Logistic Regression Modelling of Road Traffic Accident Severity: A Study on Driver Characteristics in Zimbabwe

Ever Moyo[1*], Kaitano Dzinavatonga[2], Zvikomborero Lesley Hakunavanhu[3]

[1]Mathematics and Statistics Lecturer, Department of Mathematical Sciences

[2]Physics Lecturer, Department of Physics

[3]Quality Assurance Coordinator, Department of Quality Assurance

[1,3]Zimbabwe Open University, Masvingo Region

[2]Zimbabwe Open University, Harare Region, Zimbabwe

***Corresponding Author**

DOI: https://dx.doi.org/10.47772/IJRISS.2025.91100254

Received: 21 November 2025; Accepted: 28 November 2025; Published: 06 December 2025

ABSTRACT

Road traffic accidents (RTAs) remain a major public health concern in Zimbabwe, yet little empirical work has examined the combined influence of driver behavior, demographic characteristics and environmental factors on accident severity. This study applied binary logistic regression analysis to a dataset of 500 accident-involved drivers to identify the key predictors of severe accidents. Frequency distributions summarised the characteristics of drivers, environmental conditions and vehicle status, while logistic regression quantified their influence on accident severity. The results showed that six key predictors significantly increased the likelihood of severe accidents: low driving experience, alcohol use, fatigue, mobile phone use, over speeding and wet road conditions. Over speeding emerged as the strongest predictor, with drivers who overspeed being four times more likely to be involved in a severe accident. Although age, gender, time of accident and vehicle condition were not statistically significant, they exhibited expected directional effects. The full model (M_1) significantly improved prediction compared to the null model (M0) with $\Delta\chi^2 = 77.3$ demonstrating that the included predictors collectively enhance the model's explanatory power with respect to predicting accident severity. It demonstrated an acceptable predictive accuracy with an AUC = 0.720, indicating its effectiveness in distinguishing between severe and non-severe accidents. The findings emphasize that human behavior remains the most critical determinant of accident severity in Zimbabwe, with implications for targeted interventions. The study findings highlight the need for evidence-based interventions focused on speed control, anti-drunk driving enforcement, fatigue management, mobile phone usage laws, targeted road safety campaigns and improved road infrastructure especially during wet conditions. Training and awareness programs targeting inexperienced drivers could reduce severity outcomes. The study contributes valuable insights toward improving road safety strategies in Zimbabwe and similar contexts.

Keywords: Binary Logistic Regression Analysis, Accident Severity, Driver Characteristics, Predictors, Road Safety.

INTRODUCTION

Background

Road traffic accidents (RTAs) are among the leading causes of mortality and morbidity worldwide, particularly in developing countries. Road accidents are a pervasive global issue with profound consequences for individuals, communities, and economies (Deepthi et al.2023). Zimbabwe has comprehensive road network linking the different parts of the country and providing access to neighbouring countries for imports and exports. The country is experiencing an increase in motorisation while roads have deteriorated resulting in increased road accidents (Muvuringi, 2012). Nowadays, the problem of road accident rates is one of the most important health and social policy issues concerning the countries in all continents. Each year, nearly 1.3 million people worldwide lose their life on roads, and 20–50 million sustain severe injuries, the majority of which require long-term treatment (Goniewicz et al.,2016). Goniewicz et al.,2016, discovered that the causes of road accidents include lack of control and enforcement concerning implementation of traffic regulation (primarily driving at excessive speed, driving under the influence of alcohol and not respecting the rights of other road users (mainly pedestrians and cyclists), lack of appropriate infrastructure and unroadworthy vehicles. They suggested the strategies and programmes for improving road traffic such as reducing the risk of exposure to an accident, prevention of accidents, reduction in bodily injuries sustained in accidents, and reduction of the effects of injuries by improvement of post-accident medical care.

Foya, (2019) established that the major causes of accidents in Zimbabwe were human error which included unlicensed driving, texting, alcohol consumption and driving while driving., failure to observe road and that most of the vehicles were not mechanically safe to be in the roads. Foya, (2019) concluded that all these factors have caused numerous deaths in the roads and it has become a major cost to insurance companies. State of the various roads is a major challenge to motorists, and some roads are full of potholes which makes it dangers for motorists to drive through.

Chibaro et al., (2024) assessed the effects of human factors on road traffic safety in minimizing road traffic accidents in Harare Metropolitan, Zimbabwe. They tested human factors such as improper age of drivers who tend to overtake and overspeed recklessly, over speeding, alcohol drinking, corruption and failure to maintain vehicles as a moderating variable. They identified that excessive speeding and drunk driving while driving were the worst human behaviour causing accidents in Zimbabwe.They concluded that it is necessary to enforce the speed limit and deploy speed cameras at significant intersections and problem areas. Munuhwa et al. (2020) indicated that in Botswana, the RTAs are caused by road users through speeding, unlicensed driving, using cell phones whilst driving, alcohol and drug abuse, bad state of mind and healthy, non-use of safety belts and deliberate failure to observe road regulations amongst others. Their findings also indicated that mechanically faulty vehicles, unmaintained vehicles, old vehicles, and tyre blowouts are vehicle related factors causing RTAs. Road system conditions involve potholes, stray livestock and road design attributes amongst others. They recommended educating the public on safe driving habits, punitive policies on road users breaking road traffic laws and regulations, stringent measures against livestock owners who leave stock straying in highways and public roads and regular road maintenance and vehicle maintenance to reduce RTAs. Across the world, traffic accidents cause major health problems and are of concern to health institutions; nearly 1.35 million people are killed or disabled in traffic accidents every year. This issue is growing; by 2030, road traffic injuries will be the seventh leading cause of death globally (Ahmed et al.2023).

Road traffic accidents hamper economic growth as they gobble huge financial resources which government can channel to more urgent developmental programmes. Accident severity can be categorized into three levels, slight, serious and fatal, depending on the outcome for the driver or passengers. Understanding the determinants of these severity levels is essential for policymakers, law enforcement agencies and public health authorities. Road traffic accidents (RTAs) pose significant public health and economic challenges in Zimbabwe. Understanding factors that influence the severity of accidents can help develop targeted safety interventions. This study focuses on driver characteristics as predictors of accident severity using logistic regression modelling.

Research Problem

Despite the availability of accident statistics and the prevalence of these accidents, there is limited understanding of the specific driver characteristics that influence accident severity in Zimbabwe. There is also limited quantitative research exploring how driver characteristics influence the accident severity. This gap hinders the development of targeted interventions and policies aimed at reducing the severity and incidence of road accidents. This study aims to model the relationship between driver characteristics and the severity of road traffic accidents in Zimbabwe through logistic regression modelling, to inform evidence-based strategies for improving road safety in Zimbabwe.

Research Objectives

1. To identify key driver characteristics influencing road accident severity in Zimbabwe.
2. To develop and validate a logistic regression model predicting accident severity based on driver attributes.
3. To provide recommendations for reducing the occurrence of severe road accidents.

Research Questions

1. What driver characteristics significantly influence the severity of road traffic accidents?
2. How well can logistic regression models predict accident severity outcomes?
3. What policy interventions can be derived from the model results?

Significance of the Study

This study is significant because it sheds light on the characteristics of drivers that influence the severity of accidents. This information will help road safety authorities and lawmakers create focused interventions and safety measures. The results will assist in prioritizing resource allocation for road safety programs targeted at vulnerable groups more effectively. The study enhances data-driven decision making for accident prevention by using logistic regression modelling. The results can also improve driver education programs by addressing risky behaviours that are unique to the local population. Additionally, it contributes to academic literature by applying statistical techniques in the context of road safety in Zimbabwe. Reducing accident severity has socio-economic benefits through fewer injuries, fatalities and related costs. Furthermore, the study advances regional understanding of road safety issues in Zimbabwe and comparable settings.

LITERATURE REVIEW

Logistic regression was applied to accident-related data collected from traffic police records to examine the contribution of several variables to accident severity (Al-Ghamdi, 2002). Michalaki et al., (2015) applied a generalized ordered logistic regression model to identify the factors affecting the severity of hard shoulder (HS) and main carriageway (MC) accidents on motorways. They discovered that driver fatigue is one of the factors increasing the severity of the accidents.

Nowadays, road safety is an issue particularly relevant because of the increasing occurrence of road accidents, and it is very important to analyse accident severity and the factors influencing it (Eboli et al., (2020). Human behaviour is a dominant factor in road accidents, contributing to more than 70% of such incidents (McCarty and Kim ,2024).Chen et al,(2021) investigated the factors affecting the accident severity of drivers with different driving experience in Shaanxi, China and discovered that novice drivers younger than 30 or older than 55 are prone to suffer fatal accident, but for experienced drivers, the risk of fatal accident decreases. Paleti et al., (2010) and Adavikottu and Velaga, (2021) found also that novice drivers under the influence of alcohol and driving on roads with high-speed limits triggered aggressive driving behaviour leading to severe injuries. Haerani et al. (2019) suggested that age was a moderate variable in the relationship between personality, driving behaviour and driving outcomes in the city of Makassar, the capital of the South Sulawesi province in Indonesia.Muvuringi, (2012) discovered that Zimbabwe's key risk factors that contribute to RTAs include reckless driving, violation of traffic laws, damaged vehicles, and bad roads. Chibaro et al., (2024) suggested that excessive speeding and drunk driving while driving have been identified as the worst human behaviour causing accidents in Zimbabwe. Millicent et al., (2016) findings reflected that environmental, personal and mechanical factors were the major driving factors of RTAs in Zimbabwe.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

METHODOLOGY

Research Design

A quantitative, analytical design was employed using secondary accident data on a sample of 500 drivers from 2020 to 2024, to model the relationship between accident severity (dependent variable) and driver characteristics (independent variables) using Logistic Regression Modelling in Zimbabwe. Data analysis was conducted using JASP version 0.95.4.0. which supports advanced logistic regression procedures, diagnostics and validation techniques.

Study Area

The study focused on road traffic accidents occurring across the regions of Zimbabwe. Zimbabwe is a developing country with diverse driving environments such as urban, peri-urban and rural roads. It is characterised by variability in road conditions, driver behaviours and enforcement capacity. This made it a suitable setting for assessing the determinants of accident severity.

Data Source

Secondary accident data were obtained from the official Zimbabwe Republic Police (ZRP) Accident Reports and the Traffic Safety Council of Zimbabwe (TSCZ) databases from 2020-2024 for 500 drivers. These records included detailed information on accident type, driver demographics, vehicle factors, environmental conditions and severity outcome.

Data Collection and Preparation

The data on road traffic accidents were gathered which included variables such as weather conditions, time of day, vehicle condition, road condition and driver characteristics. The target variable, accident severity was categorised as a binary outcome (severe and non-severe).

Data Cleaning and Coding

A data cleaning procedure was done in JASP.0 /1 coding was used in this study because it provides a clear, binary representation of the categorical variables used, simplifying the modelling process. It also allows the logistic regression to model the log-odds of the outcome as a linear function of predictors, making it easy to interpret the results (Dominguez-Almendros et al., 2011). The missing data was promptly inserted. The continuous variables were automatically dummy coded in JASP. The following coding scheme was employed:

Coding Schemes

Logistic regression in JASP requires correctly coded variables. The coefficient associated with1-coded variable indicates the change in log-odds when moving from the reference category (0) to the target category (1). The following conventions were done to both the dependent variable and the categorical predictor variables.

Dependent Variable (0utcome): Accident Severity

This had a binary outcome with 1 = severe accident and 0 = non-severe accident. This coding is suitable for logistic regression estimates of accident severity because it models the probability of a binary outcome.

Categorical Predictor Variables

The predictors were binary coded and those with more than two classes were automatically dummy coded in JASP.The categorical predictors were coded numerically (dummy coded) in JASP as follows:

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Table 1: Coding Scheme for the categorical Predictors

Variable	Category	code	Rationale
Age of Driver	< 25years	0	Reference
	25-45years	1	Risk
	>45years	2	Safer
Gender	Female	0	Reference
	Male	1	Risk
Driving experience	<2years	0	Reference
	2-5years	1	Risk
	>5years	2	Safer
Alcohol Use	No	0	Reference
	Yes	1	Higher Risk
Fatigue	No	0	Reference
	Yes	1	Higher Risk
Mobile Phone Use	No	0	Reference
	Yes	1	Higher Risk
Over speeding	No	0	Reference
	Yes	1	Higher Risk
Time of Accident	Day	0	Reference
	Night	1	Risk
Vehicle Condition	Good	0	Reference
	Poor	1	Higher Risk
Road Condition	Dry	0	Reference
	Wet	1	Higher Risk

Data Analysis Procedures

Exploratory Data Analysis

Frequency tables were generated for all the categorical variables. The missing values were checked and manually inserted.

Checking Logistic Regression Model Assumptions

Correlation and Multicollinearity Check

Multicollinearity is a situation in regression analysis where two or more predictor variables are highly correlated, meaning they provide overlapping or redundant information about the outcome variable. Multicollinearity affects interpretation not prediction. It does not influence the overall model fit but it affects the stability, interpretation and precision of regression coefficients. Proper diagnostic using Tolerance and VIF should always be performed before interpreting the logistic regression results. When predictors overlap, it becomes difficult to determine the unique effect of each variable. In this study multicollinearity diagnostics were conducted using Tolerance and the Variance Inflation Factor (VIF) in JASP to assess whether the predictor variables in the accident severity logistic regression model were excessively correlated. Tolerance Values < 0.10 indicate severe multicollinearity and Tolerance Values < 0.20 indicate a potential concern. Hence Acceptable ranges for Tolerance Values is any value greater than 0.20. Variance Inflation Factor (VIF) > 5 suggests moderate multicollinearity and VIF > 10 indicates serious multicollinearity. Hence Acceptance ranges for VIF is any value less than 5.

Model Specification

A binary logistic regression was employed to model the probability of severe accidents based on vehicle condition, environmental conditions and driver characteristics.

The Logistic Regression Model

Logistic regression model is a statistical model in which an evaluation is made of the relationship between: a dependent qualitative, dichotomic variable (binary or binomial logistic regression) or variable with more than two values (multinomial logistic regression) and one or more independent explanatory variables or covariates, whether qualitative or quantitative (Dominguez-Almendros et al., 2011).In this study**,** Logistic regression is a statistical modelling method used to predict the probability of a binary outcome variable, the accident severity. Binary Logistic Regression model was used to estimate the probability of an accident occurring and to understand how the predictor variables (driver age, vehicle condition, gender, driving experience, alcohol consumption, fatigue, mobile phone use when driving, over speeding, time of accident, road condition) influence the severity of the road traffic accidents. It enabled us to classify the results into significant factors and non-significant factors and to compute the Odds Ratio (OR) associated to the covariate/predictor together with the associated 95% confidence interval (CI) for interpretation (Dominguez-Almendros et al., 2011). The probability and odds aid in decision making and hypothesis testing.

The Logistic Function

The probability of the event is:

$$P(Y = (1|X) = p = \frac{1}{1+e^{-(\beta_0+\beta_1X_1+\beta_2X_2+\ldots+\beta_kX_k)}} + \epsilon$$

The Logit Transformation

The probabilities are bounded between 0 and 1, hence the binary logistic regression model predicts the log-odds of the outcome:

$$logit(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1X_1 + \beta_2X_2 + \ldots + \beta_kX_k + \epsilon$$

where

p= probability of a severe accident,

X_i= driver characteristics,$i = 1,2,3 \ldots k$.

β_i= regression coefficients, $i = 1,2,3 \ldots k$.

ϵ =the residuals

Model Fitting

The Maximum Likelihood Estimation in JASP was used to estimate the model parameters to obtain the coefficients, standard errors, Odds-Ratios, Confidence Interval and p-values that explains the relationships

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

between the variables in the data. Model fitting was assessed using the Pseudo R^2 Indices (McFadden R^2, Nagelkerke R^2, Tjur R^2, Cox & Snell R^2) in which higher values indicate a better fit, AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) in which lower values indicate a better fit and The Chi-Square Distribution, in which a low p-value indicates that the model is a good fit.

Model Evaluation and Validation

The predictive accuracy of the model was evaluated using the AUC (Area Under Curve). It helped to measure the model's ability to distinguish between positive and negative classes. If the: AUC > 0.7 -it indicates an acceptable model.

AUC > 0.8 – it indicates a good model.

AUC > 0.9 -it indicates an excellent model.

Model Interpretation

The interpretation of predictors was done using the Odds Ratio (OR),95% Confidence Interval (CI) and p-values. The confidence intervals (CI) and the p-values assess the statistical significance of predictors (Dominguez-Almendros et al., 2011). The interpretation focused on assessing the increase or decrease in the likelihood of severe accidents given specific driver characteristics. It measures how a predictor affects the odds of the event.

$Odds = \frac{p}{1-p}$ where p is the probability of a severe accident.

The coefficients (β_i) are transformed into ratios using the formula:

$OR = e^{\beta_i}$ $where\ \beta_i$ is the regression coefficient, $i = 1,2,3 \dots k$.

If the:

OR > 1: it indicates a positive association with the outcome which means the predictor increases odds of the event

OR<1: it indicates a negative association with the outcome which means the predictor decreases odds.

OR=1: it means the predictor has no effect.

Model Refinement

All the non -significance predictors were removed from the model to simplify it.The final generated model was expressed in terms of the significant predictors only.

Model Application

The generated model from the used variables consisted of the significant predictors only. It will be used to predict accident severity on new data. The high-risk scenarios will be identified for targeted interventions.

RESULTS AND DISCUSSIONS

Frequency Tables

The frequency distributions below describe the characteristics of a sample of 500 drivers involved in road traffic accidents in Zimbabwe on accident severity predictors stated in each table.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Table 2: Frequencies for Accident Severity

Accident Severity	Frequency	Percent
0	319	63.8
1	181	36.2
Missing	0	0
Total	500	100

Sources: Authors` computation using JASP 0.95.4.0.

Table 2 above shows that 63.8% of the accidents were classified as severity level 1 and 36.2% as non-severe level 0. Most recorded accidents were non-severe. However, the proportion of severe accidents (36.2%) is still high, indicating a significant public health and safety concern.

Table 3: Frequencies for Driving Experience

Driving Experience	Frequency	Percent
0	86	17.2
1	134	26.8
2	280	56.0
Missing	0	0
Total	500	100

Sources: Authors` computation using JASP 0.95.4.0.

Table 3 above shows that 56% of drivers have the highest level of experience (level 2=>5years) followed by 26.8 (level 1 = 2-5 years) and 17.2% with level 0 = <2years.This indicates that drivers with more experience are predominant in the data. This suggests that accident involvement is not limited to inexperienced drivers, experienced drivers also contribute significantly to accident statistics.

Table 4: Frequencies for Age Group

Age Group	Frequency	Percent
0	131	26.2
1	264	52.8
2	105	21.0
Missing	0	0
Total	500	100

Sources: Authors` computation using JASP 0.95.4.0.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Table 4 above shows that most drivers are in age group 1 (25-45years) with 52.8 %, followed by age group 0 (<25years) with 26,2 % and age group 2 (>45 years) with 21.0 %. Middle- aged drivers constitute the majority of accident-involved drivers. Younger and older drivers form smaller but notable proportions which may influence severity differently.

Table 5: Frequencies for Gender

Gender	Frequency	Percent
Female	103	20.6
Male	397	79.4
Missing	0	0
Total	500	100

Sources: Authors` computation using JASP 0.95.4.0.

Table 5 above shows more male drivers (79.4) than female drivers (20.6) were involved in accidents. Male drivers' dominant accident involvement. This may reflect higher exposure, risk-taking behaviour or driving frequency among men in Zimbabwe. Gender could be a significant predictor of accident severity.

Table 6: Frequencies for Alcohol Use

Alcohol	Frequency	Percent
No	437	87.4
Yes	63	12.6
Missing	0	0
Total	500	500

Sources: Authors` computation using JASP 0.95.4.0.

Table 6 above shows that 87.4% did not consume alcohol while driving but 12.6% did. A smaller proportion of drivers were to have used alcohol before the accident. Although a minority variable, alcohol use may strongly increase the odds of severe accidents.

Table 7: Frequencies for Fatigue

Fatigue	Frequency	Percent
No	410	82.0
Yes	90	18.0
Missing	0	0
Total	500	500

Sources: Authors` computation using JASP 0.95.4.0.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Table 7 above shows that 82% reported no fatigue while 18% reported fatigue. Most drivers were not in fatigue, but fatigue still represents a meaningful portion (18%) and may be associated with severe outcomes due to slowed reaction times.

Table 8: Frequencies for Mobile Phone Use

Mobile Phone Use	Frequency	Percent
No	395	79.0
Yes	105	21.0
Missing	0	0
Total	500	500

Sources: Authors` computation using JASP 0.95.4.0.

Table 8 above shows that 79.0% did not use a mobile phone while driving and 21% did. One fifth of drivers used mobile phones while driving. This distracted driving behaviour typically increases accident severity and is a relevant logistic regression predictor.

Table 9: Frequencies for Over speeding

Over speeding	Frequency	Percent
No	352	70.4
Yes	148	29.6
Missing	0	0
Total	500	500

Sources: Authors` computation using JASP 0.95.4.0.

Table 9 above shows that 70.4% did not overspeed and 29.6% did. Nearly one-third of drivers were over speeding. Since over speeding is directly linked to impact force, it is likely to be a significant predictor of severe accidents reinforcing the necessity for speed enforcement and public education.

Table 10: Frequencies for Time of Accident

Time of Accident	Frequency	Percent
Day	335	67.0
Night	165	33.0
Missing	0	0
Total	500	100

Sources: Authors` computation using JASP 0.95.4.0.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Table 10 above shows that 67.0% occurred during the day and 33.0% at night. Accidents occur more frequently during the day. However, night-time accidents, though fewer, might be associated with higher severity due to poor visibility and fatigue.

Table 11: Frequencies for Vehicle Condition

Vehicle Condition	Frequency	Percent
Good	401	80.2
Poor	99	19.8
Missing	0	0
Total	500	100

Sources: Authors` computation using JASP 0.95.4.0.

Table 11 above shows that 80.2% of vehicles were in good condition while 19.8% were in poor condition. Most vehicles were in good condition, but poor vehicle condition still accounts for a substantial portion and may sharply increase the likelihood of severe accidents necessitating vehicle safety checks.

Table 12: Frequencies for Road Condition

Road Condition	Frequency	Percent
Dry	392	78.4
Wet	108	21,6
Missing	0	0
Total	500	100

Sources: Authors` computation using JASP 0.95.4.0.

Table 12 shows that 78.4% occurred on dry roads and 21,6% on wet roads. Most accidents occurred on dry roads, but a significant proportion occur on wet surfaces. This highlights the importance of adjusting driving behaviour based on weather conditions.

Correlation and Multicollinearity Check

Table 13: Multicollinearity Diagnostic Results

	Tolerance	**VIF**
Driving Experience	0.522	1.916
Age Group	0.516	1.939
Gender	0.973	1.028
Alcohol Use	0.977	1.024
Fatigue	0.965	1.036

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Mobile Phone Use	0.963	1.038
Over speeding	0.946	1.057
Time of Accident	0.984	1.016
Vehicle Condition	0.968	1.034
Road Condition	0.948	1.055

Table 13 above shows that all predictors: driving experience, age group, gender, alcohol use, fatigue, mobile phone use, over speeding, time of accident, vehicle condition and road conditions, Tolerance values range from 0.516 to 0.977 and are well above the conventional cutoff of 0.10.VIF values range from 1.024 to 1. 939 which are also far below the commonly accepted limit of 10. These values are all well within the Acceptable thresholds. The lowest tolerance (0.516 for Age Group) is far above 0.20, meaning no risk of multicollinearity. The highest VIF (1.939 for Age Group) is far below 5, indicating a weak collinearity. Several variables (Gender, Alcohol Use, Fatigue, Mobile Phone Use) show near-perfect tolerance (≈0.97) and very low VIF (≈1.02–1.04), suggesting that they are statistically independent. The diagnostic results demonstrate that none of the predictor variables exhibit a problematic multicollinearity. The predictors do not strongly correlate with one another, meaning that each variable provides unique explanatory information about accident severity. Hence, since the multicollinearity is low, the logistic regression coefficients are stable, and the standard errors are not inflated. The effect sizes and significance values can be interpreted with confidence in subsequence regression analyses. All the predictors were retained in the model without violating the assumptions.

Logistic Regression Results for Accident Severity

Table 14: Coefficients of all the predictors

Model		**Estimate**	**Odds Ratio (OR)**	**95% Confidence Interval (CI)**	**p**
M_0	(Intercept)	-0.567	0.567	(0.473,0.681)	<0.01
M_1	(Intercept)	-1.608	0.200	(0.104,0.384)	<0.01
	Age Group	0.274	1.315	(0’912,1.897)	0.142
	Driving Experience	-0.472	0.624	(0.449,0.867)	0.005
	Gender (Male)	0.506	1.659	(0.994,2.769)	0.053
	Alcohol Use (Yes)	0.702	2.017	(1.130,3.603)	0.018
	Fatigue (Yes)	0.815	2.259	(1.367,3.734)	0.001
	Mobile Phone Use (Yes)	0.629	1.875	(1.160,3.031)	0.010
	Over speeding (Yes)	1.385	3.994	(2.600,6.136)	<0.01
	Time of Accident (Night)	0.086	1.090	(0.713,1.665)	0.691
	Vehicle Condition (Poor)	0.123	1.131	(0.685,1.867)	0.630
	Road Condition (Wet)	0.574	1.775	(1.105,2.851)	0.018

Sources: Authors` computation using JASP 0.95.4.0.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

Logistic Regression Equation with significant predictors

The dependant variable is the Accident Severity = 1 (Severe). The model was fitted using the predictor coefficients from Table 14 above. The Logistic Regression Model fitted is :

$$\begin{aligned}\text{logit}(p) = -1.608 - 0.472(\text{DrivingExperience}) + 0.702(\text{AlcoholUse-Yes}) + 0.815(\text{Fatigue-Yes}) \\ + 0.629(\text{MobilePhoneUse-Yes}) + 1.385(\text{Overspeeding-Yes}) \\ + 0.574(\text{RoadCondition-Wet}) \qquad \textit{Equation } 1\end{aligned}$$

The corresponding predicted probability Equation of a severe accident is:

$$p = \frac{1}{1+e^{-(-1.608-0.472X_1+0.702X_2+0.815X_3+0.629X_4+1.385X_5+0.574X_6)}} \qquad \text{Equation 2}$$

Where each X_i $\quad i = 1,2,3,4,5,6$,represents the coded predictor value.

The coefficients of the predictors were used to calculate the Odds Ratio (OR) for interpretation.

Interpretation of Predictors

Using results from Table 14 and Equation 1 and 2 above, the predictors were grouped into significant and non-significant. All significant predictors have p values less than 0.05 and the Confidence Intervals (CI) exclude 1and non-significant predictors p-value is greater than 0.05 and the Confidence Intervals (CI) include 1. The significant predictors were driving experience, alcohol use, fatigue, mobile phone use, over speeding and rod condition. The non-significant predictors were age group, gender(male), time of accident(night) and vehicle condition (poor).

The model intercept for the full model, $M_{1,}$ is -1.608 with OR=0.20, p < 0. 001.This is the baseline odds of experiencing severe accidents when all predictors are zero. A negative intercept suggests that without any risk factors, the baseline risk of a severe accident is low.

Significant Predictors (p< 0.05)

Table 13 results indicate that six factors significantly predicted accident severity: Driving Experience, Alcohol Use, Fatigue, Mobile Phone Use, Over speeding, and Road Condition.

1.Drving Experience (protective: OR = 0.624; 95% CI 0.449–0.867; p = 0.005). - Drivers with low experience are more likely to be involved in severe accidents. OR=0.624 means higher driving experience reduces the odds of severe accidents by 37.6% (1-0.624). This means that greater driving experience is a protective factor. It reduces the likelihood of a severe accident. The protective effect of greater driving experience aligns with Chen et al. (2021) findings that accident severity was more pronounced among novice and older drivers.

2.Alcohol Use-Yes (OR = 2.017; 95% CI 1.130–3.603; p = 0.018)– Drivers who consume alcohol before driving is more than twice as likely to be involved in a severe accident compared to sober drivers. These findings agree with Paleti et al. (2010) and McCarty & Kim (2024) who emphasize that human-related factors particularly aggressive or impaired driving, remain the primary contributors to road accidents worldwide. Alcohol impairment slows reaction time and impairs judgment which increase the severity of crashes. It also supports the work of Chibaro et al. (2024) which identified drunk driving as a critical factor in causing accidents in Zimbabwe.

3.Fatigue-Yes (OR = 2.259; 95% CI 1.367–3.734; p = 0.001). Fatigued drivers in our sample have 2.3 times higher odds of causing a severe accident underscoring fatigue's impact. Fatigue reduces driver alertness which increase the severity of crashes. This supports the argument by Eboli et al. (2020) that behavioural characteristics particularly fatigue account for the largest proportion of accident risk.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

4.Mobile Phone Use-Yes (OR = 1.875; 95% CI 1.160–3.031; p = 0.010)– Using a mobile phone while driving increases the odds of severe accidents by 88%. This is consistent with Adavikottu & Velaga (2021) and Haerani et al. (2019) findings that place distraction-related behaviours among the top behavioural contributors to severe accidents. Therefore, our mobile usage estimate fits well within the trends noted by others regionally to influence road safety issues.

5.Over Speeding-Yes (OR = 3.994; 95% CI 2.600–6.136; p < 0.01)– Over speeding is the strongest predictor in our model. Drivers who overspeed are approximately 4 times the odds of a severe crash compared to those who did not. This large effective size is consistent with regional evidence that speed is among the single strongest determinants of crash severity This aligns with findings by Michalaki et al. (2015) and Adavikottu and Velaga (2021) who reported that accident dynamics such as speed significantly increase the likelihood of high-severity crashes. Similarly, Muvuringi (2012) and Chibaro et al. (2024) identified excessive speeding as one of the most critical behavioural contributors to accidents in Zimbabwe. The consistent associations demonstrate that speed management remains a key intervention point for reducing RTA fatalities.

6.Road Condition-Wet (OR = 1.775; 95% CI 1.105–2.851; p = 0.018)– Wet Road conditions increase the odds of a severe accident by 77.5%. Regional literature commonly cites poor road surface and weather as important contributors to severe crashes. The increase we observed is consistence with the findings by Chen et al. (2021) that showed that bad weather and terrain increase the severe accident outcomes. It suggests a meaningful environmental contribution in addition to behavioural risks.

Non-significant Predictors (p > 0.05)

The non -significant predictors are not statistically significant but they show the directional effects. Factors such as Age Group, Gender, Time of Accident and Vehicle Condition did not reach statistical significance.

1.Gender-Male (p=0.053, OR=1.659) – Male drivers show 66% higher odds of severe accidents, aligning with evidence from Paleti et al. (2010) and Haerani et al. (2019) who associated young male drivers with risky behavioural patterns such as speeding or aggressive driving.

2.Age Group (p=0.142, OR=1.315, Time of Accident-Night (p=0.691, OR=1.090) and Vehicle condition-Poor (p=0.630, OR=1.131). These variables are not statistically significant in predicting severity in this model, though not significant, showed expected directional effects. This is consistent with Millicent et al. (2016) who found that both environmental and mechanical factors contribute to accident prevalence in Zimbabwe at varying degrees. While not statistically significant, the role of age as a moderating factor in behavioural outcomes, as highlighted by Haerani et al. (2019), who mentioned age as an influence of the accident severity.

Model Performance

Table 14: Performance metrics

	Value
AUC	0.720

Sources: Authors` computation using JASP 0.95.4.0.

An AUC of 0.720 indicates the acceptance predictive accuracy. The value suggests that the model has an acceptable capacity to distinguish between severe and non- severe accidents based on the given variable.

Model Validation

Table 15: Model Summary-Accident Severity

Model	Deviance	AIC	BIC	df	$\Delta\chi^2$	McFaddeR^2	Nagelkerke R^2	Tjur R^2	Cox & SnellR^2
M_0	654.6	656.6	660.8	494		0.00		0.00	
M_1	577.2	599.2	645.6	489	77.3	0.118	0.196	0.154	0.143

Sources: Authors` computation using JASP 0.95.4.0.

Note. M_1 includes Age Group, Driving Experience, Gender, Alcohol Use, Fatigue, Mobile Phone Use, Over speeding, Time of Accident, Vehicle Condition, Road Condition**.**

Table 15 above shows that the model validation results included the deviance and BIC, $\Delta\chi^2$ (Chi-square Difference) and Pseudo R^2 Measures.

Deviance

It measures the lack of fit of the model in which lower deviance indicates a better fit.The results show that M_0 (null model) has a deviance of 654,6 and M_1 (full model with predictors) has a deviance of 577.2. The reduction in deviance shows that adding the predictors significantly improves the model fit. The model with the driver characteristics (M_1) fits well the data significantly better than the null model.

AIC (Akaike Information Criterion) & BIC (Bayesian Information Criterion)

Lower AIC (599.2versus 656.6) and BIC (645.6 versus 660.8) in M_1 indicate a better model fit. This confirms that the full model balances goodness of fit and model complexity effectively.

$\Delta\chi^2$ (Chi-square Difference)

Table 15 above shows that $\Delta\chi^2 = 77.3$ with df = 5 (difference in degrees of freedom between M_0 and M_1). This is the likelihood ratio test, testing whether the full model is significantly better than the null model. It indicates that the predictors collectively improve the model's predictive ability.

Pseudo R^2 Measures

It consisted of McFadden R^2, Nagelkerke R^2, Tjur R^2 and Cox & Snell R^2. McFadden R^2 is 0.118 which indicates a low-moderate fit and it means the model explains 11.8% of the variation in accident severity. Nagelkerke R^2 is 0.196 which means the model explains 19.6% of the variation in accident severity. Tjur R^2 is 0.154 which shows a moderate discrimination between severe and non-severe accidents. Cox & Snell R^2 is 0.143 which shows that the model explains 14.3% of the variation in accident severity. These Pseudo R^2 values suggest that the model explains 12-20% of the variation in accident severity which is a better fit, given the complexity of the accident outcome. The results are consistent with modelling outcomes reported by Al-Ghamdi and 2002; Michalaki et al., 2015 that used logistic or generalized ordered logistic regression in modelling accident severity.

CONCLUSION

This study has effectively identified the predictors of road traffic accident severity in Zimbabwe using logistic regression modelling. Key factors such as Over speeding, alcohol use, fatigue, driving experience, mobile phone use and wet road conditions were found to considerably influence the likelihood of severe accidents. These behaviours align with global research that identifies human error as the dominant contributor to road crashes. Although demographic variables such as age, gender, time of accident and vehicle condition did not reach statistical significance, their directional effects remain consistent with existing literature and may become significant with larger or more balanced datasets. The logistic regression model provided an acceptable level of predictive accuracy and demonstrated that incorporating behavioral and environmental variables significantly improves the explanation of accident severity. The findings mirror evidence from Zimbabwe and other countries, reinforcing the conclusion that behavioral interventions, strict enforcement of traffic regulations and targeted training for inexperienced drivers are essential for reducing severe accident outcomes.

Limitations of Study

Despite producing important findings, this study is subject to some limitations that should be considered when interpreting the results. The data were obtained from reported road traffic accidents within specific districts, which may not fully capture all crash types occurring nationally. Minor crashes, unreported incidents, or accidents occurring in remote rural areas may be under-represented, creating possible selection bias toward more serious or police-attended cases. Several key predictors are based on self-reported or police-reported behaviours, including alcohol use, over-speeding, fatigue, and mobile phone use. These behaviours are highly sensitive and prone to under-reporting due to social desirability bias, fear of legal consequences or recall bias. Under-reporting biases risk estimates leading to uncertainty in estimated associations. These limitations highlight the need for future studies incorporating probability sampling, objective measures of driver behaviour and broader geographical coverage to strengthen the generalisability and causal interpretation.

RECOMMENDATIONS

Based on the findings of this study, the following recommendations are proposed:

1. Strengthen Speed Enforcement

Over speeding was the strongest predictor of accident severity. Deploying more speed cameras and radar systems at high-risk areas, especially during peak hours, increasing fines and penalties for repeat offenders can improve the situation.

2. Implement Anti-Drunk Driving Campaigns

Alcohol use doubled the odds of severe crashes. Conducting regular breathalyser checkpoints and implementing awareness campaigns targeting high-risk groups such as young male drivers can reduce alcohol use while driving.

3. Address Driver Fatigue

Fatigue significantly increased accident severity. Long-distance drivers are encouraged to do regular rest-breaks and promote policies that regulate hours of driving, especially for commercial drivers to combat driver fatigue. Introducing fatigue awareness campaigns during festive and peak travel seasons.

4. Enforce Laws Against Mobile Phone Use While Driving

Mobile phone distractions strongly predict severe crashes. Intensify police monitoring of mobile phone use on major roads. Promote hands-free alternatives and public educational campaigns.

5. Target Inexperienced Drivers in Training Programs

Low driving experience was associated with higher accident severity. The road traffic authorities may introduce structured defensive driving courses for new drivers and strengthen the road test to ensure better preparedness.

6. Improve Road Maintenance and Stormwater Management

Wet road conditions increased severity by 77%. The authorities may repair potholes and improve road drainage systems and install warning signs at slippery or water-logged sections.

7. Enhance Public Road Safety Education

The traffic safety authorities may implement nationwide campaigns focusing on risky behaviours highlighted in the study and promoting safe driving practices, particularly during high-risk periods such as night driving and use social media, radio and community meetings to reach diverse audiences.

8. Strengthen Data Collection and Road Safety Policy Enforcement

Improve accident data reporting systems for accurate policymaking. Align national road safety regulations with global standards recommended by WHO.

Ethical Considerations

This study adhered strictly to established ethical principles governing the use of secondary data, statistical modelling and research involving road traffic accidents records in Zimbabwe. The study relied solely on secondary accident data recorded routinely by traffic police officers during accident investigations. The informed consent was not required since the data did not involve direct interaction with human participants. However, ethical best practice dictates that such datasets must be handled responsibly with respect for privacy and confidentiality. The data were provided in anonymised form and used strictly for academic research purposes. All identifying information such as names, national identification numbers, vehicle registration numbers, phone numbers, addresses, and licence numbers was removed prior to the analysis. The researchers did not have access to any personal identifiers.

The results were presented in aggregate form (percentages, odds ratios and model coefficients), to ensure that no individual driver or accident could be traced or identified. No attempt was made to use the data for legal, punitive or discriminatory action against drivers, police officers or transport companies. The study avoided any interpretations or conclusions that could unjustly stigmatise specific groups such as gender, age categories or professional drivers. Findings were framed to improve road safety awareness and inform prevention strategies rather than assign blame. The researchers followed transparent methodological procedures, including accurate reporting of statistical models, coding schemes, assumptions and limitations. The results were not manipulated to favour any outcomes. All data transformations (coding, cleaning, recategorisation) were documented clearly. This transparency ensures replicability and strengthens the trustworthiness of the findings. The logistic regression models were used to identify risk factors, not to profile or target individual drivers. The models were validated to avoid misleading interpretations that could influence policy unfairly. Predictive results were interpreted cautiously and contextualised within broader road safety challenges in Zimbabwe. The findings will be disseminated responsibly through academic publications without revealing sensitive details. The emphasis will be on improving public road safety, driver education and accident prevention strategies.

ACKNOWLEDGEMENTS

All sources were acknowledged.

Conflicts of Interests

The authors declare that there was no conflict of interest.

Data Availability

The data used in this study is not publicly available due to the ethical considerations of the participants.

REFERENCES

1. Adavikottu, A., & Velaga, N. R. (2021). Analysis of factors influencing aggressive driver behaviour and crash involvement. Traffic Injury Prevention, 22(sup1), S21–S26.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue XI November 2025

2. Ahmed, S. K., Mohammed, M. G., Abdulqadir, S. O., El-Kader, R. G. A., El-Shall, N. A., Chandran, D., Rehman, M. E. U., & Dhama, K. (2023). Road traffic accidental injuries and deaths: A neglected global health issue. Health Science Reports, 6(5), e1240.
3. Al-Ghamdi, A. S. (2002). Using logistic regression to estimate the influence of accident factors on accident severity. Accident Analysis & Prevention, 34(6), 729–741.
4. Chen, S., Shao, H., & Ji, X. (2021). Insights into factors affecting traffic accident severity of novice and experienced drivers: A machine learning approach. International Journal of Environmental Research and Public Health, 18(23), 12725.
5. Chibaro, M., Munuhwa, S., Mupfiga, C., & Farai, F. C. (2024). Human factors and road traffic safety: An inquiry into the causes of upsurge in road traffic accidents in Harare Metropolitan Province, Zimbabwe. Asian Journal of Management, Entrepreneurship and Social Science, 4(03), 92–108.
6. Deepthi, A., Chaithrika, B., Teena, C., Manvitha, G., & Keerthana, G. S. (2023, December). Enhancing road safety: Predicting severity of accidents. In International Conference on Data Science, Machine Learning and Applications (pp. 775–783). Springer Nature Singapore.
7. Dominguez-Almendros, S., Benitez-Parejo, N., & Gonzalez-Ramirez, A. R. (2011). Logistic regression models. Allergologia et Immunopathologia, 39, 295–305.
8. Eboli, L., Forciniti, C., & Mazzulla, G. (2020). Factors influencing accident severity: An analysis by road accident type. Transportation Research Procedia, 47, 449–456.
9. Foya, D. (2019). Road traffic accidents, causes, trends and preventive measures: The case of City of Harare (2012–2018). International Journal of Research and Innovation in Social Science, 3(10), 776–792.
10. Goniewicz, K., Goniewicz, M., Pawłowski, W., & Fiedor, P. (2016). Road accident rates: Strategies and programmes for improving road traffic safety. European Journal of Trauma and Emergency Surgery, 42(4), 433–438.
11. Haerani, S., Parmitasari, R. D. A., Aponno, E. H., & Aunalal, Z. I. (2019). Moderating effects of age on personality, driving behavior, and driving outcomes. International Journal of Human Rights in Healthcare, 12(2), 91–104.
12. https://www.trafficsafety.co.zw.
13. McCarty, D., & Kim, H. W. (2024). Risky behaviors and road safety: An exploration of age and gender influences on road accident rates. PLoS ONE, 19(1), e0296663.
14. Michalaki, P., Quddus, M. A., Pitfield, D., & Huetson, A. (2015). Exploring the factors affecting motorway accident severity in England using the generalised ordered logistic regression model. Journal of Safety Research, 55, 89–97.
15. Munuhwa, S., Govere, E., Samuel, S., & Chiwira, O. (2020). Managing road traffic accidents using a systems approach: Case of Botswana—Empirical review. Journal of Economics and Sustainable Development, 10(11), 176–188.
16. Muvuringi, M. P. (2012). Road traffic accidents in Zimbabwe: Influencing factors, impact and strategies. Vrije University Amsterdam, KIT (Royal Tropical Institute).