the probability of the duration between the non-fatal crash and fatal crash occurrence as being greater than or equal to a specific time t. Mathematically, it is defined as:

$$\mathrm{F(t)} = \mathrm{P(T < t)} = 1 - \mathrm{P(T > t)} = 1 - \mathrm{S(t)} \tag{2}$$

The hazard function h(t) signifies the conditional probability that an incident will happen within the interval between time t and t = dt, given that the incident has not occurred up to time t (Washington et al., 2020) and is given by:

$$h(t) = \frac{f(t)}{1 - F(t)} \tag{3}$$

The slope of the hazard function provides insight into the relationship between the probability of an event ending and the duration of the event. Specifically, if the hazard function is decreasing with respect to the incident duration ($dh(dt)/dt < 0$), it implies that the likelihood of the incident concluding diminishes as the duration increases. In contrast, an increasing hazard function ($dh(dt)/dt > 0$) signifies that the probability of the incident ending soon rises as the duration decreases. A hazard function with a zero slope ($dh(dt)/dt = 0$) indicates that the likelihood of the incident concluding soon is independent of the duration it has lasted, suggesting a constant probability of ending regardless of the elapsed time. Mathematically, these relationships can be expressed as derivatives, offering a precise quantitative characterization of the dynamics involved.

The proportional-hazards approach has demonstrated its appropriateness in assessing the impact of explanatory variables, or covariates, on an incident. These covariates represent a multiplicative influence on specific baseline hazard functions, denoted as $h_0(t)$ (Washington et al., 2020; Greene, 2012) and is expressed as:

$$h(t|X) = h_0(t)e^{\beta X} \tag{4}$$

where $e^{\beta X}$ represents the effect of explanatory factors on the hazard, X is the vector of external contributing factors and β is the vector of estimable parameters.

The proportional-hazards and the accelerated failure time approaches have demonstrated their appropriateness in assessing the

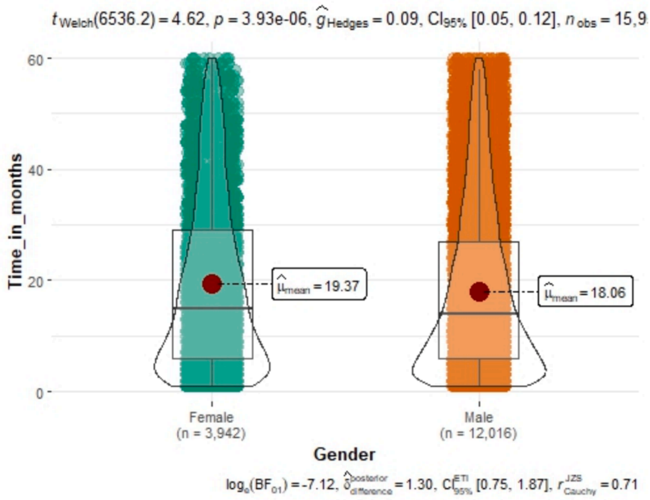

Fig. 5. A violin plot showing the time between previous non-fatal crash and fatal crash for males and female recidivist drivers.

explanatory variables impact on the hazard function. The proportional hazards model presumes the hazard ratio is constant over time, however the accelerated failure time approach assumes the explanatory factors rescale time directly into a baseline survivor function. The time until a fatal crash natural logarithm, T, is expressed as a linear function of explanatory factors in the accelerated failure time approach as (Washington et al., 2020; Greene, 2012);

$$\ln(T) = \beta\mathbf{X} + \varepsilon \tag{5}$$

where **X** is the vector of external contributing factors, β is the vector of estimable parameters and ε is the error term. The conditional survival and hazard functions could be written as (Washington et al., 2020);

$$\mathrm{S}(t|\boldsymbol{X}) = S_0[t\mathrm{EXP}(\beta\mathbf{X})] \tag{6}$$

$$\mathrm{h}(t|\boldsymbol{X}) = h_0[t\mathrm{EXP}(\beta\mathbf{X}]\mathrm{EXP}(\beta\mathbf{X}) \tag{7}$$

where $h_0(t)$ is the inherent baseline hazard function where all the explanatory variable vectors are zero.

Distributional assumption is needed to estimate fully parametric hazard and survival functions. A wide variety of distributions, including but not limited to the gamma, exponential, Weibull, and log-logistic distributions are used depending on the context and statistical fit (Washington et al., 2020). Among these distributions, the Weibull distribution stands out as the most used parametric framework for analyzing duration data. This popularity stems from its flexibility, allowing the hazard function to exhibit either a monotonically increasing trend (indicating a decreasing probability of duration between events) or a monotonically decreasing trend (indicating an increasing probability of duration between events) (Nam & Mannering, 2000). The Weibull distribution is characterized by parameters $\lambda > 0$ and > 0, and its hazard function is mathematically expressed as:

$$h(t) = (\lambda P)(\lambda t)^{P-1} \tag{8}$$

In hazard-based duration models, the conventional approach operates under the assumption that the impact of any explanatory variable remains consistent across all observations. Nevertheless, there exists a potential scenario where the duration of incidents may not exhibit homogeneity across observations, introducing a source of potential bias in model outcomes. To scrutinize the validity of the homogeneity assumption, a stochastically distributed term is incorporated into the duration models, allowing for the possibility that some or all of the model parameters vary across observations. This stochastic term is

Table 2
Descriptive statistics of variables considered for modeling males and female under different age groups.

	Variable description	Age less than 25 years old	Age between 25 years and 64 years old	Age greater than 64 years
		Mean (standard deviation)	Mean (standard deviation)	Mean (standard deviation)
Male	Time between previous non-fatal crash and subsequent fatal crash (in month)	14.26(12.96)	18.01(14.99)	14.99(22.75)
	No seat belt use at time of fatal crash (1 if yes, 0 otherwise)	0.59(0.49)	0.55(0.50)	0.50(0.40)
	Drunk at time of fatal crash (1 if yes, 0 otherwise)	0.27(0.45)	0.23(0.42)	0.42(0.06)
	Under drug Influence at time of fatal crash	0.17(0.37)	0.16(0.36)	0.36(0.06)
	Driver race/ethnicity: Hispanic (1 if yes, 0 otherwise)	0.15(0.36)	0.10(0.30)	0.3(0.05)
	Driver race/ethnicity: Caucasian (1 if yes, 0 otherwise)	0.75(0.43)	0.74(0.44)	0.44(0.81)
	Driver race/ethnicity: African American (1 if yes, 0 otherwise)	0.14(0.35)	0.17(0.37)	0.37(0.1)
	Invalid license at time of fatal crash (1 if yes, 0 otherwise)	0.16(0.37)	0.18(0.38)	0.38(0.04)
	Number of previous crashes before fatal crash	1.36(0.71)	1.35(0.76)	0.76(1.29)
	Over speeding at time of fatal crash (1 if yes, 0 otherwise)	0.43(0.5)	0.31(0.46)	0.46(0.12)
	Previous license suspension (1 if yes, 0 otherwise)	0.31(0.46)	0.33(0.47)	0.47(0.09)
	Previous DUI violation (1 if yes, 0 otherwise)	0.09(0.28)	0.1(0.3)	0.3(0.02)
	Previous speeding violation (1 if yes, 0 otherwise)	0.45(0.5)	0.32(0.47)	0.47(0.15)
	Some other previous violation (1 if yes, 0 otherwise)	0.45(0.5)	0.4(0.49)	0.49(0.29)
Female	Time between previous non-fatal crash and subsequent fatal crash (month)	15.89(14)	19.22(15.56)	15.56(23.98)
	No seat belt use at time of fatal crash (1 if yes, 0 otherwise)	0.46(0.5)	0.47(0.50)	0.50(0.23)
	Drunk at time of fatal crash (1 if yes, 0 otherwise)	0.22(0.41)	0.18(0.38)	0.38(0.03)
	Under drug Influence at time of fatal crash	0.15(0.35)	0.15(0.35)	0.35(0.04)
	Driver race/ethnicity: Hispanic (1 if yes, 0 otherwise)	0.11(0.31)	0.08(0.27)	0.27(0.03)
	Driver race/ethnicity: Caucasian (1 if yes, 0 otherwise)	0.78(0.41)	0.78(0.42)	0.42(0.84)
	Driver race/ethnicity: African American African American (1 if yes, 0 otherwise)	0.14(0.35)	0.15(0.36)	0.36(0.08)
	Invalid license at time of fatal crash (1 if yes, 0 otherwise)	0.10(0.30)	0.15(0.35)	0.35(0.02)
	Number of previous crashes before fatal crash	1.29(0.62)	1.35(0.76)	0.76(1.22)
	Over speeding at time of fatal crash (1 if yes, 0 otherwise)	0.29(0.45)	0.24(0.43)	0.43(0.07)
	Previous license suspension (1 if yes, 0 otherwise)	0.23(0.42)	0.27(0.45)	0.45(0.05)
	Previous DUI violation (1 if yes, 0 otherwise)	0.05(0.22)	0.07(0.26)	0.0014(0.03)
	Previous speeding violation (1 if yes, 0 otherwise)	0.36(0.48)	0.32(0.47)	0.47(0.10)
	Some other previous violation (1 if yes, 0 otherwise)	0.34(0.47)	0.34(0.47)	0.47(0.22)

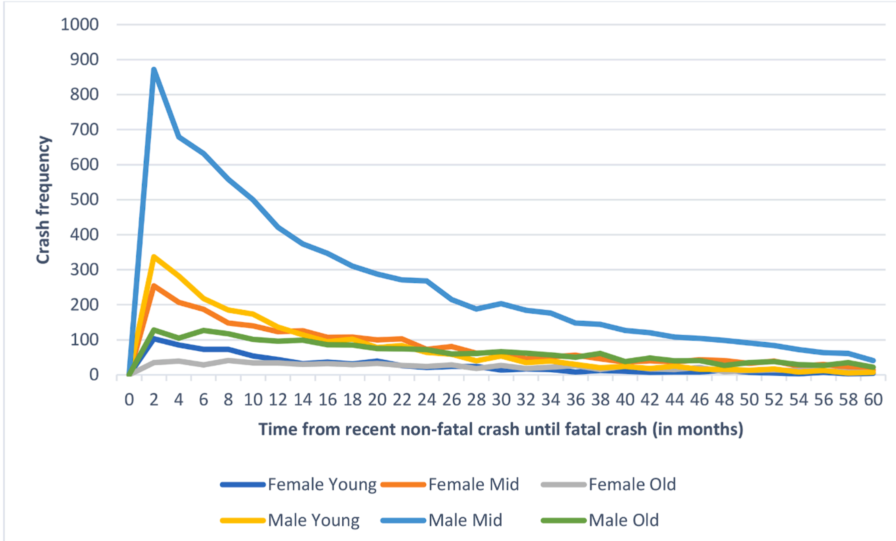

Fig. 6. Frequency distribution of time until the fatal crash after the latest non-fatal crash.

mathematically expressed as an additional component within the model structure, offering a representation that accommodates the potential heterogeneity that could affect the incident durations across the dataset by allowing explanatory variable estimates to vary across observations, and accounts for the explanatory variables' effects in the means and variances of the observation-specific variable estimates. (Alnawmasi and Mannering, 2019; Dzinyela et al., 2024; Greene, 2007; Washington et al., 2020):

$$\beta_{\mathrm{kn}} = \beta + \Theta_n \boldsymbol{Z}_{\mathrm{n}} + \sigma_n EXP(\Psi_n \boldsymbol{W}_n)\nu_n \tag{9}$$

where β_n is the mean parameter estimate across all crashes, $\boldsymbol{Z}_{\mathrm{n}}$ is a vector of crash-specific explanatory variables that captures heterogeneity in the mean that affects bicyclist injury-severity level k, Θ_n is a corresponding vector of estimable parameters, $\boldsymbol{W}_n$ is a vector‘ of crash-specific explanatory variables that captures heterogeneity in the standard deviation σ_n with corresponding parameter vector Ψ_n, and ν_n is a disturbance term. The normal distribution for the random parameters is considered due to its statistical fit superiority in previous research findings (Alnawmasi and Mannering, 2019, Pang et al., 2022). The models, incorporating random parameters with variations in both means and variances, are estimated through the simulation of maximum likelihood using 1000 Halton draws. By introducing heterogeneity in both means and variances, the models attain realistic presentation of the underlying stochastic processes associated with incident clearance durations (Washington et al., 2020). None of the random parameters were found to produce a statistically significant random parameter in the mean nor variance.

5. Likelihood ratio test

To determine if there are significant differences in subsets of the data sample, the likelihood ratio test was used. The data was first stratified by gender and then three age groups (driver age less than 25, 25–64 years and driver over 64 years). Several estimated models using data from these sample subsets were compared against a single model estimated on all subset dataset combined using the test statistics:

$$\mathrm{X}^2 = -2[LL(\beta) - \sum_{i=1}^{k} LL(\beta_k)] \tag{9}$$

where LL(β) is the log-likelihood at convergence of model estimated on full dataset and LL(β_k) is log-likelihood at convergence of model estimated on the subset dataset.

After separating the data based on gender and age groups, there was a total of six classes. Looking at the different age groups found within the gender category, the x^2 test statistic for the different age groups in the male driver's category is −2*[-16436- (−3238.3267—10624.81568—2460.11701)] = 225.4812. With a degree of freedom equal to 23, the null hypothesis that parameters are equal for all considered age groups in the male recidivist dataset can be rejected with over 99.99 % confidence. For female datasets, the x^2 = -2*[-5314.08 – (−1132.84145—3299.42152—834.19623)] = 95.25 and a degree of freedom equal to 13. This result also showed that the null hypothesis that parameters are equal for all age groups in the female dataset can be rejected at over 99.99 % confidence interval. Finally comparing all the six groups with a single model estimated with the entire data sets, the x^2 = -2*[-21753.8 – (−3238.3267—10624.81568—2460.11701–1132.84145—3299.42152—834.19623)] = 328.2496 and a degree of freedom equal to 47. This result also showed that the null hypothesis that parameters are equal for the entire datasets can be rejected at over 99.99% confidence intervals.

6. Model estimates results

Using the random parameter hazard duration model, the estimation results for male and female recidivist drivers of different age groups involved in fatal crashes were summarized in Table 3 and Table 4, respectively. The model results were estimated using simulated maximum likelihood with 1000 Halton draws. Random parameters were introduced to capture the unobserved heterogeneity associated with the crash data. The random parameters were assumed to be normally distributed. The normal distribution was observed to outperform other distributions like uniform distributions as shown in previous studies (Adanu et al., 2023; Dzinyela et al., 2023; Hojati et al., 2013). The signs on the fixed parameters, either positive or negative, can be used to interpret the hazard function. A parameter estimate with a positive sign indicates an increase in time between a previous not-fatal crash and a subsequent fatal crash or a decrease in the hazard function. Conversely, a negative sign indicates an increase in hazard function and a decrease in the time it takes for a driver to be involved in a fatal crash. However, this is different for parameters that were found to be random.

Unlike fixed parameters, the random parameters do not assume that the effect of these signs are true for all data samples. In this study, the indicator variable for previous license suspension was random in the male recidivist driver model for all age groups. The mean of −0.31 and standard deviation of 0.36, revealed that for about 19.5% of male recidivist drivers under the age of 25 years whose license have been suspended in the past, the duration between a non-fatal crash and a fatal

Table 3
Estimation results of three separate duration models (Weibull) for time until a recidivist male driver is killed in a fatal crash after his previous non-fatal crash.

Variable description	Age less than 25 years old	Age between 25 years and 64 years old	Age greater than 64 years
	Parameter estimate(t-statistics)	Parameter estimate(t-statistics)	Parameter estimate(t-statistics)
Constant	2.678(48.59)	2.956(80.06)	2.889 (46.56)
Number of previous non-fatal crashes: 1 (1 if yes, 0 otherwise)	0.256(6.48)	0.309(14.44)	0.417 (10.21)
Number of previous crashes: more than 4 (1 if yes, 0 otherwise)	−0.982(−4.18)	−0.458(−5.84)	−0.592(−3.14)
Previous speeding violation (1 if yes, 0 otherwise)	−0.252(−6.96)	−0.296(−15.08)	−0.256 (−5.47)
Standard deviation of parameter density function for Previous speeding violation	0.198(7.53)	–-	–-
Previous license suspension (1 if yes, 0 otherwise)	−0.313(−7.56)	−0.427(−18.76)	−0.328(−5.21)
Standard deviation of parameter density function for Previous license suspension	0.359(10.94)	0.456(26.50)	0.222(4.10)
Previous DUI violation (1 if yes, 0 otherwise)		0.06952(2.12)	0.31406 (2.33)
Some other previous violation (1 if yes, 0 otherwise)	−0.289(−7.78)	−0.212(−11.05)	−0.166(−4.25)
No seat belt use at time of fatal crash (1 if yes, 0 otherwise)	0.089(2.52)	–-	–-
Invalid license at time of fatal crash (1 if yes, 0 otherwise)	−0.159(−3.39)	−0.173(−7.03)	−0.207 (−2.52)
Drunk at time of fatal crash	0.082(2.01)	−0.039(−1.73)	–-
Under drug Influence at time of fatal crash	0.068(1.42)	−0.038(−1.46)	–-
Driver race/ethnicity: Caucasian (1 if yes, 0 otherwise)	0.085(2.11)	0.063(1.97)	0.085(1.52)
Driver race/ethnicity: African American –-(1 if yes, 0 otherwise)		0.0662(1.76)	0.156 (1.97)
Scale Parameter for survival function			–-
P	0.813(51.42)	0.764(96.05)	0.700(48.79)
Model statistics			
Log likelihood at convergence	–3232.22763	−10591.51653	−2459.83870
AIC	6492.5	21211.0	4943.7
Number of observations	2306	7747	1962

Table 4
Estimation results of three separate duration models (Weibull) for time until a recidivist female driver is killed in a fatal crash after his previous non-fatal crash.

Variable description	Age less than 25 years old	Age between 25 years and 64 years old	Age greater than 64 years
	Parameter estimate(t-statistics)	Parameter estimate(t-statistics)	Parameter estimate(t-statistics)
Constant	2.973(39.5)	3.044(78.97)	2.959(43.65)
One previous non-fatal crashes indicator (1 if yes, 0 otherwise)	0.136(1.90)	0.350(9.62)	0.439(6.12)
More than four previous non-fatal crashes indicator (1 if yes, 0 otherwise)	–-	−0.446(−2.30)	–-
Previous DUI violation (1 if yes, 0 otherwise)	0.208(1.39)	–-	–-
Previous speeding violation (1 if yes, 0 otherwise)	−0.409(−6.55)	−0.379(−11.14)	−0.211(−2.32)
Standard deviation of parameter density function for **Previous speeding violation**	0.390(7.73)	0.245(8.91)	
Previous license suspension (1 if yes, 0 otherwise)	−0.4556(−5.98)	−0.415(−10.71)	−0.292(−2.32)
Standard deviation of parameter density function for **Previous license suspension**	0.475(7.44)	0.474(14.69)	
Some other previous violation (1 if yes, 0 otherwise)	−0.234(−3.62)	−0.261(−7.74)	−0.110(−1.62)
Invalid license at time of fatal crash (1 if yes, 0 otherwise)	–-	−0.088(0.04589)	−0.272(−1.86)
Over speeding at time of fatal crash (1 if yes, 0 otherwise)	–-	−0.091(−2.47)	–-
Drunk at time of fatal crash	0.088(1.28)	–-	–-
Driver race/ethnicity: Hispanic (1 if yes, 0 otherwise)	−0.2631(−2.92)	–-	–-
Scale Parameter for survival function			
P	0.783(30.42)	0.736(55.3)	0.662(29.44)
Model statistics			
Log likelihood at convergence	−1128.09089	−3288.64596	−834.18207
AIC	2278.2	6599.3	1684.4
Number of observations	806	2443	692

crash is more likely to be shorter. For the remaining 80.5%, the time between a non-fatal crash and a fatal crash is more likely to be longer. Similarly for 17.5% of male recidivist drivers between 25 and 64 years whose license was previously suspended, the duration between a non-fatal crash and a fatal crash is shorter while the duration between non-fatal and fatal crash is longer for the remaining 82.5% of male recidivist drivers between 25 and 64 years with a previously suspended license. Furthermore, the mean of −0.33 and standard deviation of 0.22 revealed that for about 6.7% of male recidivist drivers older than 64 years whose license has been previously suspended, the duration between previous non-fatal crash and a subsequent fatal crash is likely to be shorter. For the rest of the 93.3% of male recidivist drivers older than 64 years whose license have been previously suspended, the time between to fatal crash after a previous non-fatal crash is longer. The variable for previous speeding violation was also random for male recidivist drivers less than 25 years model. With a mean of −0.25 and a standard deviation of 0.20, the duration between previous non-fatal crash and subsequent fatal crash for about 10.6% of male recidivist drivers less than 25 years is likely to be shorter. The duration between non-fatal crash and fatal crash for the remaining 89.45 of male recidivist drivers younger than 25 years are more likely to be longer.

For the female recidivist drivers less than 25 years, the indicator variable for previous speeding violation was also random with the mean of −0.45 and standard deviation of 0.47. This result revealed that for 16.9% of female drivers less than 25 years previously charged with speeding violation, the time between a non-fatal crash and a fatal crash is more likely to be shorter. The time between the previous non-fatal crash and fatal crash for is more likely to be longer for the remaining

83.1% of female drivers less than 25 years previously charged with speeding violation. Similarly, the duration between previous non-fatal crash and a subsequent fatal crash for about 19.2% female recidivist between 25 years and 64 years is likely to be shorter. For the rest of the 80.8% of female recidivist drivers between 25–64 years whose license has been previously suspended, the duration between previous non-fatal crash and subsequent fatal crash is longer. Furthermore, the variable indicator for female recidivist drivers less than 25 years who were previously charged with speeding violation was random with a mean of −0.41 and standard deviation of 0.39. This result revealed that for about 14.7% of female recidivist drivers less than 25 years, the expected time between a non-fatal crash and a fatal crash is shorter and the time is longer for the remaining 85.3% of female drivers in that age group.

The exponent of the coefficient of the variables provides an easy way to interpret the results similar to elasticity (Washington et al., 2020). The exponent of these coefficients reveals the percentage change in the time (i.e. duration between a non-fatal crash and a fatal crash) for a unit change in the continuous variables and a change from zero to one for the indicator predictor variables. Table 3 and 4 presents the coefficients and t-statistics in incident duration for all variables used in the male and female models. The next section discusses the results of the explanatory variables.

7. Discussion

This study has observed differences in the factors that are associated with the duration between a previous non-fatal crash and an eventual fatal crash among male and female drivers across three age categories: less than 25 years, between 25 and 65 years, and more than 65 years. One of the factors that was found to significantly influence the duration is the number of previous non-fatal crashes. In both the male and female driver models, it was observed that the percentage change in duration increased with age if the driver was involved in only one previous non-fatal crash. For instance, it was found that the duration between the two crash events increased by 29.2%, 36.3 %, and 52.7% for male drivers less than 25 years, those between 25 and 65 years, and those more than 65 years, respectively. A similar trend was observed for female drivers, although the duration increased by lower rates (14.6% for younger female drivers and 55.1% for older female drivers), compared to their male counterparts. In the case of female drivers aged between 25 and 65 years, the duration between the two crash events increased by 41.9%. These findings generally show that the likelihood of getting into a fatal crash after one non-fatal crash reduces with age, meaning that younger drivers are more likely than older drivers to be involved in a fatal crash if they were previously involved in a non-fatal crash. Additionally, the results of the study point to a lower duration for younger female drivers compared to their male counterparts. On the other hand, it was found that drivers who got into more than 4 previous non-fatal crashes have reduced duration between their most recent non-fatal crash and an eventual fatal crash, regardless of gender. This finding is intuitive as drivers who are reported to be involved in more crashes may more likely be crash-prone.

The occurrence of multiple crashes involving an individual in a span of 5 years could be indicative of a much more complex underlying behavioral factors or systemic frameworks that failed to reform the driving style and understanding of crash risk factors (Shaw & Sichel, 2013). Drivers who were cited for a previous speeding violation or license suspension had a reduced duration between their previous non-fatal crashes and an eventual fatal crash. The model results show that the duration between the two crash events is slightly shorter if the driver was previously cited for a speeding violation than for drivers who had their licenses suspended. For female drivers, the model results further found that the duration decreased with age. For instance, it was observed that the duration decreased by 33.6% and 31.6% for younger female drivers and middle-aged female drivers, respectively but the duration only decreased by 19.1% for older female drivers. Interestingly, drivers who were previously cited for DUI offense had increased duration between their previous non-fatal crashes and an eventual fatal crash. It was found that the duration increased marginally by 7.2% and by 36.9% for male drivers aged between 25 and 65 years and those more than 65 years, respectively. Younger female drivers who were cited for a DUI offense had the duration between the two crash events increased by 23.1%.

Seatbelts have been shown to save lives, and failure to use them is a citable traffic violation across almost all states in the U.S. (Farooq et al., 2021; Nichols et al.,2014). Drivers who fail to use seatbelts have been shown in previous studies (e.g., Begg & Langley,2000; Ivers et al., 2009) to be inclined to take more risks. The model results in this study found that recidivist younger male drivers have 9.3% increased duration between their previous crashes and a current fatal crash if were unbelted in the fatal crash. Similarly, the duration between the two crashes increased by 8.6% and 7% for younger male drivers who were drunk and those that were under the influence of a drug, respectively, at the time of the fatal crash. The duration also increased by 9.2 % for younger female drivers who were drunk at the time of the fatal crash. While an increased duration is observed for younger male drivers who were under the influence alcohol or drug at the time of the fatal crash, the duration is marginally decreased for middle-aged male drivers were under the influence of alcohol or drug at the time of the crash. Also, driving with no or an invalid license is a traffic violation in most states. Studies have shown that drivers who do not have a valid license at the time of a crash are more likely to be severely injured (Feng et al., 2016; Retting et al., 1999).

This study found that recidivist drivers who did not have a valid driver's license at the time of the fatal crash generally had shorter durations between their most recent non-fatal crash and the current fatal crash, regardless of age or gender. Middle-aged female drivers who were speeding at the time of their fatal crashes were found to have 8.7% reduction in the duration between their most recent non-fatal crash and the current fatal crash. This finding may be due to the increasing number of working females who may be engaged in multiple activities that require them to achieve a lot in a shorter period, hence contributing to speeding especially among middle-aged women. This age group is also mostly made up of working mothers who have increased travel responsibilities. Indeed, some previous studies have argued that generational changes in attitudes towards risk has contributed to more female drivers engaging in aggressive and risky driving behaviors (Kostyniuk et al., 1998). With respect to the driver race, the duration between crashes for younger Hispanic female drivers was found to be shorter by 23.1% whereas the duration for Caucasian drivers increased by an average of about 8% across all ages. Middle-aged and older African American male drivers have their duration between crashes increased by 6.8% and 16.9%, respectively.

The results of this study provide safety enhancement related policy implications and actionable insights. Notably, the findings underscore the importance of age as a determining factor in the duration between a non-fatal crash and a fatal crash occurrence, with younger drivers exhibiting a higher likelihood of involved with a fatal crash after a non-fatal incident. Policymakers should consider targeted interventions and educational programs geared towards younger drivers to mitigate this heightened risk. The influence of recidivism in non-fatal crashes is a critical factor, particularly for drivers with a history of more than four previous incidents. These individuals exhibit a reduced duration between their most recent non-fatal crash and a subsequent fatal crash, suggesting an increased crash-prone related behavior. This necessitates a focus on intervention strategies, possibly including driver education and rehabilitation programs, to address underlying behavioral issues contributing to repeat incidents. Citations for speeding violations or license suspensions are indicative of shorter durations between non-fatal and fatal crashes, emphasizing the need for rigorous enforcement and deterrent measures. Conversely, drivers cited for DUI offenses display an increased duration, suggesting a potential need for targeted rehabilitation programs and strict enforcement for this specific group. The study reinforces the importance of seatbelt usage, particularly for recidivist younger male drivers, as their failure to use seatbelts is associated with an increased duration between crashes leading to fatalities. This underscores the effectiveness of seatbelt enforcement in preventing severe consequences. Lastly, disparities based on race and ethnicity are evident, with implications for policy considerations for underserved communities. Efforts should be directed towards understanding and addressing these disparities, particularly focusing on factors that contribute to shorter durations for certain groups, such as younger Hispanic female drivers.

8. Summary and conclusions

Recidivist drivers pose traffic risks to themselves and other road users. Addressing the issue requires a multifaceted approach. This study presented an analysis of the factors that are associated with the duration between a previous non-fatal crash and a subsequent fatal crash. Data for the study was obtained from the FARS database for the period covering 2015–2019. In all, a total of 15,958 fatal crash observations were available for analysis after cleaning the data. To better understand how these factors that account for differences in the duration between crashes affect different driver groups, the data was divided by gender and into three age categories. Preliminary analysis revealed that on average, drivers got involved in a fatal crash within a year and a half after they get involved a non-fatal crash, with only an average of one month difference between male and female drivers. More than 75% of the crash observations involved drivers that were involved in only one previous crash before getting into a fatal crash. Random parameters hazard-based duration models were then developed for the various age and gender categories of drivers.

The model results revealed differences and similarities in the factors that influence the time interval between the previous crash and the subsequent fatal crash across the driver groups considered. Factors such as number of previous crashes, previous traffic violations, primary contributing factors and some driver demographic characteristics were found to significantly be associated with the duration between the two crash events. For example, it was found that the duration increased with driver age for drivers who were involved in only one previous crash and the duration was shorter for those that were previously involved in multiple crashes. Previous DUI violations, license suspensions, and previous speeding violations were found to be associated with shorter durations, at varying degrees depending on the driver age and gender. The duration was also observed to be longer if the fatal crash involved alcohol or drug use among younger drivers but shorter among middle-aged male drivers. Middle-aged women who got killed in a speeding-related crash were associated with shorter durations between their previous crashes and the eventual fatal crash.

These findings reveal interesting dynamics that may be linked to recidivist tendencies among some drivers involved in fatal crashes. The association between multiple previous crashes and an eventual fatal crash may be due to an increased risk-taking behavior among these driver populations, The ability to identify these population groups provide an opportunity to understand the mechanisms that may be responsible for the phenomenon. Considering that human factors have been attributed the most for risky driving and crashes, it is important to assess programs and crash countermeasures that seek to reform driver behaviors. Deficiency in such programs and countermeasures may perhaps be accountable for breeding recidivism among drivers. As such, there is the need to adopt and adapt best practices from fields such as social work and criminal justice for implementation to minimize recidivism among drivers. Additionally, multidisciplinary programs should be developed to ensure that drivers who get into crashes are encouraged to not return to driving immediately. Such drivers should undergo therapy and be re-tested to qualify to drive again. Considering the identified factors influencing the duration between these events, policy interventions should focus on age-specific and gender-specific strategies. For instance, driver license point systems may be customized to account for the gender and age of the driver, taking into consideration the previous crash records of the driver. Additionally, traffic schools for ticketed drivers should be reformed to ensure that not only are tickets dismissed, but driving behaviors are significantly transformed. These reforms could include more personalized and intensive training sessions, incorporating psychological assessments and targeted educational modules. Enhanced penalties, such as increased fines and longer license suspensions, can act as a deterrent. However, punishment alone is insufficient. Rehabilitation programs that focus on behavior modification are crucial. Leveraging technology, such as telematics devices that monitor driving habits, can provide real-time feedback and accountability for offenders. By combining stricter enforcement with educational and rehabilitative efforts, it is possible to reduce recidivism among repeat traffic law violators and improve overall road safety. Furthermore, implementing regular follow-up evaluations for high-risk drivers can help reinforce safe driving practices and reduce the likelihood of repeat offenses. Moreover, the study advocates for the adoption of best practices from fields such as social work and criminal justice to design effective programs that reform driver behavior, ultimately reducing recidivism.

Despite the valuable insights from this study, it is essential to acknowledge its limitations. The generalizability of the findings may be constrained by geographical and cultural variations within the dataset. Furthermore, the current study focused on individual-level factors, potentially overlooking broader systemic failure. The Weibull distribution, which is a more generalized form of the exponential distribution provides a more flexible means of capturing duration dependence. However, the Weibull has its limitations as it requires the hazard to be monotonic over time. In many applications, a nonmonotonic hazard is theoretically justified (Washington et al.,2020). Future research should consider a more extensive range of contextual variables such as sub-regional penalty and fine structure for traffic law violations and employ longitudinal designs to enhance the robustness and applicability of the finding. Including such variables will help to understand the effectiveness of such interventions in moderating the effects of other variables in determining the duration between crashes or the frequency of repeated offenses across the entire nation.

CRediT authorship contribution statement

Richard Dzinyela: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Emmanuel Kofi Adanu:** Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization. **Hardik Gupta:** Writing – review & editing, Writing – original draft. **Pranik Koirala:** Writing – review & editing, Writing – original draft. **Nawaf Alnawmasi:** Writing – review & editing, Writing – original draft, Visualization, Methodology. **Subasish Das:** Writing – review & editing, Writing – original draft. **Dominique Lord:** Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

Adanu, E. K., & Jones, S. (2017). Effects of human-centered factors on crash injury severities. *Journal of advanced transportation, 2017*.

Adanu, E.K., Penmetsa, P., Wood, D., Jones, S.L., 2019. Incorporating systems thinking approach in a multilevel framework for human-centered crash analysis. Transportation Research Interdisciplinary Perspectives 2, 100031.
Adanu, E.K., Dzinyela, R., Agyemang, W., 2023. A comprehensive study of child pedestrian crash outcomes in Ghana. Accid. Anal. Prev. 189, 107146.
Ali, Y., Haque, M.M., Zheng, Z., Washington, S., Yildirimoglu, M., 2019. A hazard-based duration model to quantify the impact of connected driving environment on safety during mandatory lane-changing. Transportation Research Part c: Emerging Technologies 106, 113–131.
Alnawmasi, N., Mannering, F., 2019. A statistical assessment of temporal instability in the factors determining motorcyclist injury severities. Analytic Methods in Accident Research 22, 100090.
Alzaffin, K., Kaye, S.A., Watson, A., Haque, M.M., 2023. Modelling the continuum of serious traffic injuries in police-hospital linked data by applying the random parameters hazard-based duration model. Analytic Methods in Accident Research 40, 100291.
Anastasopoulos, P. C., Haddock, J. E., Karlaftis, M. G., & Mannering, F. L. (2012). An analysis of urban travel times: a random parameters hazard-based approach. In *Proceedings of the 91st TRB Annual Meeting Compendium of Papers*.
Ayuso, M., Guillén, M., Alcañiz, M., 2010. The impact of traffic violations on the estimated cost of traffic accidents with victims. Accid. Anal. Prev. 42 (2), 709–717.
Balusu, S.K., Mannering, F., Pinjari, A., 2020. Hazard-based duration analysis of the time between motorcyclists' initial training and their first crash. Analytic Methods in Accident Research 28, 100143.
Begg, D.J., Langley, J.D., 2000. Seat-belt use and related behaviors among young adults. J. Saf. Res. 31 (4), 211–220.
Bhat, C.R., 1996. A generalized multiple durations proportional hazard model with an application to activity behavior during the evening work-to-home commute. Transp. Res. B Methodol. 30 (6), 465–480.
Bogstrand, S.T., Larsson, M., Holtan, A., Staff, T., Vindenes, V., Gjerde, H., 2015. Associations between driving under the influence of alcohol or drugs, speeding and seatbelt use among fatally injured car drivers in Norway. Accid. Anal. Prev. 78, 14–19.
Bucsuházy, K., Matuchová, E., Zůvala, R., Moravcová, P., Kostíková, M., Mikulec, R., 2020. Human factors contributing to the road traffic accident occurrence. Transp. Res. Procedia 45, 555–561.
C'de Baca, J., Miller, W. R., & Lapham, S. (2001). A multiple risk factor approach for predicting DWI recidivism. *Journal of substance abuse treatment*, *21*(4), 207-215.
C'de Baca, J., Lapham, S. C., Liang, H. C., & Skipper, B. J. (2001). Victim impact panels: do they impact drunk drivers? A follow-up of female and male, first-time and repeat offenders. *Journal of Studies on Alcohol*, *62*(5), 615-620.
Chen, T.Y., Jou, R.C., 2018. Estimating factors of individual and regional characteristics affecting the drink driving recidivism. Accid. Anal. Prev. 119, 16–22.
Das, S., Sun, X., Wang, F., Leboeuf, C., 2015. Estimating likelihood of future crashes for crash-prone drivers. Journal of Traffic and Transportation Engineering (english Edition) 2 (3), 145–157.
Dingus, T.A., Guo, F., Lee, S., Antin, J.F., Perez, M., Buchanan-King, M., Hankey, J., 2016. Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proc. Natl. Acad. Sci. 113 (10), 2636–2641.
Dzinyela, R., Adanu, E.K., Lord, D., Islam, S., 2023. Analysis of factors that influence injury severity of single and multivehicle crashes involving at-fault older drivers: A random parameters logit with heterogeneity in means and variances approach. Transportation Research Interdisciplinary Perspectives 22, 100974.
Dzinyela, R., Alnawmasi, N., Adanu, E.K., Dadashova, B., Lord, D., Mannering, F., 2024. A multi-year statistical analysis of driver injury severities in single-vehicle freeway crashes with and without airbags deployed. Analytic Methods in Accident Research 100317.
Factor, R., 2014. The effect of traffic tickets on road traffic crashes. Accid. Anal. Prev. 64, 86–91.
Farooq, M.U., Ahmed, A., Saeed, T.U., 2021. A statistical analysis of the correlates of compliance and defiance of seatbelt use. Transport. Res. F: Traffic Psychol. Behav. 77, 117–128.
Fell, J.C., 1993. Repeated dwi offenders: their involvement in fatal crashes. In: Proceedings International Council on Alcohol, Drugs and Traffic Safety Conference, Vol. 1993. International Council on Alcohol, Drugs and Traffic Safety, pp. 1044–1049.
Fell, J., 1995. Repeat DWI offenders in the United States. Traffic tech. Technology transfer series 85.
Feng, S., Li, Z., Ci, Y., Zhang, G., 2016. Risk factors affecting fatal bus accident severity: Their impact on different types of bus drivers. Accid. Anal. Prev. 86, 29–39.
Ferrante, A.M., Rosman, D.L., Marom, Y., 2001. Novice drink drivers, recidivism and crash involvement. Accid. Anal. Prev. 33 (2), 221–227.
Garib, A., Radwan, A.E., Al-Deek, H., 1997. Estimating magnitude and duration of incident delays. J. Transp. Eng. 123 (6), 459–466.
Gould, L.A., Gould, K.H., 1992. First-time and multiple-DWI offenders: a comparison of criminal history records and BAC levels. J Crim Just 20 (6), 527–539.
Greene, W.H., 2012. Econometric Analysis. Prentice Hall, Boston.
W.H. Greene LIMDEP version 9.0/NLOGIT version 4.0 Econometric Modeling Guide, Econometric Software 2007 Plainview, New York.
Hiekmann, J.M., 2012. Real-World Implications of the Safe System Approach. Roadside Safety Design and Devices 59.
Hojati, A.T., Ferreira, L., Washington, S., Charles, P., 2013. Hazard based models for freeway traffic incident duration. Accid. Anal. Prev. 52, 171–181.
Islam, N., et al., 2022. Evaluating the impact of freeway service patrol on incident clearance times: a spatial transferability test. J. Adv. Transp. 2022.
Islam, N., Adanu, E.K., Hainen, A.M., Burdette, S., Smith, R., Jones, S., 2021. A comparative analysis of freeway crash incident clearance time using random parameter and latent class hazard-based duration model. Accid. Anal. Prev. 160, 106303.
Ivers, R., Senserrick, T., Boufous, S., Stevenson, M., Chen, H.Y., Woodward, M., Norton, R., 2009. Novice drivers' risky driving behavior, risk perception, and crash risk: findings from the DRIVE study. Am. J. Public Health 99 (9), 1638–1644.
Khan, M.N., Das, S., 2024. Advancing traffic safety through the safe system approach: A systematic review. Accid. Anal. Prev. 199, 107518.
Kostyniuk, L.P., Molnar, L.J., Eby, D.W., 1998, July. Are women taking more risks while driving? A look at Michigan drivers. In: Women's Travel Issues Second National ConferenceDrachman. Institute of the University of Arizona; Morgan State University; Federal Highway Administration.
Lawpoolsri, S., Li, J., Braver, E.R., 2007. Do speeding tickets reduce the likelihood of receiving subsequent speeding tickets? A longitudinal study of speeding violators in Maryland. Traffic Inj. Prev. 8 (1), 26–34.
Lijarcio, I., Llamazares, F.J., Valle, E., Montoro, L., Useche, S.A., 2022. Assessing risk perception over recidivist traffic offenders from a multi-group approach: how gendered could it be? European Journal of Psychology Applied to Legal Context 14 (1), 33–41.
Lord, D., Qin, X., Geedipally, S.R., 2021. Highway safety analytics and modeling. Elsevier.
Manno, M., Maranda, L., Rook, A., Hirschfeld, R., Hirsh, M., 2012. The reality of teenage driving: the results of a driving educational experience for teens in the juvenile court system. J. Trauma Acute Care Surg. 73 (4), S267–S272.
McKnight, A.J., Tippetts, A.S., 1997. Accident prevention versus recidivism prevention courses for repeat traffic offenders. Accid. Anal. Prev. 29 (1), 25–31.
McMillen, D.L., Adams, M.S., Wells-Parker, E., Pang, M.G., Anderson, B.J., 1992. Personality traits and behaviors of alcohol-impaired drivers: A comparison of first and multiple offenders. Addict. Behav. 17 (5), 407–414.
Møller, M., Haustein, S., Prato, C.G., 2015. Profiling drunk driving recidivists in Denmark. Accid. Anal. Prev. 83, 125–131.
Nam, D., Mannering, F., 2000. An exploratory hazard-based analysis of highway incident duration. Transp. Res. A: Policy Pract. 34 (2), 85–102.
Nichols, J.L., Tippetts, A.S., Fell, J.C., Eichelberger, A.H., Haseltine, P.W., 2014. The effects of primary enforcement laws and fine levels on seat belt usage in the United States. Traffic Inj. Prev. 15 (6), 640–644.
Padilla, J.L., Doncel, P., Gugliotta, A., Castro, C., 2018. Which drivers are at risk? Factors that determine the profile of the reoffender driver. Accid. Anal. Prev. 119, 237–247.
Pang, J., Krathaus, A., Benedyk, I., Ahmed, S.S., Anastasopoulos, P.C., 2022. A temporal instability analysis of environmental factors affecting accident occurrences during snow events: The random parameters hazard-based duration model with means and variances heterogeneity. Analytic Methods in Accident Research 34, 100215.
Petridou, E., Moustaki, M., 2000. Human factors in the causation of road traffic crashes. Eur. J. Epidemiol. 16, 819–826.
Reason, J., Manstead, A., Stradling, S., Baxter, J., Campbell, K., 1990. Errors and violations on the roads: a real distinction? Ergonomics 33 (10–11), 1315–1332.
Retting, R.A., Ulmer, R.G., Williams, A.F., 1999. Prevalence and characteristics of red light running crashes in the United States. Accid. Anal. Prev. 31 (6), 687–694.
Roberts, P., Woolley, J., Doeke, S., 2014. Providing for Road User Error in the Safe System, No. AP-R460/14. Austroads, Sydney, New South Wales, Australia. Report AP-R460/14,.
Ryan, G.A., Ferrante, A.M., Loh, N., Cercarelli, L.R., 1996. Repeat Drink Driving Offenders in Western Australia, 1984–1994. Federal Office of Road Safety, CR, p. 168.
Schmitz, A.R., Goldim, J.R., Guimarães, L.S., Lopes, F.M., Kessler, F., Sousa, T., Pechansky, F., 2014. Factors associated with recurrence of alcohol-related traffic violations in southern Brazil. Brazilian Journal of Psychiatry 36, 199–205.
Shams, M., Rahimi-Movaghar, V., 2009. Risky driving behaviors in Tehran. Iran*Traffic Injury Prevention* 10 (1), 91–94.
Shaw, L., Sichel, H.S., 2013. Accident proneness: Research in the occurrence, causation, and prevention of road accidents. Elsevier.
Simpson, H. M., & Mayhew, D. R. (1991). The hard core drinking driver.
Sohrabi, S., Lord, D., Dadashova, B., Mannering, F., 2024. Assessing the collective safety of automated vehicle groups: A duration modeling approach of accumulated distances between crashes. Accid. Anal. Prev. 198, 107454.
Summala, H., Rajalin, S., Radun, I., 2014. Risky driving and recorded driving offences: A 24-year follow-up study. Accid. Anal. Prev. 73, 27–33.
Tao, D., Zhang, R., Qu, X., 2017. The role of personality traits and driving experience in self-reported risky driving behaviors and accident risk among Chinese drivers. Accid. Anal. Prev. 99, 228–235.
Tassoni, G., Cippitelli, M., Mirtella, D., Froldi, R., Ottaviani, G., Zampi, M., Cingolani, M., 2016. Driving under the effect of drugs: Hair analysis in order to evaluate recidivism. Forensic Sci. Int. 267, 125–128.
Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., ... & Castellan, N. J. (1979). Tri-level study of the causes of traffic accidents: Final report. Volume I: Casual factor tabulations and assessments.
Tillmann, W.A., Hobbs, G.E., 1949. The accident-prone automobile driver: a study of the psychiatric and social background. Am. J. Psychiatry 106 (5), 321–331.
Washington, S., Karlaftis, M.G., Mannering, F., Anastasopoulos, P., 2020. Statistical and econometric methods for transportation data analysis. CRC Press.
Watson, A., Freeman, J., Imberger, K., Filtness, A.J., Wilson, H., Healy, D., Cavallo, A., 2017. The effects of licence disqualification on drink-drivers: Is it the same for everyone? Accid. Anal. Prev. 107, 40–47.

Watson, B., Watson, A., Siskind, V., Fleiter, J., Soole, D., 2015. Profiling high-range speeding offenders: Investigating criminal history, personal characteristics, traffic offences, and crash history. Accid. Anal. Prev. 74, 87–96.
Yu, J., & Williford, W. R. (1995). Drunk-driving recidivism: predicting factors from arrest context and case disposition. *Journal of Studies on Alcohol*, *56*(1), 60-66.Kostyniuk, L. P., Molnar, L. J., & Eby, D. W. (1996). Are women taking more risks while driving? Proceedings from the Second National Conference on Women's Travel Issues. Federal Highway Administration, Office of Highway Information Management, Baltimore.
Zhao, N., Reimer, B., Mehler, B., D'Ambrosio, L.A., Coughlin, J.F., 2013. Self-reported and observed risky driving behaviors among frequent and infrequent cell phone users. Accid. Anal. Prev. 61, 71–77.