ORIGINAL PAPER   **Open Access**

Influence of seat-belt use on the severity of injury in traffic accidents

Check for updates

Juan Diego Febres[1] (iD), Susana García-Herrero[2*] (iD), Sixto Herrera[3] (iD), J. M. Gutiérrez[4] (iD), J. R. López-García[2] (iD) and Miguel A. Mariscal[2] (iD)

Abstract

Background: About 1.35 million people died in traffic accidents around the world in 2018, make this type of accidents the 8th cause of death in the world. Particularly, in Spain, there were 204,596 traffic accidents during 2016 and 2017, out of which 349,810 drivers were injured. The objective of this study was to understand to what extent seat belt non-use and human factors contribute to drivers injury severity.

Methodology: The results are based on the information and 2016–17 data provided by the Spain national traffic department "Dirección General de Tráfico" (DGT). The discretization model and Bayesian Networks were developed based on important variables from the literature. These variables were classified as; human factor, demographic factor, conditioning factor and seat belt use.

Results: The results showed that failure to wear the seat belt by drivers are likely to increase the risk of fatal and sever injury significantly. Moreover, distraction and road type road can contribute to the accident severity.

Keywords: Road traffic injuries, Seat belt, Bayesian networks, Traffic accidents, Human and conditioning factors

1 Introduction

Road traffic accidents are considered as one of the major life-threatening problems in the world that cause significant financial losses and long-term psychological problems. Every year, around 1.35 million people die on the road globally, making this the eighth cause of death globally, and the first leading cause of death in children and young adults aged 5 to 29 [1]. From the literature, the human factors contributing to the road traffic accidents include speeding, driving under the influence of alcohol and drugs, distraction and failure to wear seat belts and child restraint systems.

According to the Spain national traffic department "Dirección General de Tráfico" (DGT), in 2014, there were around 91,570 traffic accidents which resulted in 1688 fatalities, 9574 hospitalized injuries and 117,058 non hospitalized injuries. As 24% of the fatalities in car and vans were because of seat belt non-use, DGT aimed to reduce this figure by changing the driver's behavior and raising awareness about the importance of the use

of seat belts and child restraint system through various campaigns [2].

Previous research efforts investigated the role of human behavior in seat belt use [3]. They reported that in the United States, there is generally higher compliance with the seat belt regulations, however it is used more in the front seats than in the rear seats. It is also found that people who use public or private road transport (rear seats) in regions with less legal obligations are in need of mechanisms to promote their knowledge regarding the use of seat belt.

Furthermore, Høye [4] showed that seatbelt use can reduce fatal and non-fatal injuries in front and rear seat occupants by 60% and 44% respectively. In addition, it has been found that drivers without seatbelt are 8.3 times more likely to sustain fatal injury and 5.2 times more likely to sustain serious injury compared to the drivers who use seat belts. Bedard, Guyatt, Stones, and Hirdes [5] reported that the risk of death in traffic accident is conditioned on driver's characteristics and vehicle model. That is, women, elderly drivers, speeding, seat belt non-use, and recent model year vehicles are associated with an increased risk of fatality. Also, they found that left lateral impact is more likely to result in fatal injury than front impact. Moreover, J. M. Kim et al. [6] reported that seat belt non-use, the

* Correspondence: susanagh@ubu.es
[2]Escuela Politécnica Superior, Universidad de Burgos, Burgos, Spain
Full list of author information is available at the end of the article

driver age, the vehicle degree of deformation and the side collisions are positively associated with the serious traffic accident. Vallibhakara, Plitpolkarnpim, Suriyawongpaisal, and Thakkinstian [7] evaluated the socioeconomic factors affecting seatbelt use in Thailand. They found that people in urban and metropolitan areas tend to use seat belt more than people in rural areas. Also, gross provincial product, level of literacy and law penalty were found to be positively associated with seat belt use.

In addition to seat belt use, previous works showed that other factors corresponding to the driving behaviors such as overtaking, errors and distraction can contribute to the severity of traffic accidents [8–10]. Kaplan and Prato [11] found that drivers beyond the age of 55, female drivers, age and risky driving are likely to increase the risk of fatality. J.-K. Kim, Ulfarsson, Kim, and Shankar [12] reported the same results except that the male drivers have higher risk of fatality than female drivers. Moreover, alcohol consumption was found to increase the risk of fatality significantly.

From the methodological standpoint, numerous authors have analyzed accident severity using different methods. Zong, Xu, and Zhang [13] assessed the performance of Bayesian Networks and Regression Models in accident severity modeling. Their results showed that the Bayesian Networks outperformed the regression models and they are more suitable for the prediction of the accident severity. Previous works demonstrated that Bayesian Networks method gives validated and reliable results and it is an optimum method for assessing the probability of injury from a traffic accident [14, 15].

To this end, the aim of this study is to establish a probabilistic model based in Bayesian networks in order to predict the risk of injury and fatal injury in a traffic accident as a function of seat belt use, demographic factors, human factors, and conditioning factors (the type of vehicle, the type of road, the type of collision and the visibility of the driver).

2 Data collection

Data used in this study obtained from the Spain national traffic department [16], that collects data from the police accident reports. The database was restricted by the type of vehicle the driver was driving during the traffic accident, focusing only on those vehicles that have a seat belt. Taking these two conditions, the final database includes a total of 349,810 drivers. Driver injury severity (Accident severities) have been defined as: Fatal (FI), seriously injured (SI), lightly injured (LI), and unhurt (U). Also, as seen in Fig. 1, variables used in this study have been grouped into four factors: human factors (e.g., speed infringements, distraction), driver's characteristics (e.g., age and gender), conditioning factors (vehicle type, road type, collision type and
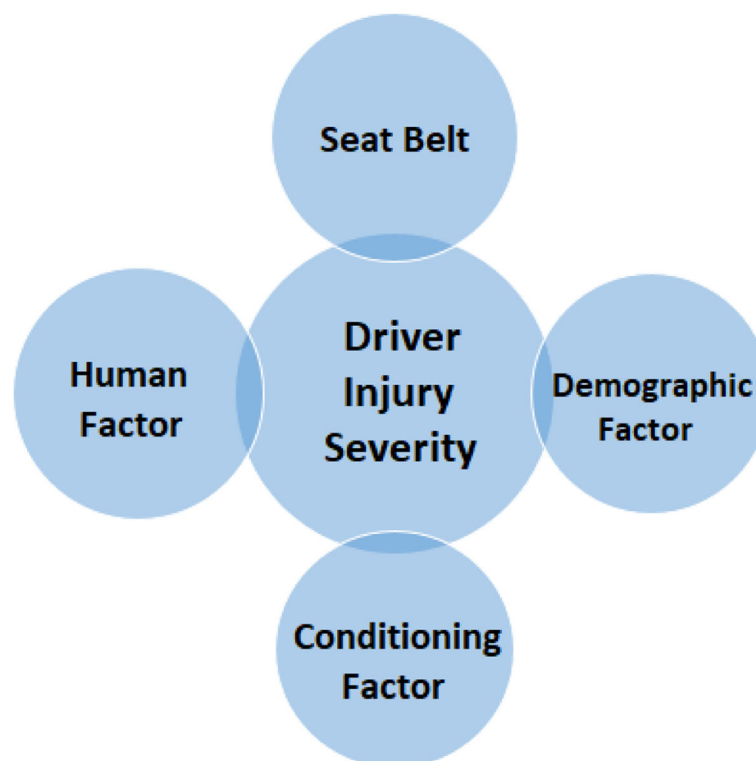
Fig. 1 Principal risk factors

the visibility of the driver) and seat belt use, **the latter belongs to the human** factors, but it has been studied separately. Figure 1, depicts the potential risk factors that can affect the driver's accident severity.

3 Methodology

3.1 Factors used in discretization model

Figure 2 shows a list of factors in the discretization model. The seat belt factor refers to the drivers that have access to use the seat belt depending on the type of car they drive; this variable is very important as it only limits the study to the vehicles that have a built-in safety belt (For example, data related to motorcycle accidents, bikes accidents, etc. were not considered). The demographic factors included 2 types of variables: age and gender. The variable 'gender' remains the same as mentioned in the police accident report. However, the variable 'age' has been grouped into four different age groups; less than 25, 26 to 40, 41 to 60, and over 60.

The human factor has been grouped in 4 class variables: speed infringements, distraction, errors and other type of infringements. The variable "speed infringements" has the same four states as shown in the police accident report; the first state is (a) none, indicating that the driver was driving with proper speed. (b) inadequate speed, that refers to exceeding speed in relation to the road condition, for example, a speed limit on a roadway might be 50 km/hr., but driving 45 km/hr. on icy roads could be dangerous. (c) over speed limit, which refers to the situation that the driver was driving over the allowable limit speed. (d) below the standard, that is the driver was driving the vehicle too slow. The group of the variables "errors" and "distraction" indicate that the driver did not make any error or were not distracted (errors occurs when the drivers do not see a traffic sign, a vehicle, a pedestrian, an

obstacle, etc.). Finally, the variable "other type of infringements", has also four states; (a) the first sate is 'none' which indicates that the driver didn't commit a specific infringement. (b) the second group indicate that the driver disrespected traffic signals. (c) the third state shows that the driver disrespected traffic norms. (d) fourth state indicates if the driver makes a reckless maneuver. All the errors and distractions included in the analysis are shown in the section of comments in Table 1.

The conditioning factor has been grouped in 4 types of variables: Vehicle type, road type, collision type and the visibility of the driver. The variable vehicle type has been discretized in 3 groups: cars, buses and trucks as they have built-in safety belt. The road type variable indicates that in what type of road occurs the traffic accident, is grouped in high speed road, medium speed road and low speed road. The variable collision type is grouped in four states that are: collision, run over, turn upside down and vehicle fall. Finally, the variable visibility has been discretized in adequate visibility, inadequate visibility and unknown. All these variables have specific content that are shown in Table 1.

Lastly, the dependent variable is driver's injury severity. Although MAIS3+ index measures the severity of injuries and is internationally recognized, in this study a variable is created that measures the severity of injuries from the data provided by the DGT. This variable has two values: firstly 'light' if the driver was slightly injured or unhurt, and secondly 'KSI' if the driver was either killed or seriously injured. It is noteworthy to mention that this study merely focus on driver's injury severity in relation to the seat belt use.

3.2 Bayesian networks

Bayesian Networks [17] are probabilistic graphical models [18] based on a directed acyclic graph (DAG)
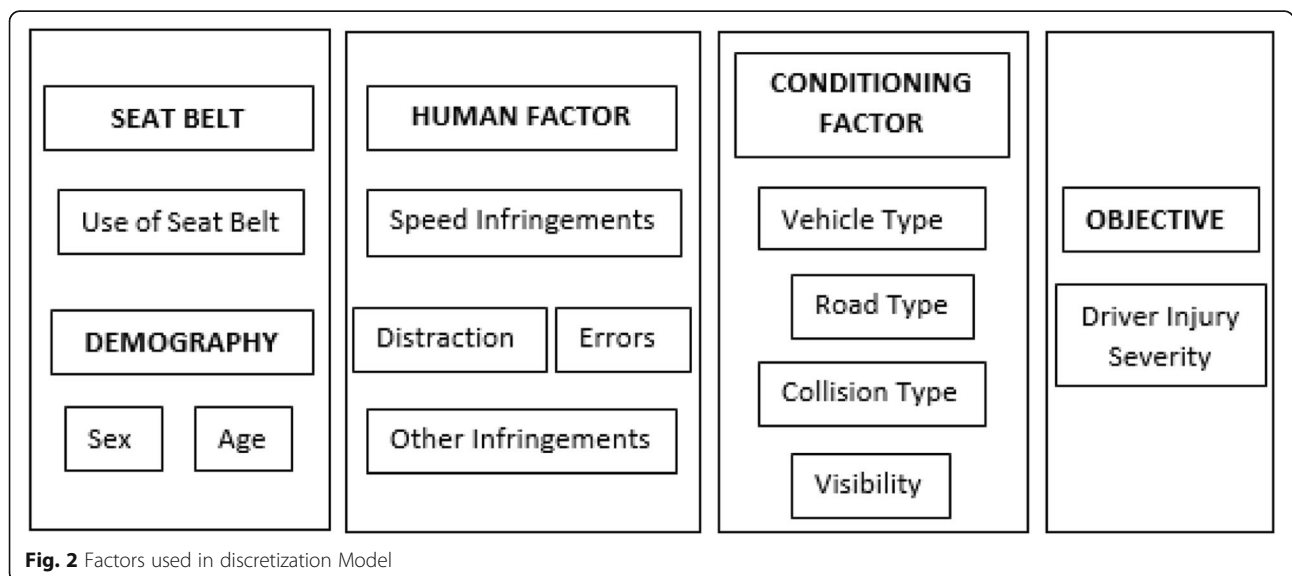
Fig. 2 Factors used in discretization Model

Table 1 Total number of cases analyzed. Processed by the Authors

Variables	N° cases			% of Reference	Comments
	2016	2017	Total Cases		
Vehicle Type					
Car	120,831	120,261	241,092	93.50%	Cars, van, all-terrain
Buses	2219	2181	4400	1.71%	Minibus <=17 passengers, bus, bi-articulated bus
Trucks	6159	6200	12,359	4.79%	Rigid Truck, Truck, articulated, articulated vehicle
Road Type					
High speed road	23,094	24,031	47,125	18.29%	Motorways, dual carriageway, others high speed roads
Medium speed road	29,460	31,604	61,064	23.70%	Conventional roads of 1 and 2 lanes
Low speed road	72,922	71,428	144,350	56.02%	Service ways, streets, neighbor roads, private roads
Others	3692	1451	5143	2.00%	Bike roads or similar
Collision Type					
Collision	101,674	99,830	201,504	78.18%	Frontal, side, and multiple collision, collision by range or against obstacle
Run Over	12,441	11,838	24,279	9.42%	Running over a person or an animal
Turn upside down	4197	4331	8528	3.31%	
Vehicle fall	1964	2021	3985	1.55%	Fallen down a mountain or in the city
Others	8892	10,541	19,433	7.54%	Get out of the way or similar
Age					
< 25	15,288	14,577	29,865	11.58%	
25–40	48,605	46,612	95,217	36.93%	
41–60	47,308	48,881	96,189	37.30%	
> 60	15,366	15,900	31,266	12.13%	
Unknown	2642	2672	5314	2.06%	
Gender					
Men	89,479	89,360	178,839	69.36%	
Women	38,520	38,317	76,837	29.80%	
Unknown	1210	965	2175	0.84%	
Seat-Belt					
Yes	95,804	96,039	191,843	74.40%	
No	3097	2803	5900	2.29%	
Not Necessary or Unknown	30,308	29,800	60,108	23.31%	
Infringement					
No infringement	42,277	40,173	82,450	31.98%	
Disrespect traffic signals	8348	8275	16,623	6.45%	Not respect the stop, yield, traffic light and others priority of step signals
Disrespect traffic norm	7179	7533	14,712	5.71%	Not respect the indications of a traffic agent, crosswalk and similar
Reckless maneuvers	15,169	15,691	30,860	11.97%	Like wrong way, zigzag circulation, wrong reverse circulation, dangerous overtake, stop without just cause, park in a prohibited places, car races, not indicate or indicate wrongly a maneuver, incorrectly change direction
Not Necessary or Unknown	56,236	56,970	113,206	43.90%	
Speed					
No infraction	58,387	57,423	115,810	44.91%	No infraction

Table 1 Total number of cases analyzed. Processed by the Authors *(Continued)*

Variables	N° cases			% of Reference	Comments
	2016	2017	Total Cases		
Inadequate	6205	5278	11,483	4.45%	Inadequate speed for road conditions
Exceeding	729	791	1520	0.59%	Exceeding the established speed
Slow	58	74	132	0.05%	Slow march / hindering circulation
Not Necessary or Unknown	63,830	65,076	128,906	49.99%	
Distraction					
No	33,529	33,582	67,111	26.03%	
Yes	8831	9647	18,478	7.17%	Like use: mobile phone, gps, handsfree, radio, dvd, smoke, others
Not Necessary or Unknown	86,849	85,413	172,262	66.81%	
Errors					
No	37,644	37,393	75,037	29.10%	
Yes	25,151	25,992	51,143	19.83%	Like don't see: a traffic sign, a vehicle, a walker, an obstacle, and others
Not Necessary or Unknown	66,410	65,256	131,666	51.06%	
Visibility					
Adequate Visibility	39,227	68,950	108,177	14.95%	
Inadequate Visibility	7041	9658	16,699	6.48%	
Unknown	82,941	50,034	132,975	51.57%	

which combine graphs and probability theories to efficiently learn the joint probability distribution of a multivariate problem involving discrete variables. On the one hand, the graphical structure given by the DAG defines the dependence (conditional or no) between the different variables considered in the model. On the other hand, these dependences reflected in the DAG define a factorization of the Joint Probability Distribution (JPD):

$$p(x_1, x_2, x_3, ..., x_n) = \prod_{i=1}^{n} p(x_i | \pi_i) \qquad (1)$$

Equation 1 stands for the Joint Probability Function of the Bayesian Network, *where* $\{x_1, ..., x_n\}$ are the variables considered in the model and π_i are the set of parents of the variable x_i given by the DAG. Finally, the DAG gives us a graphical and easily interpretable representation of the dependences between the variables.

Once the DAG and the JPD are obtained from the data [19], as new knowledge is evidenced for one or several variables of the model, it is then easily propagated to the rest of the BN to get the new probabilities (inference). In particular, the sensibility of the target variable to different pre-defined scenarios, given by combinations of the rest of variables, can be quantified by the changes of the corresponding probabilities. Moreover, a Bayes Classifier (BC) can be obtained by defining a probability threshold above/below of which the severity is considered serious/light.

All the calculations in this study have been done using the Bayes Net [20, 21] and MeteoLab [22] Toolboxes for Matlab [23].

3.3 Linear logistic regression

For the sake of the comparison, a linear logistic regression [13, 24, 25] is considered as benchmark to assess the performance of the Bayesian network. The proposed logistic regression model is a maximum-likelihood method commonly used for a binary classification problem and is given in eq. (2):

$$\text{logit(p)} = \log\left(\frac{p}{1-p}\right) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \cdot x_i \qquad (2)$$

where p is the probability of driver being severely injured, and x_i (i = 1, … n) refer to the variables considered in the model.

3.4 Validation

In order to evaluate the skill of the obtained classifier, a 10-fold cross validation approach has been considered defining a partition of the sample in 10-folds containing the 10% of the total sample. For each fold a model is obtained considering the other 90% of the sample, which is used to obtain a prediction of the fold's cases. As a result, 11 test-samples are obtained, one per fold and the latest one considering the prediction of the complete sample obtained by joining the 10 folds. Each test-

Table 2 AUC values and True Positives from Cross Validation (CV). Processed by authors

	CV-AUC	CV- TP
Bayesian Network	0.73 ± 0.03	4172 (5610)
Logistic Regression	0.87 ± 0.01	331 (5610)

sample was evaluated by using the Area Under the ROC [26] Curve (AUC), a standard measure of overall accuracy [27] for probabilistic and binary classifiers that varies from 0.5 (random guess) to 1 (perfect performance), obtaining 11 AUC values.

4 Results and discussion

This section presents the predictions of the probability of injury and fatal injury in a traffic accident as a function of seat belt use, demographic factors, human factors, and conditioning factors (type of vehicle, type of road, type of collision and the visibility of the driver). The validation of the Bayesian Network model and its corresponding graph are discussed in section 4.1. The initial probabilities are presented in section 4.2, and the sensitivity analysis results based on the Bayesian network are discussed in section 4.3 through section 4.6.

4.1 Validation and graph

Table 2 shows the results from the Bayesian Network and Logistic Regression models and their performance have been compared using AUC including the 10-fold

cross validation approach, and the number of True Positives (TP) given by the two methods.

Comparison of results indicates that, in spite of the better performance of the logistic regression in terms of AUC, this method strongly underestimates the number of accidents in which the drivers sustained severe injuries whereas the Bayesian Network is able to identify most of these cases. However, this can penalize the overall accuracy as reflected by the AUC. Moreover, the resulting Bayesian Network model is unique in a way that it allows to perform sensitivity analysis and to realize the variations in the risk of driver's injury severity due to changes in other independent variables. In contrast, as the proportion of drivers with fatal/severe injury in accident dataset is low, use of logistic regression would result in biased parameter estimates and subsequently inaccurate sensitivity analysis. Additionally, including many variables with different states to the logistic model would increase the complexity of the analysis and limit the interpretability of the results. As a result, the Bayesian Network model is used as a preferred model in the rest of the paper.

The uncertainty of the predictions obtained with the Bayesian model is based on the cross-validation step described in Section 3.4. However, the overall prediction ability of the model may not be consistent with the specific cases considered in the sensitivity analysis. To address this issue, the measure of fit (the percentage of the observations that is correctly predicted by the model) for each subsample related to each sensitivity analysis has been estimated and evaluated using cross-validation procedure. In this sense, the measure
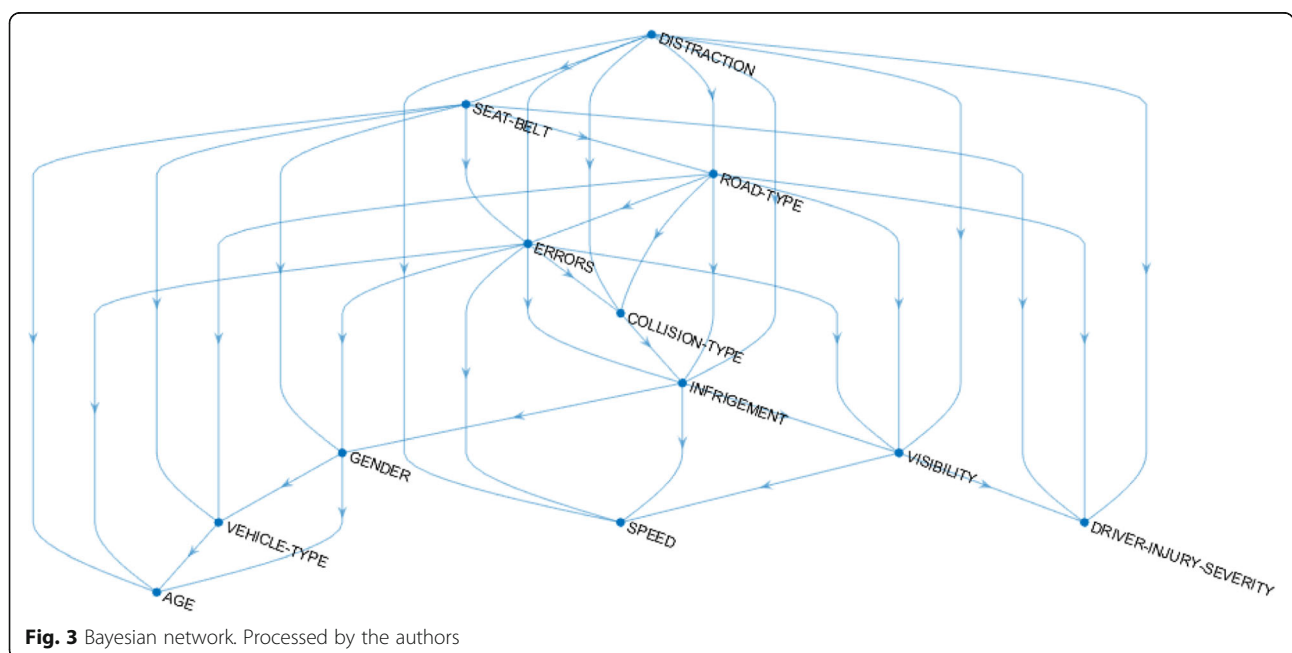
Fig. 3 Bayesian network. Processed by the authors

Table 3 A priori probability of the severity of the injury. Processed by the Authors

Variables	N° Cases	K/SI Risk	Comments
Vehicle Type			
Car	241,092	2.13%	Cars, van, all-terrain
Buses	4400	2.39%	Minibus <=17 passengers, bus, bi-articulated bus
Trucks	12,359	2.99%	Rigid Truck, Truck, articulated, articulated vehicle
Road Type			
High speed road	47,125	2.39%	Motorways, dual carriageway, others high speed roads
Medium speed road	61,064	5.53%	Conventional roads of 1 and 2 lanes
Low speed road	144,350	0.68%	Service ways, streets, neighbor roads, private roads
Others	5143	2.29%	Bike roads or similar
Collision Type			
Collision	201,504	2.12%	Frontal, side, and multiple collision, collision by range or against obstacle
Run Over	24,279	1.20%	Running over a person or an animal
Turn upside down	8528	5.15%	
Vehicle fall	3985	2.65%	Fallen down a mountain or in the city
Others	19,433	2.49%	Get out of the way or similar
Age			
< 25	29,865	2.19%	
25–40	95,217	2.17%	
41–60	96,189	2.17%	
> 60	31,266	2.25%	
Unknown	5314	1.77%	
Gender			
Men	178,839	2.22%	
Women	76,837	2.10%	
Unknown	2175	0.98%	
Seat-Belt			
Yes	191,843	2.24%	
No	5900	12.92%	
Unknown	60,108	0.90%	
Infringement			
No infringement	82,450	2.58%	
Disrespect traffic signals	16,623	2.37%	Not respect the stop, yield, traffic light and others priority of step signals
Disrespect traffic norm	14,712	2.00%	Not respect the indications of a traffic agent, crosswalk and similar
Reckless maneuvers	30,860	3.49%	Like wrong way, zigzag circulation, wrong reverse circulation, dangerous overtake, stop without just cause, park in a prohibited places, car races, not indicate or indicate wrongly a maneuver, incorrectly change direction
Unknown	113,206	1.51%	
Speed			
No infraction	115,810	2.65%	No infraction
Inadequate	11,483	3.62%	Inadequate speed for road conditions

Table 3 A priori probability of the severity of the injury. Processed by the Authors (Continued)

Variables	N° Cases	K/SI Risk	Comments
Exceeding	1520	4.08%	Exceeding the established speed
Slow	132	3.44%	Slow march / hindering circulation
Unknown	128,906	1.59%	
Distraction			
No	67,111	1.84%	
Yes	18,478	7.04%	Like use: mobile phone, gps, hands-free, radio, dvd, smoke, others
Unknown	172,262	1.78%	
Errors			
No	75,037	2.55%	
Yes	51,143	3.72%	Like don't see: a traffic sign, a vehicle, a walker, an obstacle, and others
Unknown	131,666	1.36%	
Visibility			
Adequate Visibility	178,839	3.60%	
Inadequate Visibility	76,837	5.09%	
Unknown	2175	0.64%	

of fit estimated for severe injuries falls between 84% and 91% whereas for no injuries it falls between 87% and 95%, reflecting the uncertainty of the predictions.

As shown in Fig. 3, the Bayesian Network graph depicts the significant dependencies among the variables. For example, seat belt, road type, visibility and distraction, are the only variables directly related to the dependent variable, driver injury severity. Hence, a joint analysis of is carried out on these four variables.

As indicated in Fig. 3, The graph represents the interactions among all the variables and gives information about how these variables relate to each other. This demonstrates that the network takes information from the whole data.

4.2 Initial probabilities of serious injury in a traffic accident

The study starts with a sensitivity analysis, based on the Bayesian model, to estimate the initial probabilities of suffering a serious or fatal injury for the drivers (which is called risk probability "KSI Risk") compared to minor injuries in each of the variables and model States. The results for each variables and states are summarized in Table 3.

By getting the initial probabilities from the sensitivity analysis, the most influential variables are as follows: Seat Belt (especially those who do not use seat-belt), Distraction (especially the group of drivers with distractions), Road type (especially the group of medium and high speed roads), Visibility (especially with inadequate visibility), and Speed (especially those who exceeded the

speed limits). The 'a priori' probabilities of suffering a serious and / or fatal injury in a traffic accident reach its highest level (12.92%) because of seat belt non-use (this does not account for the effects of other variables yet). This finding reinforces the study argument that seat belt use is the most influential variable.

The other variables which are associated with high serious/fatal injury probabilities are driving distractedly (7.04%), road type, specifically conventional roads with 1 and 2 lanes (5,53%), rollover accident (5,15%), inadequate visibility (5,09), speeding (4.08%) and making errors while driving (3,72%), implying that the drivers without seat belt are likely to have a fatal or serious injury when the vehicle is rollover or when they drive on medium speed roadway.

4.3 Probability of serious injury in a traffic accident based on the use of seat belt and the demographic factor

In this section demographic variables such as gender and age have been used to estimate the KSI probabilities given seat belt use. From Table 4, it can be understood that seat belt non-use would increase the risk of fatal or serious injury in men, women and all age groups more than 5 times, compared to seat belt use. Specifically, for "gender", the higher risk of suffering a serious and / or fatal injury is associated with the male drivers without seat belt, reaching (13.08%). As with "age", drivers under 25 and over 60 who do not use seat belts are more likely to sustain fatal or serious injury in a traffic accident. While driver between 41 and 60 have higher risk of fatality or serious injury compared to other age groups.

Table 4 Probabilities of suffering a serious injury, depending on the use of seat belt and the different variables of the demographic factor. Processed by the Authors

Variables	N° cases	KSI Risk medium	Use of Seat-Belt (KSI Risk)		
			Yes	No	Unknown
Gender					
Men	178,839	2.22%	2.23%	13.08%	0.91%
Women	76,837	2.10%	2.29%	12.57%	0.91%
Unknown	2175	0.98%	1.23%	3.30%	0.74%
Age					
< 25	29,865	2.19%	2.29%	15.41%	0.91%
25–40	95,217	2.17%	2.24%	13.16%	0.91%
41–60	96,189	2.17%	2.21%	11.64%	0.91%
> 60	31,266	2.25%	2.33%	15.58%	0.92%
Unknown	5314	1.77%	1.73%	13.16%	0.75%

4.4 Probability of serious injury in a traffic accident based on the use of seat belt and the conditioning factor

To analyze the probability of a serious and / or fatal injury (KSI risk) based on conditioning factor and seat belt use, a sensitivity analysis was carried out with respect to the variables "vehicle type", "road type", "collision type" and "visibility" in all its states.

As shown in Table 5 the difference between the probability of suffering a serious and / or fatal injury in a traffic accident with and without seat belt use are significant. As with "vehicle type", truck drivers without safety belt have the highest risk of fatality and serious injury (please refer to the comments in Table 5). While, bus drivers without seat belt are in lower risk of fatality or serious injury. One possible reason is that, on average, the bus drivers may travel slower compared to truck and car drivers as they carry may passengers. Also, for the "road type" variable, the two states, high and medium speed roads are found to contribute to the risk of injury and fatality of drivers who do not use seat belt. For the variable "collision type", it is clearly evident that a turn upside down traffic accident is the most dangerous situation and unbelted drivers would have an increased risk of fatality /serious injury (up to 29.79%). This could be reasonable as the risk of ejection for unbelted drivers would be high in rollover crashes. Also, from the results, "vehicle fall" is the second dangerous type of collision that threaten the life of driver who do not use seat belt. Finally, for "visibility" variable, in the two states, the probability of suffering a serious injury increases dramatically.

Table 5 Probabilities of suffering a serious injury, depending on the use of seat belt and the different variables of the conditioning factor. Processed by the Authors

Variables	N° cases	KSI Risk medium	Use of Seat-Belt (KSI Risk)			Comments
			Yes	No	Unknown	
Vehicle Type						
Car	241,092	2.13%	2.22%	14.01%	0.89%	Cars, van, all-terrain
Buses	4400	2.39%	1.39%	6.77%	0.63%	Minibus <=17 passengers, bus, bi-articulated bus
Trucks	12,359	2.99%	2.85%	17.10%	1.46%	Rigid Truck, Truck, articulated, articulated vehicle
Road-Type						
High speed road	47,125	2.39%	2.22%	20.81%	2.40%	Motorways, dual carriageway, others high speed roads
Medium speed road	61,064	5.53%	5.07%	38.04%	3.93%	Conventional roads of 1 and 2 lanes
Low speed road	144,350	0.68%	0.57%	5.84%	0.43%	Service ways, streets, neighbor roads, private roads
Others	5143	2.29%	2.13%	20.05%	1.35%	Bike roads or similar
Collision Type						
Collision	201,504	2.12%	2.19%	12.58%	0.89%	Frontal, side, and multiple collision, collision by range or against obstacle
Run Over	24,279	1.20%	1.20%	8.60%	0.58%	Running over a person or an animal
Turn upside down	8528	5.15%	4.89%	29.79%	2.65%	
Vehicle fall	3985	2.65%	2.84%	15.70%	0.92%	Fallen down a mountain or in the city
Others	19,433	2.49%	2.46%	15.21%	1.19%	Get out of the way or similar
Visibility						
Adequate visibility	178,839	3.60%	3.64%	22.84%	1.52%	
Inadequate Visibility	76,837	5.09%	5.44%	25.26%	1.94%	
Unknown	2175	0.64%	0.61%	4.29%	0.37%	

4.5 Probability of serious injury in a traffic accident based on the use of seat belt and the human factor

In this section four types of variables, including "infringement", "speed", "distraction" and "errors" were considered to estimate the probability of a serious and / or fatal injury (KSI risk) given seat belt use. From Table 6, comparison of the estimated probabilities under seat belt use and non-use conditions shows that unbelted drivers significantly have higher risk of fatality or serious injury compared to belted drivers.

For the variable "infringement", unbelted drivers making reckless maneuvers or failing to comply with traffic signal have higher risk of fatality or injury. That's because these action generally lead to serious accident (e.g., side accidents). In the specific case of speed infringements, the drivers committed any speed infringements have higher risk probabilities than drivers who complied with speed limits. Especially driving over the speed limit is the most dangerous situation if the driver does not use the seat belt. In the case of the "distraction" and "error", drivers who drive distractedly or with errors will have higher risk of fatality than drivers who drive without distraction or mistakes. However, such risk would be increased significantly if the driver does not use the seat belt.

4.6 Probability of serious injury in a traffic accident based on the use of seat belt and the most influential variables in driver injury severity

As explained in the section 4.2, seat belt use, distraction and road type, are the variables most directly related to the driver injury severity. Hence, this section presents a simultaneous analysis of these influential factors summarized in Table 7. Starting with the "high speed road" state from "road type" variable, it can be observed that the higher risk of fatality (which is associated with probability of 48.52%) occurs when a driver has a distraction, does not use the seat belt and is driving in a high speed road, like a motorway or similar (please observe the comments in Table 7). As 48.52% is the highest value in this study, therefore this would be one of the most dangerous situation in a traffic accident. In contrast, when the driver is unbelted, has a distraction and is deriving on low speed road, the risk of fatality or serious injury would be as low as 25%. This implies that the lower the driving speed the lower the impact speed which can reduce the risk of severe injury.

5 Conclusions

This study aimed to analyze the driver's injury severity as a function of seat belt use and other human factors.

Table 6 Probabilities of suffering a serious injury, depending on the use of seat belt and the different variables of the human factor. Processed by the Authors

Variables	N° cases	KSI Risk medium	Use of Seat-Belt (KSI Risk)			Comments
			Yes	No	Unknown	
Infringement						
No infringement	82,450	2.58%	2.47%	15.32%	0.98%	
Disrespect traffic signals	16,623	2.37%	2.46%	16.53%	0.94%	Not respect the stop, yield, traffic light and others priority of step signals
Disrespect traffic norm	14,712	2.00%	2.04%	14.87%	0.85%	Not respect the indications of a traffic agent, crosswalk and similar
Reckless maneuvers	30,860	3.49%	3.49%	22.47%	1.48%	Like wrong way, zigzag circulation, wrong reverse circulation, dangerous overtake, stop without just cause, park in a prohibited places, car races, not indicate or indicate wrongly a maneuver, incorrectly change direction
Unknown	113,206	1.51%	1.52%	8.43%	0.70%	
Speed						
No infraction	115,810	2.65%	2.87%	16.63%	1.03%	No infraction
Inadequate	11,483	3.62%	3.68%	21.59%	1.50%	Inadequate speed for road conditions
Exceeding	1520	4.08%	4.14%	24.67%	1.65%	Exceeding the established speed
Slow	132	3.44%	3.50%	21.25%	1.45%	Slow march / hindering circulation
Unknown	128,906	1.59%	1.60%	9.53%	0.73%	
Distraction						
No	67,111	1.84%	2.01%	8.87%	0.69%	
Yes	18,478	7.04%	6.66%	35.26%	3.79%	Like use: mobile phone, gps, hands-free, radio, dvd, smoke, others
Unknown	172,262	1.78%	1.79%	11.18%	0.79%	
Errors						
No	75,037	2.65%	2.65%	13.79%	0.97%	
Yes	51,143	3.81%	3.81%	25.63%	1.44%	Like don't see: a traffic sign, a vehicle, a walker, an obstacle, and others
Unknown	131,666	1.37%	1.37%	7.94%	0.67%	

Table 7 Probabilities of suffering a serious accident, depending on the use of seat belt and the most influential variables in driver injury severity. Processed by the Authors

Variables		Use of Seat-Belt (KSI Risk)			Comments
Road-Type	Distraction	Yes	No	Unknown	
High speed road	No	1.24%	8.73%	2.40%	Motorways, dual carriageway, others high speed roads + no distraction
	Yes	7.08%	48.52%	4.39%	Motorways, dual carriageway, others high speed roads + distraction like use: mobile phone, gps, hands-free, radio, dvd, smoke, others
	Unknown	1.99%	19.62%	2.18%	Motorways, dual carriageway, others high speed roads + unknown distraction
Medium speed road	No	3.53%	23.48%	2.64%	Conventional roads of 1 and 2 lanes + no distraction
	Yes	8.39%	38.56%	9.32%	Conventional roads of 1 and 2 lanes + distraction like use: mobile phone, gps, hands-free, radio, dvd, smoke, others
	Unknown	5.37%	42.85%	3.62%	Conventional roads of 1 and 2 lanes + unknown distraction
Low speed road	No	0.66%	3.57%	0.39%	Service ways, streets, neighbor roads, private roads + no distraction
	Yes	2.24%	25.13%	2.15%	Service ways, streets, neighbor roads, private roads + distraction like use: mobile phone, gps, hands-free, radio, dvd, smoke, others
	Unknown	0.48%	5.24%	0.35%	Service ways, streets, neighbor roads, private roads + unknown distraction
Others	No	2.64%	0.00%	0,00%	Bike roads or similar + no distraction
	Yes	6.49%	42.86%	0.99%	Bike roads or similar + distraction like use: mobile phone, gps, hands-free, radio, dvd, smoke, others
	Unknown	1.89%	21.97%	1.56%	Bike roads or similar + unknown distraction

Analysis of data showed that the high probability 'a priori' of serious and fatal injury, 12.92%, happens when the driver does not use the seat belt. In addition to "seat belt", the variables that are directly connected with the driver injury severity are "distraction" and "road type" which have a priori probability of risk of 7.04% and 5.53% respectively.

Regarding the demographic factor (sex and age), male drivers and the drivers under 25 and over 60 years old are more likely to suffer a serious and / or a fatal injury in a traffic accident. As with male drivers, being under 25 or over 60 and failing to wear the seat belt can increase the risk of death or severe injury to 15,41% and 15.58% respectively. For conditioning factor, the most worrisome case is when the driver is driving on a medium speed road without seat belt. Among the "collision type" and "vehicle type" States, the risk of sustaining a serious injury increases when the traffic accident is a "turn upside down" and the driver is driving a "truck". Also, when the drivers do not wear a seat belt, the probability of suffering a serious/fatal injury increases significantly, which is consistent with previous study [4]. In addition to "seat belt use", distraction and high-speed road are the most influential variables when the driver is not using the seat belt. These findings are consistent with previous works conducted around the world, implying that human factors along with other factors such as road and vehicle type are the major causes of road accident [28].

The majority of previous works on traffic safety adopted frequentist approaches (e.g., multivariate regression, logit models, etc.) to assess the impact of "seat belt use" in combination with other factors on injury severity. In this research, we compared the performance of the Logistic Regression and Bayesian Network using their estimated AUCs. In spite of the better performance of the Logistic Regression in terms of AUC (0.87 for Logistic Regression versus 0.73 for Bayesian Network), the Bayesian Network better estimates the number of accidents in which the drivers sustained severe injuries (4172 of 5610 for Bayesian Network versus 331 of 5610 for Logistic Regression). Moreover, the proposed Bayesian Network model provides the probabilities for injury severity while accounting for the interactions among the variables (especially influential variables such as seat belt use, distraction, and road type).

This study's findings can be used by transportation authorities and decision makers in order to establish effective policies. For example, it is recommended to; (a) use incentive and educational programs for young and old adults to promote their awareness about the seat belt use, (b) obligate the car manufacturers to equip the vehicles with seat belt interlock devices to prevent the car from being started unless front seat occupants have fastened their safety belts, (c) obligate the car manufactures to reduce the risk of distraction from electronic systems in vehicles, and (d) increase the security checkpoints and roadside cameras on high-speed roadways.

Febres *et al. European Transport Research Review*   (2020) 12:9

Page 12 of 12

Authors contributions
Conceptualization: SG-H and JDF; Data curation: SG-H, SH and JF; Funding acquisition: SG-H; Investigation: SG-H, JDF and JMG; Supervision: JMG and SG; Writing: JDF, SG-H and SH; Writing – review & editing: JRL-G and MÁM. All authors read and approved the final manuscript.

Availability of data and materials
The institution in Spain on charge to register the road traffic accidents and the injuries derived from them is the "Dirección General de Tráfico (DGT)". This institution provide the data base used in this study. The microdata files can be obtained through the following link: http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/ficheros-microdatos-accidentalidad/

Competing interests
We wish to confirm that there are no known conflicts of interest associated with this publication title "Influence of the seat-belt use on the severity of the injury in traffic accidents". We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed.

Author details
[1]Dep. of Chemistry and Exact Sciences, Universidad Técnica Particular de Loja, Loja, Ecuador. [2]Escuela Politécnica Superior, Universidad de Burgos, Burgos, Spain. [3]Departamento de Matemática Aplicada y Ciencias de la Computación, Universidad de Cantabria, Santander, Spain. [4]Instituto de Física de Cantabria, Consejo Superior de Investigaciones Científicas, Santander, Spain.

References
1. W.H.O. (2018). *Global status report on road safety 2018*: World Health Organization.
2. DGT. (2015). Principales Cifras de la Siniestralidad Vial. España . Retrieved from http://www.dgt.es/Galerias/prensa/2015/07/NP-principales-cifras-siniestralidad-2014.pdf.
3. Beck, L. F., Kresnow, M. J., & Bergen, G. (2019). Belief about seat belt use and seat belt wearing behavior among front and rear seat passengers in the United States. *J Saf Res, 68*, 81–88. https://doi.org/10.1016/j.jsr.2018.12.007.
4. Høye, A. (2016). How would increasing seat belt use affect the number of killed or seriously injured light vehicle occupants? *Accident Analysis & Prevention, 88*, 175-186. Doi: https://doi.org/10.1016/j.aap.2015.12.022.
5. Bedard, M., Guyatt, G. H., Stones, M. J., & Hirdes, J. P. (2002). The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. *Accid Anal Prev, 34*(6), 717–727. https://doi.org/10.1016/s0001-4575(01)00072-0.
6. Kim, J. M., Kim, S. C., Lee, K. H., Kim, H. J., Kim, H., Lee, S. W., et al. (2019). Preventive effects of seat belts on traumatic brain injury in motor vehicle collisions classified by crash severities and collision directions. *Eur J Trauma Emerg Surg*. https://doi.org/10.1007/s00068-019-01095-4.
7. Vallibhakara, S. A. O., Plitpolkarnpim, A., Suriyawongpaisal, P., & Thakkinstian, A. (2018). The nationwide surveillance of seat belt usage and encouraging factors of increasing the seat belt rate in Thailand: A road safety survey. *J Med Assoc Thail, 101*(6), 809–819.
8. Cardamone, A. S., Eboli, L., Forciniti, C., & Mazzulla, G. (2017). How usual behaviour can affect perceived drivers' psychological state while driving. *Transport, 32*(1), 13–22.
9. De Oña, J., De Oña, R., Eboli, L., Forciniti, C., & Mazzulla, G. (2014). How to identify the key factors that affect driver perception of accident risk. A comparison between Italian and Spanish driver behavior. *Accid Anal Prev, 73*, 225–235. https://doi.org/10.1016/j.aap.2014.09.020.
10. Febres, J. D., Mohamadi, F., Mariscal, M., Herrera, S., & García-Herrero, S. (2019). *The Role of Journey Purpose in Road Traffic Injuries: A Bayesian Network Approach Journal of Advanced Transportation*, 2019.
11. Kaplan, S., & Prato, C. G. (2012). Risk factors associated with bus accident severity in the United States: A generalized ordered logit model. *J Saf Res, 43*(3), 171–180.
12. Kim, J.-K., Ulfarsson, G. F., Kim, S., & Shankar, V. N. (2013). Driver-injury severity in single-vehicle crashes in California: A mixed logit analysis of heterogeneity due to age and gender. *Accid Anal Prev, 50*, 1073–1081.
13. Zong, F., Xu, H., & Zhang, H. (2013). Prediction for Traffic Accident Severity: Comparing the Bayesian Network and Regression Models. Mathematical Problems in Engineering. https://doi.org/10.1155/2013/475194
14. de Oña, J., Lopez, G., Mujalli, R., & Calvo, F. J. (2013). Analysis of traffic accidents on rural highways using latent class clustering and Bayesian networks. *Accid Anal Prev, 51*, 1–10. https://doi.org/10.1016/j.aap.2012.10.016.
15. Zong, F., Xu, H., & Zhang, H. (2013). *Prediction for traffic accident*. Severity: Comparing the Bayesian Network and Regression Models. Mathematical Problems in Engineering. https://doi.org/10.1155/2013/475194.
16. DGT. (2016). Anuario Estadístico de Accidentes.
17. Castillo, E., Gutiérrez, J. M., & Hadi, A. S. (1997). Sensitivity analysis in discrete Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 27*(4), 412–423.
18. Koller, & Friedman. (2009). *Probabilistic Graphical Models: Principles and Techniques*: MIT Press.
19. Neapolitan, R. E. (2004). *Learning bayesian networks* (Vol. 38): Pearson prentice hall upper Saddle River, NJ.
20. Murphy, K. (2001). The bayes net toolbox for matlab. *Computing science and statistics, 33*(2), 1024–1034.
21. Toolbox, f M. (2001). from https://github.com/bayesnet/bnt
22. Gutiérrez, J. M., Cano, R., Cofiño, A. S., & Sordo, C. M. (2004). *Redes probabilísticas y neuronales en las ciencias atmosféricas*: Ministerio de Medio Ambiente, Secretaría General Técnica.
23. Matlab. (2014). MATLAB para inteligencia artificial. from https://es.mathworks.com/
24. Kadilar, G. O. (2016). Effect of driver, roadway, collision, and vehicle characteristics on crash severity: A conditional logistic regression approach. *Int J Inj Control Saf Promot, 23*(2), 135–144.
25. Kononen, D. W., Flannagan, C. A., & Wang, S. C. (2011). Identification and validation of a logistic regression model for predicting serious injuries associated with motor vehicle crashes. *Accid Anal Prev, 43*(1), 112–122.
26. Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recogn Lett, 27*(8), 861–874.
27. Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36.
28. Yaacob, N. F. F., Rusli, N., & Bohari, S. N. (2018). A Review Analysis of Accident Factor on Road Accident Cases Using Haddon Matrix Approach. Paper presented at the proceedings of the second international conference on the future of ASEAN (ICoFA) 2017–volume 2.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.