Traffic & Transportation

Riding the Edge – Unveiling the Key Factors behind Injury Severity in Single-Vehicle Motorcycle Crashes

Marija FERKO[1], Dario BABIĆ[2], Ali PIRDAVANI[3]

Original Scientific Paper
Submitted: 10 Feb 2025
Accepted: 4 Apr 2025

[1] Corresponding author, mferko@fpz.unizg.hr, University of Zagreb Faculty of Transport and Traffic Sciences, Zagreb, Croatia; marija.ferko@uhasselt.be, UHasselt, Faculty of Engineering Technology, Diepenbeek, Belgium
[2] dario.babic@fpz.unizg.hr, University of Zagreb Faculty of Transport and Traffic Sciences, Zagreb, Croatia
[3] ali.pirdavani@uhasselt.be, UHasselt, Faculty of Engineering Technology, Diepenbeek, Belgium; UHasselt, Transportation Research Institute (IMOB), Hasselt, Diepenbeek, Belgium

Publisher:
Faculty of Transport and Traffic Sciences, University of Zagreb

ABSTRACT

Motorcyclists are among the most vulnerable road users, with single-vehicle crashes often resulting in severe or fatal injuries. Although some earlier studies focus on motorcyclists, crash statistics show significant space for motorcyclist safety improvement. This study examines the factors influencing crash severity by analysing police-reported crash data in the Republic of Croatia from 2017 to 2022, incorporating rider characteristics, roadway and environmental conditions, and crash circumstances. A data-driven approach was applied to assess the relative importance of these factors in determining single-vehicle motorcycle crash injury severity. This study utilised two different modelling approaches – multinomial logistic regression and random forest modelling to investigate the factors influencing crash severity. The results highlight that rider age, road type, speed conditions and alcohol consumption significantly influence crash outcomes. Older riders and crashes occurring on county and local roads were more likely to result in severe or fatal injuries. At the same time, inappropriate speed and collisions with roadside objects further increased the likelihood of fatal outcomes. The findings suggest the necessity of targeted interventions, including enhanced speed management, infrastructure improvements, stricter enforcement of alcohol regulations and advanced rider safety programs. To complement this research, future research should integrate more detailed behavioural and vehicle-specific data to refine injury severity predictions. This study provides valuable insights for policymakers and transportation safety professionals seeking to mitigate the severity of motorcycle crashes and enhance overall road safety.

KEYWORDS

motorcycle; crash severity; single-vehicle crash; safety.

1. INTRODUCTION

In road traffic, about 1.19 million people are killed each year worldwide, and half of those killed fall into the category of vulnerable road users, i.e. pedestrians, cyclists and motorcyclists [1]. Motorcycle crashes tend to lead to more severe injuries and fatalities compared to crashes that involve other types of motor vehicles, mainly because motorcyclists lack physical protection in a crash event [2]. Additionally, motorcyclists suffer specific injuries in crashes depending on the type of collision [3]. From 2010 to 2019, motorcycle fatalities decreased by 14% in the European Union, while overall road fatalities fell by 23%, leading to a slight increase in the proportion of motorcycle-related deaths [4]. A particularly worrying fact is that motorcyclists account for 30% of total road fatalities in Croatia [5]. According to the European Commission, the proportion of motorcycle crashes on rural roads in the EU was 54% in 2020, compared with 39% on urban roads [6].

Single-vehicle crashes, a type of road crash in which only one vehicle and no other road user is involved, have a higher fatality rate compared to other types of motorcycle crashes [7]. Those crashes can occur due to the rider losing control, falling during a turn or sudden motorcycle deceleration. Generally, motorcycles have a higher power-to-weight ratio than cars and are capable of very high speeds, which results in a high risk of severe injury [8]. In addition, pavement quality and infrastructural elements, such as longitudinal ridging or grooving of the road surface and raised road markings, can produce steering instability for motorcycles [2]. Also, research confirms that the presence of motorcycles can raise the likelihood of injury crashes by approximately 30% and nearly double the probability of fatal crashes [9].

When analysing motorcycle crash data, considerable efforts have been made to improve road infrastructure safety. Road authorities and road designers need prediction tools to analyse the potential safety issues, identify safety improvements, and estimate the potential effect of these improvements. Possible interventions include designing safer infrastructure, incorporating road safety features, improving post-crash care and improving driver/rider behaviour. Therefore, road crash prediction models and, more broadly, analytical or explanatory approaches that identify key risk factors serve as valuable tools for understanding the frequency and severity of crashes. Unlike relying solely on descriptive statistics, these models enable exploring more profound, underlying relationships between variables. Generally, several modelling procedures have been applied in the literature to estimate the severity of motorcyclists' injuries. Crash severity models explore the relationship between crash severity injury and contributory factors such as driver behaviour, vehicle characteristics, roadway geometry and road-environmental conditions.

The European Commission has also recognised the importance of researching motorcycle crashes [10]. Their study team suggests further research into injury causation mechanisms (including pre-injury events and differences between fatal and non-fatal injuries), the influence of travel patterns and risk factors identified, practical measures for reducing severe injuries and preventing fatalities, and a policy benchmarking review. To our knowledge, there is a limited number of similar studies in Europe, and, as stated earlier, motorcyclist safety is a significant concern, especially in the southern parts of this region. Before conducting the analysis, extensive research of the existing literature was carried out to familiarise oneself with previous research and to learn which approaches and data were used by other authors. This study aims to explore which factors and to what extent influence the severity of single-vehicle motorcycle crashes. For that purpose, a crash dataset from Croatia is utilised, and two modelling approaches are used to investigate the factors influencing the crash severity of motorcyclists.

2. LITERATURE REVIEW

To comprehend the context of similar research, we did an extensive literature review, focusing on previous findings and methods for revealing crash severity influencing factors. To fully understand the mechanism of reducing the severity of motorcycle crashes, there is a need to identify key factors that influence the crash severity.

2.1 Key factors influencing single-vehicle motorcycle crash severity

A combination of rider characteristics, environmental conditions, road infrastructure and crash circumstances influences the severity of single-vehicle motorcycle crashes. Speeding has been identified as the most critical factor, with studies showing that unadjusted speed is the primary cause of over half of all fatal single-vehicle motorcycle crashes [11, 12]. Research also found that motorcyclist violations, including speeding, overtaking and U-turns, increase fatal injuries in single-vehicle crashes [7]. Another study showed that behavioural factors such as speeding, alcohol consumption (even within legal limits) and helmet non-use further exacerbate risks, with helmet non-use increasing the likelihood of fatality by 5.58 times [13]. Blood alcohol content (BAC) above 0.2 g/L significantly increased the risk of severe injuries, with nighttime and weekend crashes further heightening fatality likelihood [14].

Rider characteristics, such as age and experience, are significant predictors of crash severity. Crashes involving young motorcyclists tend to have a lower risk of fatal injuries compared to older motorcyclists [7]. Studies consistently show that older riders face a higher risk of fatal injuries, i.e. they are more likely to sustain fatal injuries compared to younger riders [15-17]. Some research shows that pillion passengers cause a lower chance of fatal crashes [7], but others deny it [15].

Collisions with fixed objects such as trees, poles or traffic devices present the highest fatality risks, with a likelihood of fatality up to 12.28 times greater [13]. The probability of the fatal outcome increased by 12.2%

and 19.2% in the case of a collision between a motorcycle and fixed objects [18]. Moreover, the study showed an increased probability of fatal outcomes of 9.5% in run-off-road crashes. Significant risk factors include crash type, road width, terrain, road alignment and collision point, with fixed-object collisions being the most critical [19]. Furthermore, higher fatality risks are associated with non-urban roads, curved segments and high-speed limits (>60 km/h) [19]. In the study on single-vehicle motorcycle crashes in Japan, fixed-object collisions accounted for the majority of fatalities, with a fatality rate four times higher than self-skidding crashes [20]. Fatal crashes frequently occurred on car-only roads with segments 9–19.5 meters wide, curves and dry road surfaces [20]. Additionally, motorcycle-barrier crashes were found to be at risk of incurring severe injuries [21]. Higher road standards increase the probability of severe injuries and fatalities for motorcycle-involved crashes [22]. Moreover, the rider ejection, two-way roads with no physical separation, single vehicle and curve-aligned roadways are associated with a higher likelihood of fatal crashes [23]. Non-urban roads and roads without speed regulations also show higher fatality risks [11, 20]. High-speed roads (over 50 mph) are associated with a greater probability of fatal crashes [12].

Some temporal and visibility effects have also been noted. Due to increased driving risk during adverse weather conditions, many studies have been dedicated to the investigation of weather effects on the crash occurrence of motor vehicles. The influence of rain was not a strong indicator of motorcyclist crash probability, indicating that the occurrence of rain decreased the possibility of crashes [24]. The good weather was statistically significant and related to severe injury crashes [25]. Also, the probability of fatal/severe injury increases for crashes occurring during dry weather conditions, in the early morning hours or late afternoon and/or early evening hours and on weekdays [26]. Environmental factors such as dark conditions without street lighting significantly elevate the risk of fatal/serious injuries [17]. Riding at night, speeding, alcohol consumption and falling asleep are critical contributors to severe crashes, exacerbated by reduced visibility and challenging road curvature [27]. Conversely, daytime crashes were influenced by speeding, with holidays and peak traffic hours slightly reducing severity due to slower traffic speeds.

In summary, the most influential factors contributing to severe and fatal injuries in single-vehicle motorcycle crashes include speeding, collisions with fixed objects, non-urban road environments, nighttime and weekend riding, alcohol consumption, helmet non-use, older rider age and high-speed roadway conditions.

2.2 Statistical approaches to crash severity analysis

This subsection summarises statistical and machine learning models used in crash severity analysis. Also, attention is paid to the use of variables and datasets. In crash severity modelling, various statistical approaches have been investigated, such as binary and multinomial logit models, nested logit models, and mixed and ordered probit models, and more advanced approaches like machine learning methods have been used. The data used in modelling motorcycle crash severity typically include many details relating to the crash occurrence, including the number of vehicles involved, the age of crash victims, weather conditions, types of vehicles involved and crash type (e.g. run-off-road, head-on, side collision, etc.). Those details can be integrated into statistical models, such as discrete choice models, the most commonly used tools to analyse injury severity [28]. Since the dependent variable (i.e. crash severity) usually has multiple outcome categories (e.g. fatal, different injury levels and property damage), using logit and probit models to model the crash severity is standard practice [29]. Motorcycle crash data could lack some factors that are not observable or relevant and are not reported in crash data. Such information can include the size and speed of the motorcycle, rider experience and training. To address the problem of unobserved heterogeneity, studies on crash injury severity analysis have used mixed logit models, latent-class models with random parameters, latent-class models and bivariate/multivariate models with heterogeneity in means and variances. Random parameters (mixed logit) and finite mixture (latent class) methods can accommodate individual unobserved heterogeneity by allowing parameters to differ across observations [21].

Pervez et al. (2022) performed a random parameter logit model with heterogeneity in means and variances to identify the key factors influencing the injury severity of single-vehicle motorcycle crashes [7]. The study was performed in Pakistan, and crash data were extracted from a crash injury research project during a one-year period. The variables for the model analysis were chosen using two criteria: characteristics that have been studied previously in the literature and variables based on the local environment. Consequently, variables were categorised into four main categories: temporal, motorcyclist, roadway and crash characteristics. Similarly, Seyfi et al. (2023) used a random parameters (mixed) logit model (RPL) to evaluate the severity of injuries in motorcycle crashes [15]. A total of 10,897 non-intersection motorcycle crash datasets from the beginning of 2009 to November 2020 in the State of Victoria, Australia, were analysed in this study.

A study conducted by Wang (2022) utilised binary logit and mixed logit models to analyse factors influencing motorcyclist fatalities in single-motorcycle (SM), motorcycle-motorcycle (MM) and motorcycle-vehicle (MV) crashes [13].

Farid and Ksaibati (2021) modelled the severities of motorcycle crashes in Wyoming that occurred from 2008 to 2017 using random parameters [16]. They implemented binary logistic regression and mixed binary logistic regression modelling structures. The mixed binary logistic regression model exhibited a better fit than the binary logistic regression model according to the AIC (a measure to assess model fit) results. Shaheed and Gkritza (2014) used a latent class multinominal logit model to investigate the factors that affect crash severity outcomes in 3,644 single-vehicle motorcycle crashes from Iowa, USA, from the 2001–2008 period [21].

Quddus et al. (2001) conducted an analysis of motorcycle injury and vehicle damage severity using ordered probit models to explain how variations in the characteristics of the roadway, the rider, environmental factors and the motorcycle can lead to variations in different levels of injury severity and damage to the motorcycle [22]. An extensive database of 27,570 motorcycle crashes in Singapore was extracted from police reports over 9 years. The study included two variables: the severity of motorcycle crash injuries (three categorical levels: fatal, seriously injured and slightly injured) and the severity of damage to the motorcycle (four categorical levels: total wreck, extensive damage, slight damage and no damage).

The probability of a fatal outcome in motorcycle and fixed object crashes was also proven by Sivasankaran et al. (2021) in studying potential risk factors contributing to the severity of single-vehicle motorcycle crashes with an ordered logit model [18]. The dataset consisted of 16,542 single-vehicle motorcycle crashes from 2009–2017 in Tamilnadu, India.

Cheng et al. (2017) did a study using alternative Bayesian multivariate crash frequency models and weather data for predicting motorcycle crash severity [24]. They observed motorcycle crashes in San Francisco (2008–2013) and categorised them into four levels of severity: fatal, serious injury, other visible injury and complaint of pain. Five models were developed using a full Bayesian formulation accounting for correlations commonly seen in crash data and then compared for fitness and performance. Wei et al. (2021) applied a full Bayesian spatial random parameters logit model to analyse rural single-vehicle crash severity across passenger cars, motorcycles, pickups and trucks, utilising data from 29,525 crashes in Shandong Province, China (2015–2020) [17]. Crash severity was categorised into three levels (no injury, slight injury and fatal/serious injuries), focusing on fatal/serious outcomes.

Another approach is to apply deep learning to other aspects of transportation engineering and transportation safety analysis. Das et al. (2018) employed a deep learning technique to analyse 5 years of single-vehicle crashes in Louisiana [23]. The study contributes to the existing injury severity modelling literature by developing a deep learning framework called DeepScooter to predict motorcycle-involved crash severities. The severity of crashes is recorded as five injury levels (commonly known as the KABCO injury scale).

Some researchers used machine learning techniques to evaluate motorcycle injury severity [30, 31]. The study adopted random forest, support vector machine, multivariate adaptive regression splines and binary logistic regression techniques to predict injury severity outcomes of crashes [30]. Although all the models perform similarly in predicting motorcycle crashes, the random forest model was the best-fit model according to the misclassification rate and AUC (area under the curve) measure. Later, they used a random forest model to identify the factors influencing motorcycle crash injury severity. A study in Japan applied a deep Q-learning network-based imbalanced classification (DQNIC) model to predict fatal single-vehicle motorcycle crashes using data from 2019–2022 [19]. Wisutwattanasak et al. (2024) used the XGBoost-SHAP algorithm to analyse single-vehicle motorcycle crashes in Thailand under daytime and nighttime conditions, revealing key factors influencing crash severity [27].

Santos et al. (2024) conducted a study using data from Portugal (2010–2019) [14]. They applied machine learning techniques to analyse factors influencing motorcycle crash injury severity, identifying alcohol consumption, road type and helmet use as key contributors. Among seven machine learning models evaluated, random forest (RF) and logistic regression (LR) developed with a balanced dataset outperformed others in predicting fatalities.

2.3 Conclusions based on literature review

The reviewed studies highlight the complexity of methodologies used to analyse the severity of single-vehicle motorcycle crashes. Various statistical and machine learning approaches have been applied, including logit models (binary, multinomial and mixed logit), ordered probit models and deep learning techniques. While traditional statistical models, such as multinomial logit and probit models, provide interpretable relationships between predictor variables and crash severity, machine learning methods, including random forest, XGBoost and deep learning algorithms, offer enhanced predictive performance by capturing complex interactions. The diversity of methodologies underscores the ongoing need to balance interpretability and predictive accuracy in crash severity modelling. However, the existing research is still relatively modest, especially in the European area.

Regarding contributing factors, the findings consistently highlight key variables that influence crash severity. Speeding, alcohol consumption, road curvature and collisions with fixed objects are among the most critical determinants of severe and fatal motorcycle crashes. Nighttime riding, road surface conditions and rider demographics (such as older age groups) have been frequently associated with increased crash severity. Still, a gap concerning the role of road design and infrastructure in crash outcomes could be noted. Differences in study findings suggest further research to explore context-specific risk factors and their interactions.

The research problem in this study extends beyond the general understanding that motorcyclists are vulnerable road users. While motorcycle crashes have been widely studied, the focus has primarily been on multi-vehicle crashes. Single-vehicle motorcycle crashes present distinct risk factors and severity patterns that require dedicated investigation. Although similar research questions have been previously investigated, our study produces a novel analysis with insights comparable to other studies, utilising an uninvestigated crash dataset. Specifically, the study seeks to answer the following research questions:

1) Which factors significantly influence injury severity in single-vehicle motorcycle crashes?
2) Do influencing factors differ worldwide?
3) How do traditional statistical models (MLR) compare with machine learning models (RF) in predicting crash severity outcomes?

3. METHODOLOGY

The methodology section outlines this study's research design, data collection and analysis procedures. It explains the statistical models and techniques applied to identify the significant factors influencing crash severity, ensuring a comprehensive and systematic approach to the analysis.

3.1 Research data

The data used in this study comprises police (Ministry of the Interior, Republic of Croatia) records of road crashes involving a single motorcycle (single-vehicle crashes) from 2017 to 2022 (except for 2020 because of the incomplete dataset). In the mentioned period, 7,676 crashes involving motorcycles were recorded. After that, multi-vehicle crashes and crashes with an unknown GPS location/road type were filtered out, and a total of 2,170 single-vehicle crashes were analysed in this study. Data with incorrect/unrecorded GPS coordinates were excluded to avoid misclassifying the road type. The data describe crashes that occurred on the Croatian road network, which consists of motorways (primary roads), state and county roads (primary and secondary roads), lower-ranking roads (local and uncategorised) and urban roads/streets. It should be noted that all roads except motorways and urban roads are predominantly in rural areas. The dataset includes fundamental information on environmental conditions (road surface condition, weather, part of the year, visibility), rider characteristics (citizenship, gender, alcohol intoxication, age), infrastructure-related factors (road type, road characteristics, speed limit) and crash related data (crash type, crash circumstance, crash outcome severity) relevant to each crash. The list of used variables is shown in *Tables 1* and *2*. The comprehensive data provide a basis for examining the key determinants and patterns associated with motorcycle single-vehicle crashes.

3.2 Analysis methods

Following a review of previous research, random forest (RF) and multinomial logistic regression (MLR) methods were selected to utilise in the further analysis. Although RF and MLR are well-known methods and have been proven helpful for this type of research, we did not find a reference where they were used on a single-vehicle motorcycle crash dataset to compare their efficacy. The selection of (RF) and (MLR) was also influenced by the study objectives, the nature of the dataset, and the inherent differences between these two modelling approaches. RF was chosen for its ability to handle non-linear relationships, detect complex interactions among variables and rank feature importance, making it helpful in identifying key predictors of crash severity. Meanwhile, MLR was selected due to its interpretability and ability to estimate the likelihood of each crash severity level while controlling for multiple factors. All statistical analyses were conducted using SPSS (version 29.0) and R (version 4.3.3).

Multinomial logistic regression (MLR) is a widely used statistical method for modelling categorical dependent variables with more than two outcomes. Unlike binary logistic regression, which deals with two categories, MLR extends the approach by estimating the probability of an outcome falling into one of several possible categories relative to a reference category [32, 33]. It estimates a set of regression coefficients for each category, quantifying how predictor variables influence the likelihood of each outcome. This makes it particularly suitable for analysing crash severity, where the response variable can take multiple levels, such as no injury, mild injury or severe/fatal injury. The model's coefficients are interpreted through odds ratios, which indicate how changes in predictor variables affect the probability of a crash, resulting in a particular severity level. This robust method accommodates continuous and categorical predictors, making it suitable for complex real-world problems [34]. The advantages of MLR are its ability to handle categorical and continuous predictors simultaneously and its high level of interpretability compared to many machine learning models.

Random forest enhances the decision tree approach by employing an ensemble method that generates multiple decision trees [33, 34]. A decision tree is a supervised learning technique primarily used for classification, though it can also be applied to regression. It begins with a root node representing the first decision point, where the dataset is split based on the feature that best separates the data into distinct classes. Each subsequent split leads to another decision node, further refining the classification or a terminal node that assigns a final class prediction. The process of dividing the data into binary partitions is known as recursive partitioning. Unlike traditional decision trees, random forest constructs multiple trees, with each tree trained on a random subset of features rather than using the entire dataset. The final classification outcome is determined through majority voting across all trees, ensuring greater robustness and reducing overfitting [35, 36]. The number of predictor variables randomly selected at each node when constructing individual decision trees (mtry) is a key hyperparameter that controls the diversity of the trees in the ensemble [39-41]. A smaller number increases randomness, reducing tree correlation and improving generalisation, but may lead to weaker individual trees. On the other hand, a larger number allows trees to be more similar to each other, potentially improving accuracy but increasing the risk of overfitting. The Gini index, or impurity, measures how well a decision tree node separates the classes [39-41]. It quantifies the probability that a randomly chosen element in a node would be incorrectly classified if randomly labelled based on the distribution of class labels. In a graphical representation, the mean decrease in Gini measures how important each variable is in doing splits during the decision tree construction process within the random forest model. In other words, a greater mean decrease in Gini indicates that the variable plays a stronger role in classification.

To evaluate model performance, receiver operating characteristic (ROC) curves were generated for both, with the area under the curve (AUC) calculated as a performance metric. The AUC quantifies a classifier's ability to distinguish between classes, ranging from 0.5 (random chance) to 1.0 (perfect classification). Since the AUC considers the full range of classification thresholds, it is a reliable measure of model effectiveness. The AUC is computed by summing the areas of successive trapezoids under the ROC curve, providing a comprehensive assessment of classifier performance [42].

4. RESULTS

The descriptive analysis provides an overview of road and infrastructure characteristics, circumstantial factors and rider characteristics related to motorcycle crashes (*Table 1*). Most crashes occurred on state roads (39.6%), with 50.8% on curved road sections and 55.6% in areas with a 50 km/h speed limit. Dry and clean surfaces were predominant (85.3%), and clear weather was reported in 58.3% of cases. Regarding circumstantial factors, crashes were more frequent between April and September (78.1%) and during the day (70.6%). Run-off-road crashes (59.6%) were the most common, and inappropriate speed (69.3%) was the leading contributing factor. Nearly half of all crashes (46.5%) resulted in death or severe injuries. Rider characteristics showed that the vast majority of motorcyclists involved were male (95.9%), and 23.0% had alcohol in their system at the time of the crash.

Table 1 – Descriptive statistics of categorical variables

Variable	Description	Code	Sample (N)	Percentage (%)
Road type	Motorway	1	57	2.63%
	State road	2	859	39.59%
	County road	3	395	18.20%
	Local road	4	167	7.70%
	Uncategorised	5	504	23.23%
	Urban street	6	188	8.66%
Road characteristics	Curve	17	1102	50.78%
	Road stretch	18	668	30.78%
	Other	23	46	2.12%
	Intersection	41	275	12.67%
	Other intersection	42	20	0.92%
	Road object	43	18	0.83%
	Railway crossing	44	8	0.37%
	Pedestrian/bicycle-related	47	33	1.52%
Speed limit (km/h)	<= 30	1	155	7.14%
	40	2	303	13.96%
	50	3	1207	55.62%
	60	4	189	8.71%
	>= 70	5	316	14.56%
Road surface condition	Dry – clean surface	1	1851	85.30%
	Dry – sand, gravel on the surface	2	90	4.15%
	Wet surface	3	196	9.03%
	Other conditions	4	33	1.52%
Weather conditions	Clear	1	1265	58.29%
	Cloudy	2	312	14.38%
	Rain	3	84	3.87%
	Other conditions	4	509	23.46%
Part of the year	April - September	1	1694	78.06%
	October - March	2	476	21.94%
Part of the day	Day	1	1531	70.55%
	Night	2	573	26.41%
	Dusk	3	51	2.35%
	Dawn	4	15	0.69%
Crash type	Run-off-road	1	1293	59.59%
	Hitting an object on the road / a roadside object	2	279	12.86%
	Other crash types	3	598	27.56%
Crash circumstances	Inappropriate speed	1	1503	69.26%
	Late hazard detection / Sudden braking	2	119	5.48%
	Other circumstances	3	548	25.25%
Crash severity*	No injuries	0	330	15.21%
	Mild injuries	1	832	38.34%
	Death or severe injuries	2	1008	46.45%
Citizen of Croatia	Yes	1	1744	80.37%
	No	2	425	19.59%
	Unknown	3	1	0.05%
Gender	Male	1	2082	95.94%
	Female	2	88	4.06%
Alcohol in blood	No	0	1671	77.00%
	Yes	1	499	23.00%

** Crash severity – dependent variable*

The descriptive statistics for rider age indicate that the dataset includes 2,170 motorcyclists aged 14 to 86 (*Table 2*). The mean age is 39.13 years, with a standard deviation of 14.66, suggesting a relatively broad age distribution. The skewness value of 0.395 indicates a slight positive skew, meaning that the age distribution has a longer tail toward older riders. The kurtosis value of -0.643 suggests a relatively flat distribution compared to a normal distribution. These statistics highlight the presence of younger and older motorcyclists in the dataset, concentrating on middle-aged riders.

Table 2 – Descriptive statistics of riders' age

	N	Min	Max	Mean	Std. dev.	Variance	Skewness	Kurtosis
Rider age (yrs.)	2170	14	86	39.13	14.661	214.932	0.395	-0.643

The following sections examine the explanatory value of individual independent variables about different crash outcomes using random forest and multinomial logistic regression. While random forest will highlight variable importance and capture complex, non-linear relationships, multinomial logistic regression will provide interpretable estimates of how specific factors influence the crash severity levels. Rather than focusing solely on predictive accuracy, this analysis aims to enhance the understanding of key risk factors and their relative contributions to crash outcomes and offer valuable insights for road safety improvements.

4.1 Random forest (RF)

An RF classification model was developed to gain insight into the importance of independent variables and predict crash severity using a dataset of motorcyclist-involved crashes. The model was trained on 1,100 trees with three variables randomly selected at each split (mtry = 3). A feature importance analysis was conducted to further interpret the model's decision-making process. *Figure 1* illustrates the variable importance rankings from the RF model, measured by the mean decrease in the Gini index. Rider age is the most influential predictor, indicating that a rider's age substantially contributes to the model's ability to distinguish crash severity outcomes. Road type and speed limit are closely followed, underscoring the importance of roadway conditions and regulatory factors. Other notable predictors include road characteristics, weather conditions and crash type, while gender appears to have the lowest relative importance among the examined variables.

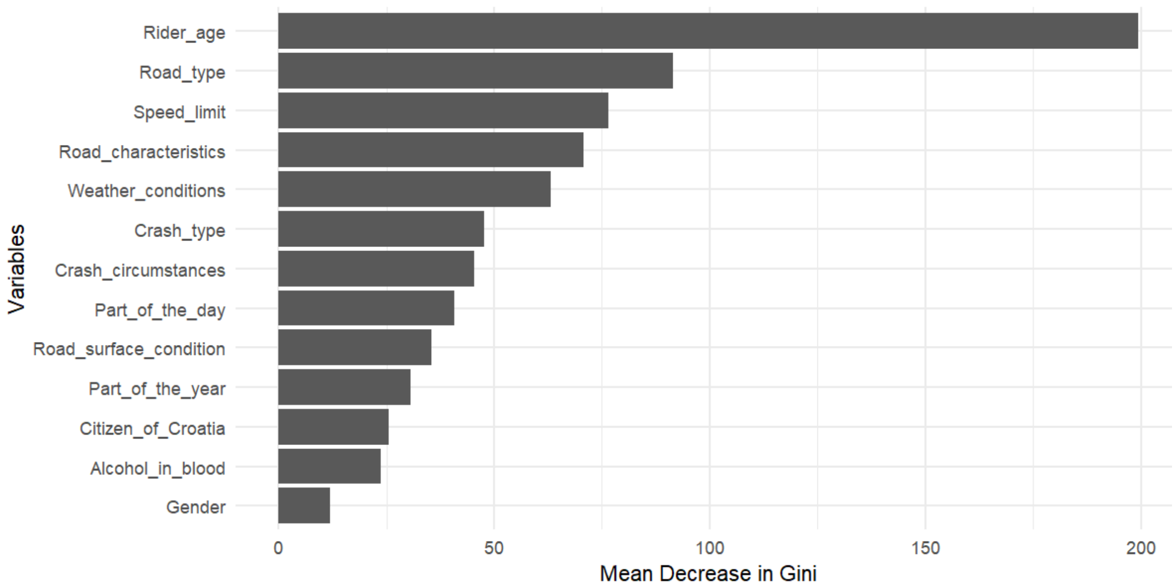


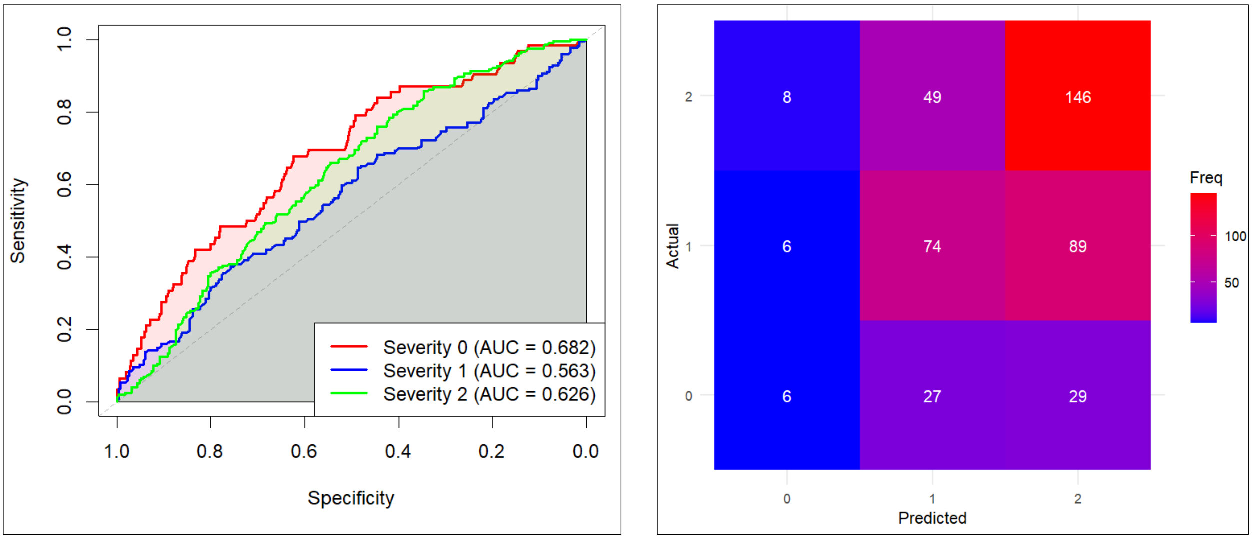
Figure 1 – Variable importance according to the random forest model

Table 3 presents the performance metrics for predicting crash severity outcomes. The model demonstrates the highest sensitivity for the death or severe injury category (0.724), indicating its relatively strong ability to identify severe crashes. However, sensitivity is considerably lower for no injuries (0.113), suggesting the model struggles to correctly classify cases where no injuries occur. Mild injuries fall in between, with a sensitivity of 0.444, reflecting moderate recognition of this category. In terms of specificity, the model performs best for no injuries (0.968), effectively excluding cases that do not belong to this category. Precision, which measures the proportion of correctly predicted cases within each category, is highest for severe injuries (0.557) and lowest for no injuries (0.368), suggesting that predictions for the no-injury category include a higher number of false positives. The F1-score balances sensitivity and precision and follows a similar trend, with the best performance observed for death or severe injury (0.630). The model's overall accuracy is 0.528, meaning that the model correctly classifies cases 52.8% of the time. While the model effectively captures patterns related to severe or fatal injuries, its performance in distinguishing no-injury cases indicates room for improvement.

Table 3 – Performance metrics of the random forest model for injury severity outcomes

Outcome (dependent variable)	Sensitivity	Specificity	Precision	F1-score	Accuracy
0 – No injuries	0.113	0.968	0.368	0.173	0.528
1 – Mild injuries	0.444	0.713	0.497	0.469	0.528
2 – Death or severe injury	0.724	0.494	0.557	0.630	0.528

Figure 2 presents a) the receiver operating characteristic (ROC) curves and b) the confusion matrix from the random forest model for each crash severity category. The areas under the curve (AUCs) indicate moderate classification performance, with values of 0.682 for no injuries, 0.563 for mild injuries and 0.626 for severe or fatal injuries. These results suggest that the model is relatively better at distinguishing no-injury cases, while its ability to classify mild injuries is notably weaker. In the confusion matrix, the diagonal elements represent correctly classified cases, whereas off-diagonal elements indicate misclassifications. The colour intensity reflects misclassification frequency, with darker shades indicating higher counts. While the model captures general patterns in crash severity classification, further improvements might enhance its ability to better distinguish between mild and severe injury outcomes.



a) *b)*

Figure 2 – a) ROC curves and b) confusion matrix for the random forest model

4.2 Multinomial logistic regression (MLR)

The results of the model fitting criteria indicate that the final multinomial logistic regression model provides a significantly better fit than the intercept-only model, as evidenced by the reduction in -2 Log Likelihood (from 4358.017 to 4085.802) and the significant likelihood ratio test ($\chi^2 = 272.215$, df = 70, $p < 0.001$). The goodness-of-fit tests further support the adequacy of the model, with both the Pearson ($\chi^2 = 4285.248$, df = 4178, p = 0.121) and Deviance ($\chi^2 = 4061.529$, df = 4178, p = 0.900) tests yielding non-significant results, indicating that the model does not show significant lack of fit. The pseudo-R-square values suggest that the model explains a modest proportion of the variance in the dependent variable (Cox and Snell $R^2 = 0.118$, Nagelkerke $R^2 = 0.136$, McFadden $R^2 = 0.062$), highlighting that while the model provides meaningful insights, additional factors may influence the outcome. Nonetheless, the significant improvement over the null model and the acceptable fit statistics support the model's suitability for analysing the relationship between predictor variables and crash severity outcomes.

Several factors, including road type, alcohol consumption, gender, rider age, nationality, crash circumstances, crash type and road surface conditions, are significantly associated with the outcome variable (*Table 4*). Road characteristics are borderline significant, meaning they may still contribute meaningfully but do not reach the conventional 0.05 threshold. These findings suggest that these factors should be carefully considered when analysing crash risks and their severity. At the same time, other variables, such as part of the day and weather conditions, appear less influential in this specific model.

Table 4 – Likelihood ratio tests for predictors in multinomial logistic regression

Variable	Chi-square	df	Sig.
Road type	36.726	10	<0.001
Part of the day	4.743	6	0.577
Alcohol in blood	7.183	2	0.028
Gender	11.174	2	0.004
Rider age	14.612	2	<0.001
Citizen of Croatia	21.632	4	<0.001
Road characteristics	23.608	14	0.051
Road surface condition	16.239	6	0.013
Weather conditions	5.543	6	0.476
Crash circumstances	35.054	4	<0.001
Crash type	25.640	4	<0.001
Part of the year	3.158	2	0.206
Speed limit	7.095	8	0.526

Table 5 presents the multinomial logistic regression results for predictors of crash severity, with death or severe injuries as the reference category. For clarity of results, only those parameters that showed statistical significance ($p <= 0.05$) are shown in the table. Each coefficient (B) and odds ratio (Exp(B)) describe how a predictor variable affects the likelihood of experiencing either no injuries or mild injuries compared to death or severe injuries.

Several comparisons to the reference road type show strong negative associations in the no injuries category. For instance, crashes on a county road have a coefficient of -1.240 ($p < 0.001$; Exp(B) = 0.289), suggesting about a 71% reduction in the odds of no injuries compared to severe/fatal injuries on that road type. A similar pattern appears for mild injuries, albeit with slightly smaller effect sizes (e.g. state road: -0.629, $p = 0.002$; Exp(B) = 0.533).

Negative coefficients dominate the no injuries category among all road characteristics (e.g. curves, intersections, road objects). For example, encountering a road object yields a coefficient of −3.227 ($p = 0.011$; Exp(B) = 0.040), indicating that crashes involving a road object are associated with a 96% lower chance of no injuries than severe or fatal injuries. Inappropriate speed negatively affects the odds of no injuries (-0.561, $p < 0.001$; Exp(B) = 0.571), confirming that higher speeds increase the risk of severe or fatal injuries. Conversely, late hazard detection/sudden braking is positively associated with no injury crashes (1.016, $p = 0.001$; Exp(B) = 2.761), meaning that such circumstances increase the odds of experiencing no injuries nearly threefold compared to severe injuries. Run-off-road crashes have negative coefficients in both the no injuries (−0.730, $p < 0.001$; Exp(B) = 0.482) and mild injuries (-0.300, $p = 0.016$; Exp(B) = 0.741) categories, indicating a higher likelihood of severe outcomes when a crash involves leaving the roadway.

For no injuries, male riders have a positive coefficient (1.341, $p = 0.007$; Exp(B) = 3.824), suggesting that males are nearly four times more likely to experience no injuries rather than severe injuries compared to females. Additionally, the absence of alcohol in the rider's blood increases the odds of no injuries (0.428, $p = 0.024$; Exp(B) = 1.535) and mild injuries (0.260, $p = 0.041$; Exp(B) = 1.297), reinforcing the protective effect of sobriety in reducing severe crash outcomes.

For no injuries, the coefficient for rider age is -0.017 ($p < 0.001$), with an odds ratio of 0.983, indicating that each additional year of age reduces the odds of experiencing no injuries rather than severe injuries by approximately 1.7%. For mild injuries, the coefficient is -0.009 ($p = 0.009$), with an odds ratio of 0.991, reflecting a 0.9% decrease in the odds of mild injuries for each additional year of age.

Older age, hazardous road characteristics (e.g. curves, intersections, road objects) and inappropriate speed increase the likelihood of severe or fatal injuries. On the other hand, late hazard detection, male gender and the absence of alcohol were shown to increase the likelihood of less severe crash outcomes.

Table 5 – Multinomial logistic regression parameter estimates for crash injury severity

Crash severity	Variable	B	Std. error	Wald	Sig.	Exp(B)	95% Confidence interval for Exp(B)	
							Lower bound	Upper bound
0 – No injuries	Rider age	-0.017	0.005	12.325	0.000	0.983	0.973	0.992
	Road type = State road	-1.016	0.251	16.412	0.000	0.362	0.221	0.592
	Road type = County road	-1.240	0.282	19.382	0.000	0.289	0.167	0.503
	Road type = Local road	-1.127	0.343	10.816	0.001	0.324	0.166	0.634
	Road type = Uncategorised	-0.683	0.249	7.503	0.006	0.505	0.310	0.823
	Road characteristics = Curve	-1.490	0.528	7.962	0.005	0.225	0.080	0.635
	Road characteristics = Road stretch	-1.242	0.525	5.609	0.018	0.289	0.103	0.807
	Road characteristics = Other	-1.464	0.663	4.872	0.027	0.231	0.063	0.849
	Road characteristics = Intersection	-1.114	0.540	4.260	0.039	0.328	0.114	0.945
	Road characteristics = Other intersection	-2.760	1.182	5.452	0.020	0.063	0.006	0.642
	Road characteristics = Road object	-3.227	1.267	6.480	0.011	0.040	0.003	0.476
	Crash circumstances = Inappropriate speed	-0.561	0.159	12.378	0.000	0.571	0.418	0.780
	Crash circumstances = Late hazard detection / Sudden braking	1.016	0.302	11.303	0.001	2.761	1.527	4.992
	Crash type = Run-off-road	-0.730	0.169	18.681	0.000	0.482	0.346	0.671
	Gender = Male	1.341	0.495	7.358	0.007	3.824	1.451	10.081
	Alcohol in blood = No	0.428	0.190	5.078	0.024	1.535	1.057	2.227
1 – Mild injuries	Rider age	-0.009	0.003	6.743	0.009	0.991	0.984	0.998
	Road type = State road	-0.629	0.205	9.362	0.002	0.533	0.357	0.798
	Road type = County road	-0.538	0.217	6.166	0.013	0.584	0.382	0.893
	Road type = Local road	-0.690	0.258	7.146	0.008	0.501	0.302	0.832
	Road characteristics = Curve	-0.982	0.483	4.143	0.042	0.374	0.145	0.964
	Crash type = Run-off-road	-0.300	0.125	5.782	0.016	0.741	0.580	0.946
	Alcohol in blood = No	0.260	0.128	4.163	0.041	1.297	1.010	1.665

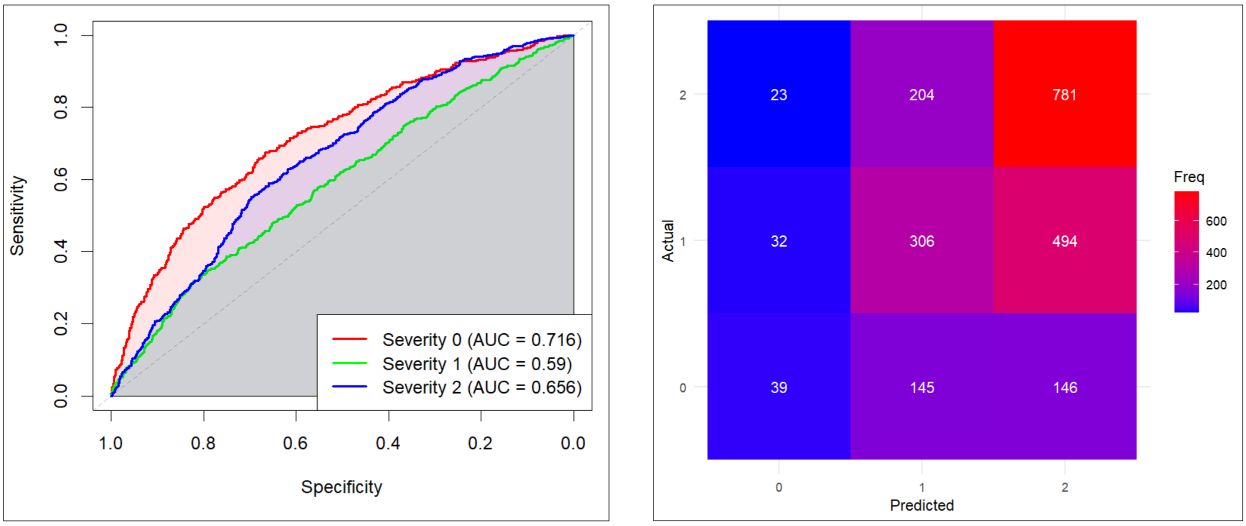
The reference category is 2 – Death or severe injuries.

Table 6 summarises the performance metrics of the MLR model for predicting injury severity outcomes (no injuries, mild injuries, death or severe injury). The model displayed its most substantial sensitivity for the most severe category, suggesting greater effectiveness at detecting cases resulting in severe outcomes. In contrast, sensitivity for mild injuries was more limited and especially low for cases with no injuries. Specificity was highest for no injuries, indicating strong performance excluding non-injury cases. This metric decreased for mild injuries and declined further for the most severe category, implying that some cases classified as severe may belong to other groups. A similar trend emerged for precision, which was comparatively higher for severe outcomes and lower for mild and no injuries. F1-scores, capturing the balance between sensitivity and precision, were also most significant for the most severe category, somewhat lower for mild injuries, and lowest for no injuries, aligning with the overall classification pattern.

Table 6 – Performance metrics of the multinomial logistic regression model for injury severity outcomes

Outcome (dependent variable)	Sensitivity	Specificity	Precision	F1-score	Accuracy
0 – No injuries	0.118	0.970	0.415	0.184	0.519
1 – Mild injuries	0.368	0.739	0.467	0.412	0.519
2 – Death or severe injury	0.775	0.449	0.550	0.643	0.519

Figure 3 illustrates the ROC curves for the three crash severity categories under the multinomial logistic regression model. The model demonstrates a fair level of discrimination for no injuries (AUC = 0.716) and a moderate ability to distinguish death or severe injury (AUC = 0.656) from other outcomes. In contrast, it performs relatively poorly in differentiating mild injuries from the other two categories (AUC = 0.590). This pattern suggests that while the model is most effective at identifying no-injury cases and can moderately discriminate severe outcomes, it struggles to classify mild injuries accurately. The trade-off between sensitivity and specificity is most favourable for the no-injury category, as indicated by the red curve lying furthest from the diagonal.



a) *b)*

Figure 3 – a) ROC curves and b) confusion matrix for the multinomial logistic regression model

5. DISCUSSION

This study explored the factors influencing the severity of single-vehicle motorcycle crashes using two distinct modelling approaches: random forest (RF) classification and multinomial logistic regression (MLR). The findings provide insights into the key predictors of crash severity and their relative contributions to injury outcomes.

5.1 Comparison of modelling approaches

The RF model effectively identified the most influential variables affecting crash severity. The mean decrease in the Gini index highlighted rider age, road type and speed limit as the strongest predictors. This aligns with previous research indicating that older riders face a higher risk of severe crashes due to decreased reaction time and physical vulnerability [13, 43, 44]. At the same time, roadway and speed conditions significantly influence crash outcomes. However, the random forest model exhibited moderate classification accuracy, with misclassification rates highest for mild injuries, suggesting that while it effectively captures complex interactions, its ability to distinguish between injury levels precisely remains limited. Conversely, the MLR model allowed an interpretable assessment of how individual factors influence crash severity. Several variables, including road type, alcohol consumption, gender, rider age, crash type and speed limit, significantly impacted injury severity. The model better predicted deaths and severe crashes than distinguishing mild injuries from no injuries. The ROC analysis confirmed this pattern, with higher AUC values for no and severe/fatal injuries but a lower value for mild injuries, indicating difficulties in differentiating intermediate severity cases. It has to be noted that some similar studies that used machine learning methods showed more favourable results but utilised a bigger dataset [14, 45].

While RF provides robust variable importance rankings and captures non-linear relationships, MLR offers clearer interpretability regarding the direction and magnitude of predictor effects. Considering both models in a complementary way strengthens the reliability of findings by balancing interpretability and predictive performance. MLR helps explain how individual factors contribute to different injury severity levels, while RF detects complex interactions that traditional parametric models may overlook. MLR is well-suited for policy discussions because it provides statistical significance levels and effect estimates, making it easier to communicate results to policymakers. On the other hand, RF is valuable for identifying the most influential variables, even in non-linearity and interactions. The conclusions are similar to comparable studies' findings, saying that RF outperformed LR in terms of predictive ability, while LR provided stronger interpretability of individual risk factors [14, 45].

5.2 Key factors

This study confirms the importance of rider age, road infrastructure, speed limits and alcohol consumption in determining crash severity, findings that are consistent with prior studies on motorcycle safety.

Literature shows that younger riders tend to behave riskily and more carelessly and are more likely to be involved in crashes [46-48]. On the other hand, our findings revealed that older motorcyclists exhibited a higher likelihood of severe injuries, which is a trend observed in other studies [15, 17, 49, 50]. Older riders are more likely to experience fatal or severe injuries, possibly due to reduced reaction time and increased physical vulnerability. This reinforces the need for targeted safety measures for ageing riders, including training programs focused on hazard perception and reaction time and promoting the usage of quality and certified personal safety equipment.

The findings also indicate that riders who are not under alcohol influence are less likely to sustain severe or fatal injuries, i.e. the absence of alcohol in a rider's organism reduces the likelihood of death and severe injuries. This is a trend consistent with other authors' research, confirming that alcohol in motorcyclists' blood might lead to severe or fatal crashes [13-15, 51]. Further, research confirms that road users' skills are significantly impaired with alcohol in the blood [52]. Intoxicated riders exhibited slower response times and allowed greater margins for error, resulting in more frequent performance mistakes during tasks [53]. These results highlight the continued importance of alcohol enforcement measures, such as random breath testing and public awareness campaigns, to reduce crash severity outcomes.

Contrary to some previous studies [54], male riders were found to have higher odds of no injuries than death or severe injuries, possibly due to greater physical resilience. However, it should be added that only 4% of the sample was female, which may have influenced this outcome. This is similar to results from Pakistan and Taiwan, where male motorcyclists were overrepresented in crashes but had slightly lower fatality rates than expected [8, 26].

This study shows that weather, season and visibility conditions do not influence the severity of the observed motorcycle crashes. This is somewhat unexpected and contrary to earlier findings [7, 17, 27, 55]. Generally, motorcyclists tend to ride during optimal weather and visibility conditions (i.e. warm, dry, daytime), which might explain why no association was found between crash severity and those factors.

Crashes on county and local roads were more likely to result in severe injuries, which aligns with earlier studies, where non-urban roads were associated with more fatal crashes [7, 11], as well as riskier behaviour [56]. Additionally, crashes occurring at curves or involving road objects increased the likelihood of severe outcomes significantly, further supporting previous findings on the dangers of high-speed cornering and roadside hazards [17, 20, 26, 57]. It is expected that crashes with more serious consequences can occur on lower-ranked roads, given that they are generally less well-maintained, have less police control, and route elements may not fully comply with road design rules.

Inappropriate speed was identified as a major risk factor, significantly increasing the likelihood of severe injuries [51]. This is consistent with previous research, which also highlighted fixed-object collisions and high-speed crashes as critical contributors to fatal injuries [7, 18, 19]. The current study further confirms that run-off-road crashes are particularly hazardous, in line with previous research [58]. Although speeding was recognised as an influencing factor, posted speed limits did not significantly impact crash severity outcomes. However, some studies indicate that the posted speed limit is positively associated with severe and fatal crash outcomes [59].

One potential explanation for this discrepancy is the distinction between posted limits and actual rider behaviour. The concept of inappropriate speed reflects context-specific riding that is too fast for road or traffic conditions rather than exceeding a legal threshold. Riders may exceed safe speeds even within legal limits, particularly on seemingly low-risk rural roads or during favourable weather, which could explain why the speed limit variable lacked significance. On the other hand, research shows that motorcyclists tend to overspeed more than other vehicle groups [60]. The lower the speed limit, the greater the speed excess (i.e. the amount by which the speed is exceeded) [61]. Considering the aforementioned and the vulnerability of motorcyclists, it is justified to conclude that speed significantly impacts crash outcomes, that is, the severity of the consequences.

These findings underscore the need for enforcement strategies and rider education campaigns beyond numerical limits, emphasising speed adaptation based on real-time conditions. When discussing educational strategies to improve motorcycle safety, it is essential to emphasise the importance of practical, skills-based training and not only theoretical instruction. Participation in certified motorcycle safety programs, particularly those incorporating controlled environments such as closed-course driving ranges, offers substantial benefits beyond the general goal of crash avoidance. For example, riders can practice appropriate responses to common hazards such as unexpected curves, loose gravel, road surface irregularities or sudden changes in pavement conditions. In addition to enhancing hazard perception and crash prevention, these programs should teach motorcyclists how to respond if a crash or near-crash situation becomes unavoidable. Learning how to react under pressure (such as managing panic, controlling the motorcycle during loss of traction or assuming a safer body posture during impact) can significantly reduce the severity of injuries. Structured training also enables riders to evaluate their skill levels in a safe and supportive environment, fostering self-awareness and encouraging more adaptive riding behaviour. A key component is the emphasis on situational awareness and the ability to adjust speed and positioning in response to rapidly changing road and traffic conditions. Literature confirms that training programs result in riders having fewer crashes and less severe motorcycle crashes [62-64].

Although the findings are broadly consistent with previous research, the detected differences justify the assumption of regional influence on traffic safety. Furthermore, this study contributes to the growing body of research on motorcycle crash severity by employing both RF and MLR to analyse the key factors influencing injury outcomes in single-vehicle motorcycle crashes.

In summary, the identified risk factors, such as rider age, alcohol consumption, inappropriate speed, road type and crash configuration, directly influence motorcycle crash severity by affecting both the likelihood of losing control and the rider's physical vulnerability upon impact. These factors shape the dynamics of crash occurrence and outcomes, either by increasing exposure to high-risk scenarios or reducing the rider's capacity to avoid or withstand a severe collision. Encouraging both proactive and reactive riding competencies, such as educational interventions, contributes meaningfully to reducing motorcycle crash severity and potentially the occurrence. Furthermore, improving road design and maintenance levels and adjusting law enforcement could also contribute significantly to a more motorcyclist-friendly environment. Additionally, infrastructure design should support appropriate speed behaviour through explicit road signalling and geometry that encourages safe speeds.

5.3 Limitations and future research

While this study provides valuable insights, several limitations should be considered. Firstly, the study relies on police-reported crash data, which may contain errors or lack details on rider behaviour, helmet use and motorcycle condition. These factors could further influence injury severity. However, police databases remain one of the most comprehensive and widely used sources for crash analysis. Furthermore, a more extensive dataset might be beneficial for greater accuracy and robustness of prediction models. Finally, the study focuses on single-vehicle motorcycle crashes in Croatia, which may limit the generalisability of the results. Therefore, expanding the dataset and international comparisons could enhance the robustness of the findings. Future research could integrate hospital data, naturalistic riding data or survey-based insights to complement the current findings. Although our current analysis focused exclusively on single-vehicle crashes, we acknowledge the value of comparing these with multi-vehicle crashes, especially in cases where the motorcyclist is at fault. Therefore, future research is planned to address this topic.

Regardless of the acknowledged limitations, the findings from our study provide valuable insight into what direction to take to reduce the severity of single-vehicle motorcycle crashes, especially when it comes to fatalities or severe injuries.

6. CONCLUSIONS

This study investigated the factors influencing the severity of single-vehicle motorcycle crashes using multinomial logistic regression and random forest classification. The findings highlight the significant role of rider demographics, road infrastructure, environmental conditions and crash circumstances in determining injury severity. The random forest model identified rider age, road type and speed limit as the most influential predictors of crash severity, emphasising the importance of roadway conditions and rider behaviour. Meanwhile, multinomial logistic regression provided interpretable insights, revealing that older riders, crashes on county and local roads, inappropriate speed and collisions with roadside objects significantly increase the likelihood of severe injuries or fatalities. Additionally, the study confirmed that the absence of alcohol consumption is strongly associated with less severe crash outcomes, reinforcing the need for stricter enforcement of impaired driving laws. Both models exhibited moderate classification performance, with higher predictive accuracy for severe and fatal crashes than for mild injuries. Despite this, the key variables identified were statistically significant in explaining crash severity, as indicated by MLR likelihood ratio tests, where road type ($p < 0.001$), rider age ($p < 0.001$), alcohol in the blood ($p = 0.028$), gender ($p = 0.004$), crash circumstances ($p < 0.001$) and crash type ($p < 0.001$) were all strong predictors.

The findings of this study have important practical implications for road safety and policy development. Enhanced speed management through stricter enforcement, automated speed control and improved road design can be crucial in reducing the severity of motorcycle crashes, particularly those resulting in fatal injuries. Infrastructure improvements on high-risk roads, especially county and local roads, should prioritise enhanced signalling and improved pavement conditions to mitigate crash risks. Given the increased vulnerability of older riders, targeted safety campaigns focusing on hazard perception and advanced riding techniques could help reduce the likelihood of severe crash consequences in this demographic. Furthermore, alcohol enforcement measures, including random breath testing and awareness campaigns, remain critical for addressing high-risk crashes associated with impaired driving. Additionally, run-off-road crash prevention strategies should be implemented, incorporating better roadside design, horizontal signalling and motorcyclist-friendly safety barriers to minimise the severity of crashes occurring on curves or involving fixed objects. These findings underscore the need for comprehensive, data-driven interventions to improve motorcycle safety and reduce crash severity on various road types.

In conclusion, this study underscores the need for multi-faceted interventions to enhance motorcycle safety. By integrating data-driven policy measures, improved infrastructure and rider education, policymakers can reduce the severity of motorcycle crashes and enhance road safety for motorcyclists.

ACKNOWLEDGEMENTS

This study was supported by the Special Research Fund (BOF) of Hasselt University with the BOF number “BOF21BL03”.

REFERENCES

[1] World Health Organization. *Road Traffic Injuries*. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries [Accessed 11th Apr. 2023].

[2] Elliot MA, et al. *Motorcycle safety: A scoping study*. TRL Report TRL581. 2003.

[3] Petit L, et al. A review of common motorcycle collision mechanisms of injury. *EFORT Open Reviews*. 2020;5(9):544-548. DOI: 10.1302/2058-5241.5.190090.

[4] Road Safety Thematic Report – Motorcycles. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport. Brussels, Belgium; 2023.

[5] Bulletin on Road Traffic Safety in 2023. Ministry of the Interior of the Republic of Croatia. Zagreb, Croatia; 2024.

[6] Facts and Figures – Motorcyclists and Moped Riders – 2021. Brussels, Belgium; 2021. https://road-safety.transport.ec.europa.eu/system/files/2022-03/FF_powered_twowheelers_20220209.pdf [Accessed 16th May 2023].

[7] Pervez A, Lee J, Huang H, Zhai X. What factors would make single-vehicle motorcycle crashes fatal? Empirical evidence from Pakistan. *International Journal of Environmental Research and Public Health*. 2022;19(10):5813. DOI: 10.3390/ijerph19105813.

[8] Lin HY, et al. Environmental factors associated with severe motorcycle crash injury in university neighborhoods: A multicenter study in Taiwan. *International Journal of Environmental Research and Public Health*. 2022;19(16). DOI: 10.3390/ijerph191610274.

[9] Afshar F, Seyedabrishami S, Moridpour S. Application of extremely randomised trees for exploring influential factors on variant crash severity data. *Scientific Reports*. 2022;12(1):11476. DOI: 10.1038/s41598-022-15693-7.

[10] Aarts LT, et al. *Study on serious road traffic injuries in the EU*. Brussels, Belgium.

[11] Tollazzi T, Parežnik LB, Gruden C, Renčelj M. In-depth analysis of fatal motorcycle accidents—Case study in Slovenia. *Sustainability*. 2025;17(3):876. DOI: 10.3390/su17030876.

[12] Savolainen P, Mannering F. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. *Accident Analysis & Prevention.* 2007;39(5):955-963. DOI: 10.1016/j.aap.2006.12.016.

[13] Wang MH. Investigating the difference in factors contributing to the likelihood of motorcyclist fatalities in single motorcycle and multiple vehicle crashes. *International Journal of Environmental Research and Public Health*. 2022;19(14):8411. DOI: 10.3390/ijerph19148411.

[14] Santos K, Firme B, Dias JP, Amado C. Analysis of motorcycle accident injury severity and performance comparison of machine learning algorithms. *Transportation Research Record: Journal of the Transportation Research Board*. 2024;2678(1):736-748. DOI: 10.1177/03611981231172507.

[15] Seyfi MA, Aghabayk K, Karimi Mamaghan AM, Shiwakoti N. Modeling the motorcycle crash severity on nonintersection urban roadways in the Australian state of Victoria using a random parameters logit model. *Journal of Advanced Transportation*. 2023;2023:1-12. DOI: 10.1155/2023/2250590.

[16] Farid A, Ksaibati K. Modeling severities of motorcycle crashes using random parameters. *Journal of Traffic and Transportation Engineering (English Edition)*. 2021;8(2):225-236. DOI: 10.1016/j.jtte.2020.01.001.

[17] Wei F, et al. Investigating rural single-vehicle crash severity by vehicle types using full Bayesian spatial random parameters logit model. *Applied Sciences*. 2021;11(17):7819. DOI: 10.3390/app11177819.

[18] Sivasankaran SK, Rangam H, Balasubramanian V. Investigation of factors contributing to injury severity in single vehicle motorcycle crashes in India. *International Journal of Injury Control and Safety Promotion*. 2021;28(2):243-254. DOI: 10.1080/17457300.2021.1908367.

[19] Zulherman D, Yang J, Shimizu K, Yokota Y. Deep Q-learning network-based imbalanced classification for fatality prediction of single-vehicle motorcycle crashes. *International Journal of Intelligent Transportation Systems Research*. 2024. DOI: 10.1007/s13177-024-00440-1.

[20] Zulherman D, Yang J, Yokota Y. Examining the factors affecting single-vehicle motorcycle crashes using association rules mining: Case study of Japan. In: *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE; 2022. p. 1222-1227. DOI: 10.1109/ITSC55140.2022.9922376.

[21] Shaheed MS, Gkritza K. A latent class analysis of single-vehicle motorcycle crash severity outcomes. *Analytical Methods in Accident Research*. 2014;2:30-38. DOI: 10.1016/j.amar.2014.03.002.

[22] Quddus MA, Noland RB, Chin HC. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. 2002. www.elsevier.com/locate/jsr.

[23] Das S, et al. Using deep learning in severity analysis of at-fault motorcycle rider crashes. *Transportation Research Record: Journal of the Transportation Research Board*. 2018;2672(34):122-134. DOI: 10.1177/0361198118797212.

[24] Cheng W, et al. Predicting motorcycle crash injury severity using weather data and alternative Bayesian multivariate crash frequency models. *Accident Analysis & Prevention*. 2017;108:172-180. DOI: 10.1016/j.aap.2017.08.032.

[25] Pai CW, Saleh W. Modelling motorcyclist injury severity by various crash types at T-junctions in the UK. *Safety Science*. 2008;46(8):1234-1247. DOI: 10.1016/j.ssci.2007.07.005.

[26] Waseem M, Ahmed A, Saeed TU. Factors affecting motorcyclists' injury severities: An empirical assessment using random parameters logit model with heterogeneity in means and variances. *Accident Analysis & Prevention*. 2019;123:12-19. DOI: 10.1016/j.aap.2018.10.022.

[27] Wisutwattanasak P, et al. Factors Affecting Single and Multivehicle Motorcycle Crashes: Insights from Day and Night Analysis Using XGBoost-SHAP Algorithm. *Big Data and Cognitive Computing*. 2024;8(10):128. DOI: 10.3390/bdcc8100128.

[28] Anarkooli AJ, Hosseinpour M, Kardar A. Investigation of factors affecting the injury severity of single-vehicle rollover crashes: A random-effects generalized ordered probit model. *Accident Analysis & Prevention*. 2017;106:399-410. DOI: 10.1016/j.aap.2017.07.008.

[29] Abdulhafedh A. Road crash prediction models: Different statistical modeling approaches. *Journal of Transportation Technologies*. 2017;7(2):190-205. DOI: 10.4236/jtts.2017.72014.

[30] Rezapour M, Mehrara Molan A, Ksaibati K. Analyzing injury severity of motorcycle at-fault crashes using machine learning techniques, decision tree and logistic regression models. *International Journal of Transportation Science and Technology*. 2020;9(2):89-99. DOI: 10.1016/j.ijtst.2019.10.002.

[31] Rezapour M, Farid A, Nazneen S, Ksaibati K. Using machine learning techniques for evaluation of motorcycle injury severity. *IATSS Research*. 2021;45(3):277-285. DOI: 10.1016/j.iatssr.2020.07.004.

[32] Upton GJG. Multinomial logistic regression. In: *Categorical Data Analysis by Example*. Wiley; 2016. p. 109-124. DOI: 10.1002/9781119450382.ch10.

[33] El-Habil AM. An application on multinomial logistic regression model. *Pakistan Journal of Statistics and Operation Research*. 2012;8(2):271. DOI: 10.18187/pjsor.v8i2.234.

[34] Abdillah A, et al. Application of multinomial logistic regression to analyze learning difficulties in statistics courses. *Journal of Physics: Conference Series*. 2020;1490(1):012012. DOI: 10.1088/1742-6596/1490/1/012012.

[35] Breiman L. Random forests. *Machine Learning*. 2001;45(1):5-32. DOI: 10.1023/A:1010933404324.

[36] Sarker IH, Salah K. AppsPred: Predicting context-aware smartphone apps using random forest learning. *Internet of Things*. 2019;8:100106. DOI: 10.1016/j.iot.2019.100106.

[37] Liu CY, Ku CY, Wu TY, Ku YC. An advanced soil classification method employing the random forest technique in machine learning. *Applied Sciences*. 2024;14(16):7202. DOI: 10.3390/app14167202.

[38] Dutta P, Paul S, Kumar A. Comparative analysis of various supervised machine learning techniques for diagnosis of COVID-19. In: *Electronic Devices, Circuits, and Systems for Biomedical Applications*. Elsevier; 2021. p. 521-540. DOI: 10.1016/B978-0-323-85172-5.00020-4.

[39] Louppe G, Wehenkel L, Sutera A, Geurts P. Understanding variable importances in forests of randomized trees. In: *Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'13)*. Curran Associates Inc.; 2013.

[40] Sikdar S, Hooker G, Kadiyali V. Variable importance measures for multivariate random forests. *Journal of Data Science*. 2025;243-263. DOI: 10.6339/24-JDS1152.

[41] Nembrini S, König IR, Wright MN. The revival of the Gini importance? *Bioinformatics*. 2018;34(21):3711-3718. DOI: 10.1093/bioinformatics/bty373.

[42] Melo F. Area under the ROC Curve. In: *Encyclopedia of Systems Biology*. Springer; 2013. p. 38-39. DOI: 10.1007/978-1-4419-9863-7_209.

[43] Pai CW, Saleh W. An analysis of motorcyclist injury severity under various traffic control measures at three-legged junctions in the UK. *Safety Science*. 2007;45(8):832-847. DOI: 10.1016/j.ssci.2006.08.021.

[44] Nunn S. Death by motorcycle: Background, behavioral, and situational correlates of fatal motorcycle collisions. *Journal of Forensic Sciences*. 2011;56(2):429-437. DOI: 10.1111/j.1556-4029.2010.01657.x.

[45] Wahab L, Jiang H. Severity prediction of motorcycle crashes with machine learning methods. *International Journal of Crashworthiness*. 2020;25(5):485-492. DOI: 10.1080/13588265.2019.1616885.

[46] Abdul Basit H, et al. Assessment of risk-taking behaviour of young motorcyclists at un-signalised intersections – A partial least square structural equation modelling approach. *Promet – Traffic & Transportation*. 2022;34(1):135-147. DOI: 10.7307/ptt.v34i1.3872.

[47] Liu CC, Hosking SG, Lenné MG. Hazard perception abilities of experienced and novice motorcyclists: An interactive simulator experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2009;12(4):325-334. DOI: 10.1016/j.trf.2009.04.003.

[48] Chang HL, Yeh TH. Motorcyclist accident involvement by age, gender, and risky behaviors in Taipei, Taiwan. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2007;10(2):109-122. DOI: 10.1016/j.trf.2006.08.001.

[49] Wang C, et al. Differences in single-vehicle motorcycle crashes caused by distraction and overspeed behaviors: Considering temporal shifts and unobserved heterogeneity in prediction. *International Journal of Injury Control and Safety Promotion*. 2023;30(3):375-391. DOI: 10.1080/17457300.2023.2200768.

[50] Champahom T, et al. A comparison of contributing factors between young and old riders of motorcycle crash severity on local roads. *Sustainability*. 2023;15(3):2708. DOI: 10.3390/su15032708.

[51] Shankar V, Mannering F. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. *Journal of Safety Research*. 1996;27(3):183-194. DOI: 10.1016/0022-4375(96)00010-2.

[52] Vu AT, Nguyen MT, Nguyen DVM, Khuat VH. Investigating the effect of blood alcohol concentration on motorcyclist's riding performance using an advanced motorcycle simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2020;73:1–14. DOI: 10.1016/j.trf.2020.06.010.

[53] Creaser JI, et al. Effects of alcohol impairment on motorcycle riding skills. *Accident Analysis and Prevention.* 2009:41(5):906–913. DOI: 10.1016/j.aap.2009.04.007.

[54] Champahom T, et al. Temporal instability of motorcycle crash fatalities on local roadways: A random parameters approach with heterogeneity in means and variances. *International Journal of Environmental Research and Public Health*. 2023;20(5):3845. DOI: 10.3390/ijerph20053845.

[55] Se C, et al. Empirical comparison of the effects of urban and rural crashes on motorcyclist injury severities: A correlated random parameters ordered probit approach with heterogeneity in means. *Accident Analysis & Prevention*. 2021;161:106352. DOI: 10.1016/j.aap.2021.106352.

[56] He J, Shi X, Xu Z, Hang W. Investigation and analysis of motorcycle safety in rural China. *Transportation Research Record: Journal of the Transportation Research Board*. 2012;2317(1):97-103. DOI: 10.3141/2317-12.

[57] Blackman RA, Haworth NL. Comparison of moped, scooter and motorcycle crash risk and crash severity. *Accident Analysis & Prevention*. 2013;57:1-9. DOI: 10.1016/j.aap.2013.03.026.

[58] Ferko M, et al. Influence of road safety barriers on the severity of motorcyclist injuries in horizontal curves. *Sustainability*. 2022;14(22):14790. DOI: 10.3390/su142214790.

[59] Zahid M, et al. Factors affecting injury severity in motorcycle crashes: Different age groups analysis using Catboost and SHAP techniques. *Traffic Injury Prevention*. 2024;25(3):472-481. DOI: 10.1080/15389588.2023.2297168.

[60] Broughton PS, et al. Conditions for speeding behaviour: A comparison of car drivers and powered two wheeled riders. *Transportation Research Part F: Traffic Psychology and Behaviour.* 2009;12(5):417–427. DOI: 10.1016/j.trf.2009.07.001.

[61] Ferko M, Pirdavani A, Babić D, Babić D. Exploring factors influencing speeding on rural roads: A multivariable approach. *Infrastructures (Basel)*, 2024;9(12):222. DOI: 10.3390/infrastructures9120222.

[62] McDavid JC, Lohrmann BA, Lohrmann G. Does motorcycle training reduce accidents? Evidence from a longitudinal quasi-experimental study. *Journal of Safety Research.* 1989;20(2):61–72. DOI: 10.1016/0022-4375(89)90032-7.

[63] Rowden P, Watson B, Haworth N. Measuring the effect of motorcycle rider training on psychosocial influences for risk taking. *27th International Congress of Applied Psychology*, *11-16 Jul. 2010, Melbourne, Australia*. 2010.

[64] Rowden P, Watson B. Motorcycle rider training and perceptions of skill. *2008 Australasian Road Safety Research, Policing and Education Conference, Nov 2008, Adelaide, SA, Australia.* 2008.

Vožnja na rubu: Ključni čimbenici koji utječu na težinu ozljeda u prometnim nesrećama s jednim vozilom – motociklom

Marija FERKO, Dario BABIĆ, Ali PIRDAVANI

Sažetak

Motociklisti se ubrajaju među najranjivije sudionike u prometu, pri čemu prometne nesreće u kojima sudjeluje samo motocikl često rezultiraju teškim ili smrtnim ozljedama. Iako su se ranija istraživanja djelomično bavila ovom temom, statistike o prometnim nesrećama ukazuju na značajan prostor za unapređenje sigurnosti motociklista. Ovo istraživanje analizira čimbenike koji utječu na težinu ishoda nesreće korištenjem podataka prikupljenim od strane policije u Republici Hrvatskoj u razdoblju od 2017. do 2022. godine, uzimajući u obzir karakteristike vozača, uvjete na cestama i okoliš, kao i okolnosti nesreće. Primijenjen je pristup temeljen na podacima radi procjene relativne važnosti ovih čimbenika u određivanju težine ozljeda u nesrećama s jednim vozilom – motociklom. U analizi su korištena dva različita modela – multinomijalna logistička regresija i model slučajne šume (engl. random forest) – kako bi se istražili čimbenici koji utječu na ishod nesreće. Rezultati pokazuju da dob vozača, vrsta ceste, uvjeti brzine i prisutnost alkohola značajno utječu na težinu ozljeda. Stariji vozači i nesreće koje se događaju na županijskim i lokalnim cestama imaju veću vjerojatnost da će rezultirati teškim ili smrtnim ishodom. Također, neprimjerena brzina i sudari s objektima uz cestu dodatno povećavaju vjerojatnost smrtnog ishoda. Nalazi istraživanja ukazuju na potrebu za ciljanom intervencijom, uključujući unaprijeđeno upravljanje brzinom, unapređenje infrastrukture, strožu provedbu zakona o konzumaciji alkohola i napredne programe edukacije motociklista. Kao nadopuna ovom istraživanju, buduća bi istraživanja trebala uključiti detaljnije podatke o ponašanju vozača i karakteristikama vozila radi preciznijeg predviđanja težine ozljeda. Ovo istraživanje pruža uvide za donositelje politika i stručnjake u području prometne sigurnosti s ciljem smanjenja težine posljedica motociklističkih nesreća i unaprjeđenja ukupne sigurnosti na cestama.

Ključne riječi
motocikl; težina prometne nesreće; prometna nesreća s jednim vozilom; sigurnost.