Social, Economic, and Legislative Factors and Global Road Traffic Fatalities

Mohammad Reza Rahmanian Haghighi
Shiraz University of Medical Sciences
Mohammad Sayari
Shiraz University of Medical Sciences
Sulmaz Ghahramani
Shiraz University of Medical Sciences
Kamran Bagheri Lankarani (kblankarani@yahoo.com)
Shiraz University of Medical Sciences

Research article

Keywords: HDI, Education, Income, Life expectancy, GINI index, Road Traffic Fatalities

Posted Date: March 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-15580/v1

Version of Record: A version of this preprint was published on September 17th, 2020. See the published version at https://doi.org/10.1186/s12889-020-09491-x.

Social, Economic, and Legislative Factors and Global Road Traffic Fatalities

Mohammad Reza Rahmanian Haghighi[1]

Mohammad Sayari[1]

Sulmaz Ghahramani[1]

Kamran Bagheri Lankarani*[1]

Summary

Background: Road traffic fatalities (RTF) is the 8th cause of mortality around the world. At the end of the Decade of Action, it would be of utmost importance to revisit our knowledge on the determinants of RTF.

Methods: We used Road Safety Development Index (RSDI) which accounts for the interactions between system, human and products to assess the RTF in 115 and 113 countries in 2013 and 2016, respectively. To analyze data, three statistical procedures (linear regression, Classification and Regression Trees (CART), and Multivariate Adaptive Regression Splines (MARS)) were employed.

Results: CART has the best performance amongst all others followed by MARS for 2013 and 2016 data set with an R_2 around 0.83. Results show that any increase in human development index (HDI) was associated with RTF reduction. Comparing RTF data of 2013 and 2016, 8 countries

experienced a change of more than 30%. RTF change in these countries showed significant relationship with GINI index. Considering the three components of HDI, it is revealed that education explained most of RTF variation in CART model followed by income and life expectancy.

Conclusion: Policy makers can make provisions to reduce RTF in the long run by focusing on enhancing the three components of HDI, mainly education. However, there is a need to investigate the correlation among these three components with RTF with different time-trend procedures.

Keywords HDI; Education; Income; Life expectancy; GINI index; Road Traffic Fatalities

Introduction

According to recent World Health Organization (WHO) reports, road traffic fatality (RTF) was the 8th leading cause of death worldwide in 2016 (1). In 2013, approximately 1.4 million people were killed in Road Traffic Accidents (RTA). Although 46% of motorized vehicle are registered in high-income countries, only 10% of RTF has occurred in these countries. Middle- and low-income countries with 53 and 1 percent registered vehicles have 74 and 16 percent of RTF, respectively (2).

Infrastructures such as land-usage planning, road layout, designing for road function and vehicle safety are important factors in preventing RTF but too much focus has been devoted on road users' characteristics and behaviors(3). Most researches were conducted at micro levels on these factors, reporting the effect of road users' age, level of education, gender, drug and alcohol consumption, fatigue and other factors (3-5). Research on social aspects of road accidents at community level have shown that low socioeconomic status (SES) is associated with higher RTF as people living in deprived areas are more exposed to road traffic injuries and receive less efficient care (6, 7).

Confounding factors such as the use of older vehicles and vehicles with lower safety conditions might also exacerbate the situation (8). Vulnerable groups such as pedestrians, cyclists and motorcyclists which consist nearly half of the victims in road traffic accidents are also more concentrated in deprived neighborhoods (2).

The perspective about road traffic accidents has changed in the last century from focusing only on individual behaviors to the importance of supportive environment. It was recognized that individual users do not have the ability to reduce the risk of accidents and governments have to play a more active role in reducing road traffic accidents through legislative processes and various forms of regulation (9).

RTF should be approached as a social problem within the context of each individual community (10). In this context, culture and cultural attitude toward road traffic are important factors, which should be taken into consideration. This would affect the priorities in interventions and legislations which may differ between and within countries.

Contrary to earlier accident causation theories, there is now more emphasis on the complex interaction of causalities. In the system theory of accident causation, three main components interact: person, machine and environment (11). Haddon work recognizing the importance of the whole area of "mishaps" with all extra-rational conditions in accidents should be considered a turning point as he proposed a systematic approach rather than telescopic interventions in reducing RTF (12). Based on Haddon's approach, the major source of errors can be identified, and appropriate interventions can be utilized to reduce the risk of accidents or to mitigate the adverse effects of accidents including crash severity, fatalities and injuries (13).

Along the same line, Al-Haji (2007) proposed Road Safety Development Index (RSDI) as a comprehensive conceptual framework to approach road traffic accidents. In RSDI, three dimensions of performance, which include human, product and system performance are considered. Al-Haji considered the system as a whole with the interactions between system, human and products. Accordingly, "system performance" would be the effectiveness of interventions in improving car safety, safer roads and safer use. Table 1 summarizes factors based on RSDI (11).

Table 1. RSDI Factors (Goetsch 2011).

RSDI components	Factors
Human performance	• Safer road user's "behavior"
Product performance	• Percentage change of death trend • Personal risk "death per population" • Traffic risk "death per vehicle"
System performance	• Safer roads • Safer vehicles • Enforcement performance • Organizational performance • Socioeconomic performance

The current research aimed to investigate the main factors associated with RTF at global level considering social, economic, and legislative factors based on RSDI.

Methodology

Data sources and indices

We used the latest published data for safer road and mobility, safer vehicles, and safer road users in the "Global status report on road safety 2015" and "Global status report on road safety 2018", published by world health organization (WHO). The data of 180 countries were collected in 2013 and 2016 reports.

After list-wise deletion (eliminating samples with any missing values), 115 and 113 countries were selected from 2015 and2018 reports, respectively. Three categories were chosen (safer road and mobility, safer vehicles, and safer road users) from each country in the target years. The selected indices for each factor are shown in Table 2. Data on organizational performance (one of the RSDI components) was not available. "Death per population" was chosen for product performance. Data as "death per population" and "death per vehicles" factors were analyzed to obtain a more holistic view on the outcome.

Table 2.Safe System (factors and their indices).

SAFER ROADS AND MOBILITY	SAFER VEHICLES	SAFER ROAD USERS
Formal audits required for new road construction projects (2013)	Frontal impact standard (for both years)	National speed limit law (for both years)
Regular inspections of existing road infrastructure (2013)	Electronic stability control (for both years)	National drink–driving law (for both years)

Policies to promote walking or cycling (2013)	Pedestrian protection (for both years)	National motorcycle helmet law (for both years)
Policies to encourage investment in public transport (2013)	Motorcycle anti-lock braking system (2016)	National seat-belt law (for both years)
Policies to separate road users and protect VRUs (2013)		National child restraint law (for both years)
Audits or star rating required for new road Infrastructure (2016)		National law on mobile phone use while driving (for both years)
Design standards for the safety of pedestrians / Cyclists (2016)		National drug-driving law (for both years)
Inspections / star rating of existing roads (2016)		
Investments to upgrade high risk locations (2016)		
Policies & investment in urban public transport (2016)		

HDI (which consists of three indexes of life expectancy index, education index, and Gross National Income (GNI) index) (14), urban population (per 100,000), GINI index and unemployment rate were extracted from the United Nations Development Program and World

Bank data center as socioeconomic performance indices in target years. The data from world happiness and homicide rate (per 100,000 people) as indices of safer road user's "behavior" were considered too. We used the latest data available for each country. The RSDI components, factor and indices used, and the data sources are summarized in Table 3.

Table 3. RSDI components, Indices and data resources.

RSDI components	Selected factors	Indices	Data sources
Human performance	Safer road user's "behavior"	Happiness	World happiness: Trends, explanations and distribution (2013) and (2016) (15, 16)
		Homicides (per 100,000 people)	World Bank(17)
Product performance	Personal risk "death per population"	Mortality caused by road traffic injury (per 100,000 people)	Global status report on road safety 2015 and 2018 (2, 18)
System performance	Safer roads		Global status report on road safety 2015 and 2018 (2, 18)
	Safer vehicles		
	Enforcement performance		
		HDI	UNDP(19)

	Socioeconomic performance	Urban population (% of total)	World Bank(20)
		GINI index (World Bank estimate)	World Bank(21)
		Unemployment, total (% of total labor force) (modeled ILO estimate)	World Bank(22)

ILO, International Labour Organization

Statistical analysis

In Road Safety analysis, regression analysis such as linear regression models and Poisson regression has been the most conventional procedures to determine the factors affecting mortality rate. However, they have to meet certain assumptions and if we ignore the pre-defined assumptions, estimation would be inaccurate" (23).

In order to overcome these limitations, other procedures such as Classification and Regression Trees (CART) and Multivariate Adaptive Regression Splines (MARS) were applied in Road Safety analysis. They are considered decision tree procedures, and MARS can be viewed as the modified version of CART (24). Acciani et al. (2011) showed that MARS and CART perform

computationally well with small datasets (25). In this research, two non-parametric statistical procedures, CART and MARS, were employed to determine and classify factors associated with mortality rate. Furthermore, the prediction indices were compared to a stepwise multivariate linear regression (SMLR).

CART

CART is a tree-based procedure, which can be applied for continuous and discrete variables. CART algorithm uses all data sets to build child nodes by splitting the subsets of the entire predictor variables. The goal is to obtain a maximum homogeneous subset of the data with regard to the dependent variable. Especially, when the target variable is continuous, CART uses the least squared deviation (LSD) criterion to build an optimal tree. The split is chosen to maximize the value of LSD criterion function. The split procedure performs recursively into terminal nodes (26).

MARS model

MARS is a nonparametric data mining procedure, which combines the classical linear regression, the spline functions and the recursive partitioning. It uses a set of piecewise function called basis function (BF) to determine relationships between a set of independent variables and the target variable (27, 28). The general expression of MARS is defined as follow:

$$\hat{y} = \beta_0 + \sum_{m=1}^{M} \beta_m \, h_m(x)$$

Where $\hat{y}$ is the target variable predicted by the model, M is the number of selected BFs, β_0 is the constant term, β_m is the coefficient of the m-th BF and a $h_m(x)$ is one or more functions defined as follow:

$$h(x) = \max(0, X - t) = \begin{cases} X - t, & if\ t < X \\ 0, & otherwise \end{cases}$$

Or

$$h(x) = \max(0, t - X) = \begin{cases} t - X, & if\ X < t \\ 0, & otherwise \end{cases}.$$

The optimal model is chosen by Generalized Cross-Validation criterion (GCV) (29).

Prediction performance indices

In order to compare and assess the prediction performance of the models, the performance indices such as correlation coefficient ($r = cov(y_i, \hat{y}_i)/\sigma_{y_i}\sigma_{\hat{y}_i}$) , root mean-squared error ($RMSE = \sqrt{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2/n}$), mean absolute error ($MAE = \sum_{i=1}^{n}|\hat{y}_i - y_i|/n$), relative absolute error ($RAE = \sum_{i=1}^{n}|\hat{y}_i - y_i| / \sum_{i=1}^{n}|y_i - \bar{y}|$) and R-squared ($R^2 = 1 - (\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 / \sum_{i=1}^{n}(y_i - \bar{y})^2)$) were estimated, where y_i is the i −th actual value of the dependent variable, $\hat{y}_i$ is the corresponding predicted value and n is the number of observations (25).

Results

The data set

Data set includes HDI and its components (Education, Income and Life Expectancy), Happiness, GINI index, Urban Population, Unemployment, Homicide, Safe road, Safe vehicle and Safe user as independent variables. Using stated procedures, an initial analysis was performed found that the HDI was the most important variable amongst all independent variables (Additional file 1: Tables S1 to S6). Afterward, to identify the main factors related to RTF, HDI components have been considered in the analysis. Moreover, in order to assess the predictive capacity of the models, a

10-fold cross-validation was carried out. Table 4 presents common descriptive statistics for all the data set for 115 countries in 2013 and 113 countries in 2016.

Table 4.Common descriptive statistics of the variables.

	Minimum		Maximum		Mean		Std. Deviation	
Year	2013	2016	2013	2016	2013	2016	2013	2016
Mortality	2.8	2.7	36.2	35.9	16.561	16.122	9.219	9.2045
HDI	.340	.351	.946	.951	.7152	.7252	.156	.157
GINI	25.4	25.0	63.4	63.0	37.457	37.078	7.902	7.4880
Homicide	.183	2.905	74.28	7.526	6.489	5.351	9.944	1.1829
Happiness	2.936	.2835	7.693	82.842	5.460	6.217	1.105	10.684
Urban population	15.437	12.388	97.776	97.919	58.908	59.635	20.949	21.485
Unemployment	.3192	.524	28.996	26.55	8.425	7.629	6.410	5.747
Safe road	0	.50	5	5.00	3.16	3.668	1.405	1.131
Safe vehicle	0	3.00	3	7.00	1.07	6.257	1.387	.998
Safe user	2	0.00	7	4.00	6.33	1.442	.915	1.817
Education	.204	.212	.941	.940	.66396	.67466	.178088	.181785
Income	.287	.287	.975	.984	.69570	.70135	.174642	.176655
Life Expectancy	.468	.514	.975	.981	.80230	.81301	.125288	.117707

SMLR analysis

Stepwise Multivariate Linear Regression analysis was initiated with eleven independent variables. The entry and exit criteria (the P-value of F-statistic) were set to 0.05 and 0.1, respectively. In 2013, the final model consisted of five variables that explained the dependent variable (Mortality

rate). Specifically, Income, Safe vehicle, GINI index, Life Expectancy and Safe user were selected, represented in the following equation:

$$Mortality = 45.38 - 13.61 * Income - 1.77 * safe\ vehicle + 0.17 * GINI$$

$$-17.91 * Life\ Expectancy - 1.52 * safe\ user \quad (1)$$

As can be seen in Eq. (1), all of the coefficients carry the expected signs. The R-squared (R^2) value is the proportion of the variation of target variable which can be explained by its explanatory variables. The R^2 value for Eq. (1) is 0.746, indicating a good model fit to the data (Table 5). Furthermore, Table 5 illustrates partial influence of the selected variables. Table 5 shows that 65% of the variation of the dependent variable is explained by Income.

Table 5. Stepwise Multivariate Linear Regression: model summary (2013).

Model	R	R Square	Adjusted R Square	R Square Change
1	.804[a]	.646	.643	.646
2	.832[b]	.693	.688	.047
3	.849[c]	.720	.713	.027
4	.857[d]	.734	.724	.013
5	.864[e]	.746	.734	.012

a. Predictors: (Constant), Income 2013
b. Predictors: (Constant), Income 2013, safevehicle2013
c. Predictors: (Constant), Income 2013, safevehicle2013, GINI2013
d. Predictors: (Constant), Income 2013, safevehicle2013, GINI2013, Life Expectancy 2013
e. Predictors: (Constant), Income 2013, safevehicle2013, GINI2013, Life Expectancy 2013, safeuser2013

In 2016, the final model consists of three variables that explain mortality rate. Specifically, Income, GINI index and Life Expectancy were selected, represented in the following equation:

$$Mortality = 37.47 - 26.91 * Income + 0.35 * GINI - 19.16 * \text{Life Expectancy} \quad (2)$$

As can be seen in Eq. (2), all of the coefficients carry the expected signs. The R^2 value for Eq. (2) is 0.784, representing a good model fit to the data (Table 6). As can be seen in Table 6, 67% of the variation of the dependent variable is explained by Income.

Table 6. Stepwise Multivariate Linear Regression: model summary (2016).

Model	R	R Square	Adjusted R Square	R Square Change
1	.821[a]	.674	.671	.674
2	.877[b]	.769	.765	.095
3	.885[c]	.784	.778	.015

a. Predictors: (Constant), Income 2016
b. Predictors: (Constant), Income 2016, GINI2016
c. Predictors: (Constant), Income 2016, GINI2016, Life Expectancy 2016

CART

CART analysis was performed using the eleven independent variables. Furthermore, in order to obtain an optimal model, a 10-fold cross-validation was carried out. In 2013, the result of CART is a tree with 7 non-terminal nodes and 8 terminal nodes (Figure 1). From the 11 independent variables, CART used Income, Life Expectancy, Urban population, Unemployment and Homicide to build the optimal model.

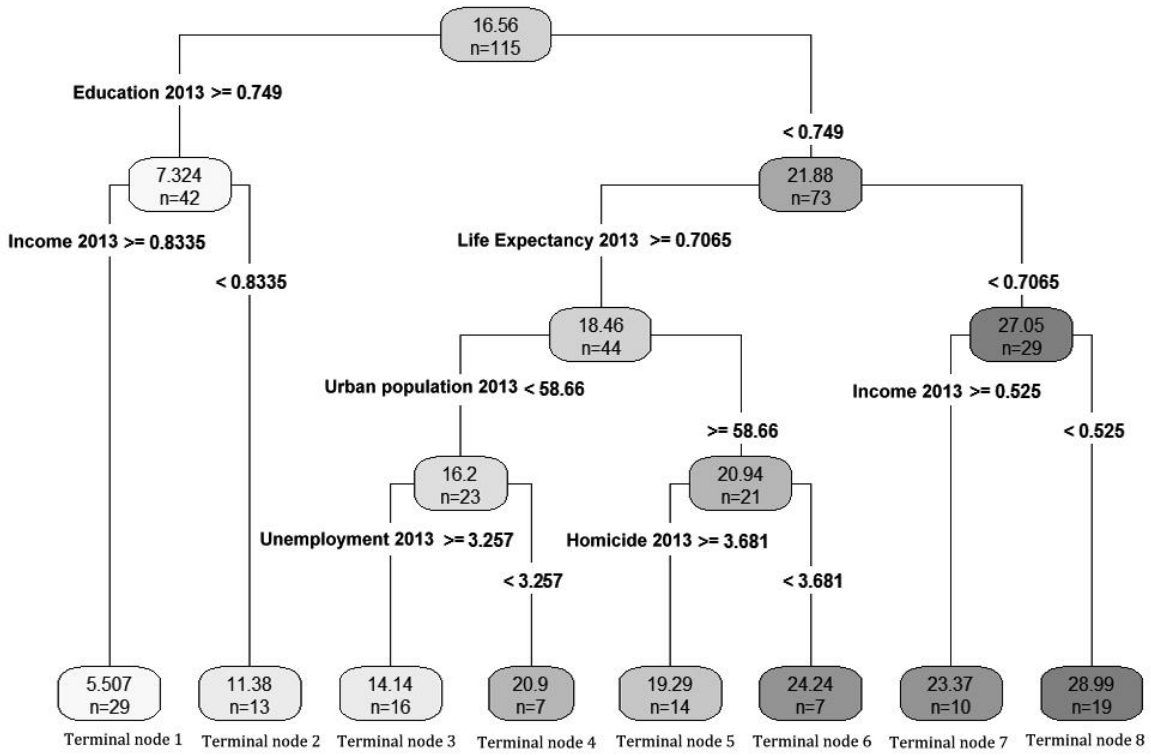

Figure 1. Optimal tree created by CART (2013).

In 2016, the result of CART is a tree with 6 non-terminal nodes and 7 terminal nodes (Figure 2). CART used Education, Income, Life Expectancy, Unemployment and Happiness to build the optimal model. The rules and the mean of Mortality rate from the final tree are available in Additional file 1: Tables S7 and S8.

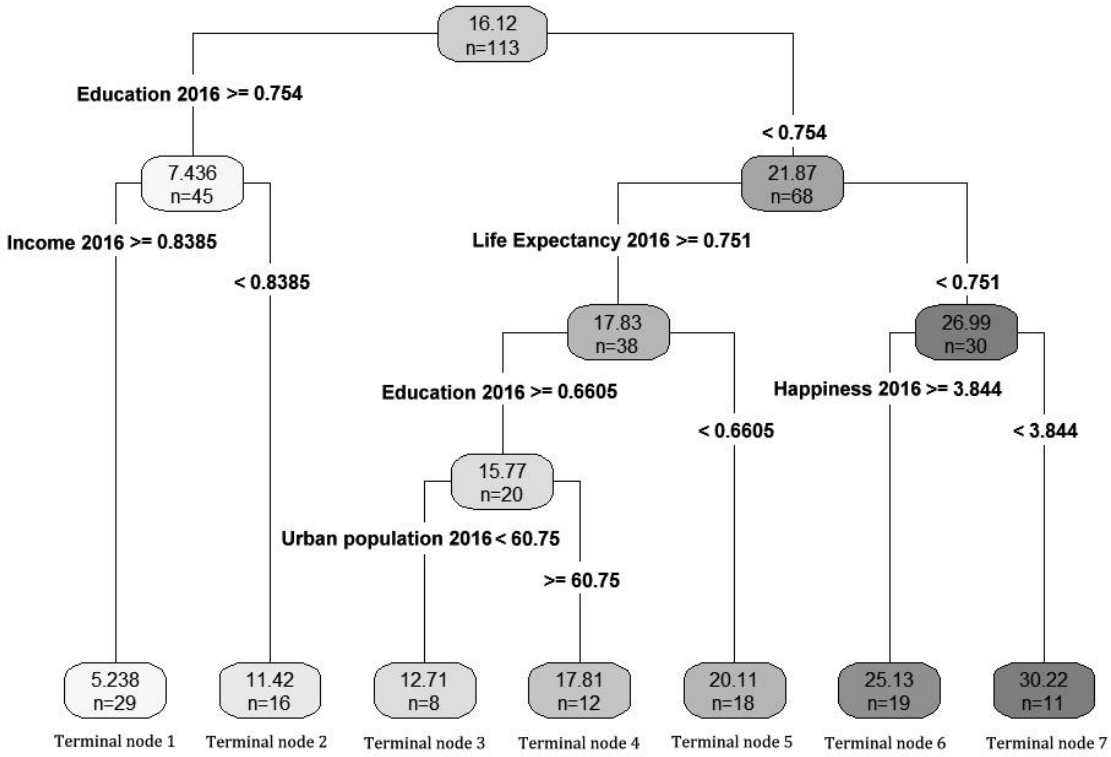

Figure 2. Optimal tree created by CART (2016).

MARS

A first-order MARS was carried out, so the basis functions of the models consist of linear splines. In order to create smaller models during the pruning step, the GCV criterion was replaced with 10-fold cross-validation. In Table 7, the basis functions and their coefficients are shown in detail for 2013 data set. Then, the MARS prediction function is represented in the following equation:

$$Y = 16.99 - 1.2 * BF1 + 0.09 * BF2 + 191.51 * BF3 - 437.70 * BF4$$

$$+ 225.58 * BF5 + 47.72 * BF6 + 36.05 * BF7 - 66.44 * BF8 - 36.37 * BF9$$

Table 7. Basis functions of the MARS and their coefficients (2013).

Variables	Basis Function	coefficients
(Intercept)		16.99
BF1	safevehicle2013	-1.2
BF2	h(urbanpopulation2013-38.979)	0.09
BF3	h(Education2013-0.583)	191.51209
BF4	h(Education2013-0.623)	-437.70114
BF5	h(Education2013-0.654)	225.58727
BF6	h(Income2013-0.59)	47.72427
BF7	h(0.745-Income2013)	36.05110
BF8	h(Income2013-0.745)	-66.43674
BF9	h(LifeExpectancy2013-0.613)	-36.36855

To illustrate the interpretation of MARS outcomes consider the BF2 given in Table 7

$$\mathrm{BF2} = \max(0, \text{urban population} - 38.979) = \begin{cases} \text{GINI} - 38.979, & if\ 38.979 < \text{urban population} \\ 0, & otherwise \end{cases}.$$

Therefore, if the GINI index of a country is 40.979, then the MARS model predicts the mortality rate increase by 0.18 (i.e., 0.09*(40.979–38.979)); otherwise, if the GINI index of a country is less than 38.979, then GINI index has no effect on mortality rate.

As shown in Table 7, the MARS model contains 9 basis functions. It can be observed that five variables play an important role in determining mortality rate. These variables include Safe vehicle, urban population, Education, Income, and Life Expectancy.

Table 8 shows the basis functions and their coefficients for 2016 data set. The MARS prediction function is represented in the follow equation:

$$Y = 20.9 - 0.29 * BF1 + 3.68 * BF2 - 30.82 * BF3 + 29.53 * BF4 - 52.24 * BF5.$$

Table 8. Basis functions of the MARS and their coefficients (2016).

Variables	Basis Function	coefficients
(Intercept)		20.9
BF1	h(45-GINI2016)	-0.29
BF2	h(5.121-Happiness2016)	3.68
BF3	h(Education2016-0.631)	-30.82
BF4	h(0.549-Income2016)	29.53
BF5	h(LifeExpectancy2016-0.865)	-52.24

As can be seen in Table 8, the MARS model contains 5 basis functions. It can be observed that five variables play an important role in determining mortality rate. These variables are GINI index, Happiness, Education, Income, and Life Expectancy. For more details on the impact of each basis function on mortality rate for each data set refer to Additional file 1: Tables S9 and S10.

Changes in road traffic fatality

Comparing RTF data of 2013 and 2016 indicates that more than 20% of the change occurred in 18 countries, and more than 30% in 8 countries. Table 9 shows countries with more than 30% of change in RTF with corresponding growth rate in GINI index and HDI between these years.

Table 9. Countries with more than 30% change in RTF with correspondent growth rate in GINI index and HDI.

Country	RTF rate 2013	RTF rate 2016	RTF Growth rate	GINI 2013	GINI 2016	GINI Growth rate	HDI 2013	HDI 2016	HDI Growth Rate

Iran	32.1	20.5	-36.14	37.4	40.0	6.95	0.784	0.796	1.53
Belarus	13.7	8.9	-35.04	26.6	25.3	-4.89	0.804	0.805	0.12
Bolivia	23.2	15.5	-33.19	47.6	44.6	-6.30	0.668	0.689	3.14
Macedonia	9.4	6.4	-31.91	36.2	35.6	-1.66	0.743	0.756	1.75
Kyrgyzstan	22.0	15.4	-30.00	28.8	26.8	-6.94	0.658	0.669	1.67
India	16.6	22.6	36.14	35.7	35.7	0.00	0.607	0.636	4.78
Turkey	8.9	12.3	38.20	40.2	41.9	4.23	0.771	0.787	2.07
Iceland	4.6	6.6	43.48	25.4	27.8	9.45	0.920	0.933	1.41

Prediction Performance of the models

The indices described in the previous section are shown in Table 10 for each model. The values of indices indicated a good fit to the data for each model. It can be observed that CART has the best performance among other methods.

Table 10. Prediction performance measures of the models.

	r		**RMSE**		**MAE**		**RAE**		**R_2**	
Year / Model	**2013**	**2016**	**2013**	**2016**	**2013**	**2016**	**2013**	**2016**	**2013**	**2016**
SMLR	0.86	0.88	4.62	4.27	3.21	3.11	0.40	0.39	0.75	0.78
CART	0.91	0.91	3.80	3.71	2.75	2.71	0.34	0.34	0.82	0.84
MARS	0.90	0.90	4.02	3.90	2.93	2.80	0.36	0.35	0.81	0.82

Variable importance measure (VIM) is one of the useful outputs from the CART and MARS models, which reflects the effect of predictor variables on the model. Table 11 indicates the relative

variable importance computed for the 9 independent variables in 2013 and 2016. There were considerable differences between the models regarding the importance of independent variables, which will be discussed in the next section.

Table 11. Importance of variables included in the CART and MARS model.

	Importance in CART		**Importance in MARS**	
Year / Variable	**2013**	**2016**	**2013**	**2016**
Education	21	21	100	100
Income	19	18	19.6	10.6
Life Expectancy	17	17	38.7	15
Safe vehicle	14	13	5	unused
Happiness	12	5	unused	38.5
Homicide	10	12	unused	unused
urban population	3	2	16.5	unused
unemployment	2	1	unused	unused
GINI	1	unused	unused	19.5
Safe user	unused	11	unused	unused
Safe road	unused	unused	unused	unused

Discussion

This study aims to investigate the main factors related to Road Traffic Fatality (RTF) around the world. Based on RSDI framework, we considered human performance, product performance and

system performance. In this context, two main factors which play a critical role in RTF in both 2013 and 2016 years were HDI and GINI indexes. Amongst the selected factors, the results showed that HDI was the main determinant, which was related to RTF. The results indicate that from 2013 to 2016, the mortality rate decreased along with GINI index, while HDI increased in this period (Additional file 1: Tables S3 to S5). In both years, HDI can explain more than 66 percent of RTF variations (Additional file 1: Tables S1 and S2).

To the best of our knowledge, there is scanty research similar in scope to the present study, and none as broad in data collection as our study. Most studies use panel data to determine factors which influence RTF. However, in these studies the number of indexes or the number of countries are limited. Al_Haji found a strong relationship between HDI and RSDI (Road Safety Development Index) and mentioned that by boosting the income of countries, more safer vehicles will be used and their investment on road infrastructure will be promoted (10, 30). Yannis (2014) investigated 27 European countries between 1975 to 2011 and find out that by increasing GDP (one of HDI components), RTF also increase in these countries (31). We could also see this trend in middle-income countries (32). Bester (2001) found that passenger car ownership, HDI and the percentage of other vehicles had an impact on road traffic death and HDI was able to explain 53% of the variation of road death (33). While in our study we found that HDI in SMLR was able to explain nearly 66% of the variation of mortality rate in 2013 and 72% in 2016. Melinder (2007) investigated 15 European countries and found a relationship amongst fatal death in road traffic accidents, wealth and religion. She concluded that economic level has an impact on RTF to some point and other factors subsequently play a role (34). A much broader study on 176 countries accounted for seven factors influencing road traffic injuries. These factors include income level of

a country and some other factors that could be categorized under the subheading of the "safer road users" (35).

Data for 2013 and 2016 reveals that RTF changed more than ±20% in 18 countries and more than ±30% in 8 countries (Additional file 1: Tables S11 to S16). In this regard, by considering different variables related to RTF, only GINI index seemed to have a relationship with RTF in these 8 countries. RTF decreased dramatically in 5 countries in these three years more than 30%, namely Iran, Belarus, Bolivia, Macedonia and Kyrgyzstan. Except Iran, other four countries in these years experienced a reduction in their GINI index. On the other hand, RTF in three countries increased more than 30%, namely India, Turkey and Iceland. Except India, other two countries in these years are faced with an increasing rate in GINI index. Two countries (Iran and India) were excluded from this analysis. In the case of Iran (however not limited to Iran), this could be explained by inconsistency between the reported data of the country and WHO estimation. However, it should be mentioned that from 2006 to 2012 Iran faced a decrease in the absolute number of deaths (approximately 27%) (36). On the other hand, the most recent available data for GINI index for India is attributed to the year 2011. Hence, as the most recent data have been used in this study (mentioned in methodology section), for GINI index in 2013 and 2016, there was no difference between these two years. These results show the importance of GINI index mainly for countries with more than 30% changes. However, there is a need to investigate the relation between RTF and GINI index in a long run.

In this study, all countries included for analysis - except Iceland- were developing countries with the HDI between 0.6 and 0.8 and we could see that the most radical changes occurred in these countries and it shows that developing countries might be more fragile to GINI index changes.

In this study, three different procedures for analyzing data in 115 countries were applied in 2013 and for 2016, 113 countries were assessed. As HDI is the main factor related to RTF, three components of HDI (Income, Education and Life expectancy) were considered alongside other factors. By using MLR, income is the main factor related to RTF in both years and it could explain 64% and 67% of RTF variation in 2013 and 2016, respectively. It is noticeable that when income is the main factor to explain RTF, GINI index also affects our output in both years. In 2013, safe vehicle, life expectancy and safe users were also related to RTF. However, these factors were not present in 2016.

Other two procedures have better explanation than MLR. CART and MARS models reveal that education is the main factor related to RTF among other factors. Variable importance table shows that for both years education, income and life expectancy (HDI components) are the main factors related to RTF respectively. For 2013, the cut-off point of education index is 0.749 and countries which have higher education index have lower RTF. In this context, countries with income index of more than 0.8335 have the lowest RTF rate (29 countries). On the other hand, for countries with education index of lower than 0.749, life expectancy has a significant relationship with RTF. Life expectancy cut-off point for these countries is 0.7065 and countries with higher life expectancy have lower rate of RTF. These results have been approximately confirmed by MARS model, where education index, income index and life expectancy index have the highest coefficient than other factors. This trend can also be seen in 2016. In this year, for CART model, education index cut-off point is 0.754 and countries with income index of higher than 0.8385 have the lowest RTF (29 countries). On the other hand, countries with education index of lower than 0.754, life expectancy with the cut-off point of more than 0.751 have lower RTF than other countries in this category. In MARS model, these three variables also have the highest impact on RTF.

Limitations

This is a cross-sectional study and there is a need for a longitudinal study to assess the impact of different factors on mortality rate. Although we tried to consider nearly 180 countries, missing data reduced our cases to approximately 115 countries. Finally, some important factors such as road network size of countries or their investment in road industry were not accessible for most countries in 2013 and 2016; hence, we did not consider these factors.

Conclusion

To sum up, although legislative factors, urban population and happiness can play an important role in the prediction of RTF for some countries, we found that HDI -- as signs of development -- can be the core predictors for 2013 and 2016 in most countries. Considering HDI components gives us a wider view about factors related to RTF. By comparing MLR, CART, and MARS models, it can be seen that CART provides a better explanation than others with R_2 of nearly 83% for both years. Based on CART results, education can play a central role on decreasing RTF in countries and by investing on education, countries could reduce their road fatality. For countries with a high rate of education index, income plays an important role too. This can be due to their investigation on road safety and using better vehicles. For countries with lower education, better medical care can reduce their vulnerability to RTF. It is noticeable that by considering HDI indexes, GINI index has a negligible impact on RTF.

Additional file

Additional file 1 : Microsoft Word file (doc) providing details of the supplementary tables. (DOCX 59 kb)

Abbreviations

RTF: Road traffic fatalities; RSDI: Road Safety Development Index; CART: Classification and Regression Trees; MARS: Multivariate Adaptive Regression Splines; HDI: Human Development Index; WHO: World Health Organization; RTA: Road Traffic Accidents; SES: socioeconomic status; SMLR: Stepwise Multivariate Linear Regression; LSD: Least Squared Deviation; BF: basis function; GCV: Generalized Cross-Validation criterion; RMSE: Root Mean-Squared Error; MAE: Mean Absolute Error; RAE: Relative Absolute Error

Acknowledgements

Not applicable.

Authors' Contributions

Conceptualization: KBL, MRRH and SG; Data curation: MRRH and SG; Formal analysis: MS, MRRH and SG; Investigation: KBL, MRRH, SG and MS; Methodology: MRRH, SG and MS; Supervision: KBL and SG; Validation: SG, KBL and MRRH; Visualization: MRRH and SG; Writing – original draft: KBL, MRRH, SG and MS; Writing – review & editing: KBL, MRRH, SG and MS. All authors read and approved the final manuscript.

Funding

None.

Availability of data and materials

The datasets used and/or analyzed during the study are available from the corresponding author on reasonable request.

Ethical approval and consent to participate

All procedures performed in this study were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Approval was obtained from the Ethics Committee of Shiraz University of Medical Sciences. Informed consent was obtained from all individual participants included in the study. Ethics committee approval: IR.SUMS.REC.1397.487

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] Health Policy Research Center, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran

Reference:

1. Organization WH. GLOBAL HEALTH ESTIMATES 2016 SUMMARY TABLES: GLOBAL DEATHS BY CAUSE, AGE AND SEX, 2000-2016. Geneva, Switzerland: World Health Organization; 2018.
2. Organization WH. Global status report on road safety 2015: World Health Organization; 2015.
3. Mohan D. Road traffic injury prevention training manual: World Health Organization; 2006.
4. Begg DJ, Langley JD, Williams SM. A longitudinal study of lifestyle factors as predictors of injuries and crashes among young adults. Accident Analysis & Prevention. 1999;31(1):1-11.
5. Hasselberg M, Vaez M, Laflamme L. Socioeconomic aspects of the circumstances and consequences of car crashes among young adults. Social science & medicine. 2005;60(2):287-95.
6. Hosking J, Ameratunga S, Exeter D, Stewart J, Bell A. Ethnic, socioeconomic and geographical inequalities in road traffic injury rates in the Auckland region. Australian and New Zealand journal of public health. 2013;37(2):162-7.
7. Klaitman SS, Solomonov E, Yaloz A, Biswas S. The Incidence of Road Traffic Crashes Among Young People Aged 15–20 Years: Differences in Behavior, Lifestyle and Sociodemographic Indices in the Galilee and the Golan. Frontiers in public health. 2018;6.
8. Whitlock G, Norton R, Clark T, Pledger M, Jackson R, MacMahon S. Motor vehicle driver injury and socioeconomic status: a cohort study with prospective and retrospective driver injuries. Journal of Epidemiology & Community Health. 2003;57(7):512-6.
9. Leveson N. A new accident model for engineering safer systems. Safety science. 2004;42(4):237-70.
10. Al-Haji G. Road safety development index: theory, philosophy and practice: Linköping University Electronic Press; 2007.
11. Goetsch DL. Occupational Safety and Health for Technologists, Engineers, and. 2011:32-50.
12. Haddon Jr W. The changing approach to the epidemiology, prevention, and amelioration of trauma: the transition to approaches etiologically rather than descriptively based. American journal of public health and the Nations health. 1968;58(8):1431-8.
13. Peden M, Scurfield R, Sleet D, Mohan D, Hyder AA, Jarawan E, et al. World report on road traffic injury prevention. World Health Organization Geneva; 2004.
14. Programme UND. Human development indices and indicators: 2018 Statistical update. United Nations Development Programme South Africa; 2018.
15. Helliwell JF, Wang S. World happiness: Trends, explanations and distribution. World happiness report. 2013;201(3):8-37.
16. Helliwell JF, Layard PR, Sachs J. World happiness report 2016 update: Sustainable Development Solutions Network; 2016.
17. Bank W. Homicides (per 100,000 people).
18. Organization WH. Global status report on road safety 2018: World Health Organization; 2018.
19. PROGRAMME UND. Human Development Data (1990-2017).
20. Bank W. Urban population (% of total).

21. Bank W. GINI index (World Bank estimate).
22. Bank W. Unemployment, total (% of total labor force) (modeled ILO estimate).
23. Chang L-Y, Chen W-C. Data mining of tree-based models to analyze freeway accident frequency. Journal of safety research. 2005;36(4):365-75.
24. Hastie T, Tibshirani R. J., F.(2001). The Elements of Statistical Learning–Data Mining, Inference and Prediction. Springer; 2001.
25. Acciani C, Fucilli V, Sardaro R. Data mining in real estate appraisal: a model tree and multivariate adaptive regression spline approach. Aestimum. 2011:27-45.
26. Breiman L. Classification and regression trees: Routledge; 2017.
27. Friedman JH. Multivariate adaptive regression splines. The annals of statistics. 1991:1-67.
28. Trevor H, Robert T, JH F. The elements of statistical learning: data mining, inference, and prediction. New York, NY: Springer; 2009.
29. Craven P, Wahba G. Smoothing noisy data with spline functions. Numerische mathematik. 1978;31(4):377-403.
30. Al Haji G. Towards a Road Safety Development Index (RSDI): Development of an International Index to Measure Road Safety Performance: Linköping University Electronic Press; 2005.
31. Yannis G, Papadimitriou E, Folla KJSs. Effect of GDP changes on road traffic fatalities. 2014;63:42-9.
32. Ali Q, Yaseen MR, Khan MTIJTrpAp, practice. The causality of road traffic fatalities with its determinants in upper middle income countries: A continent-wide comparison. 2019;119:301-12.
33. Bester CJ. Explaining national road fatalities. Accident Analysis & Prevention. 2001;33(5):663-72.
34. Melinder K. Socio-cultural characteristics of high versus low risk societies regarding road traffic safety. Safety science. 2007;45(3):397-414.
35. Jafari SA, Jahandideh S, Jahandideh M, Asadabadi EB. Prediction of road traffic death rate using neural networks optimised by genetic algorithm. International journal of injury control and safety promotion. 2015;22(2):153-7.
36. Lankarani KB, Sarikhani Y, Heydari ST, Joulaie H, Maharlouei N, Peimani P, et al. Traffic accidents in Iran, a decade of progress but still challenges ahead. 2014;28:96.

Figures

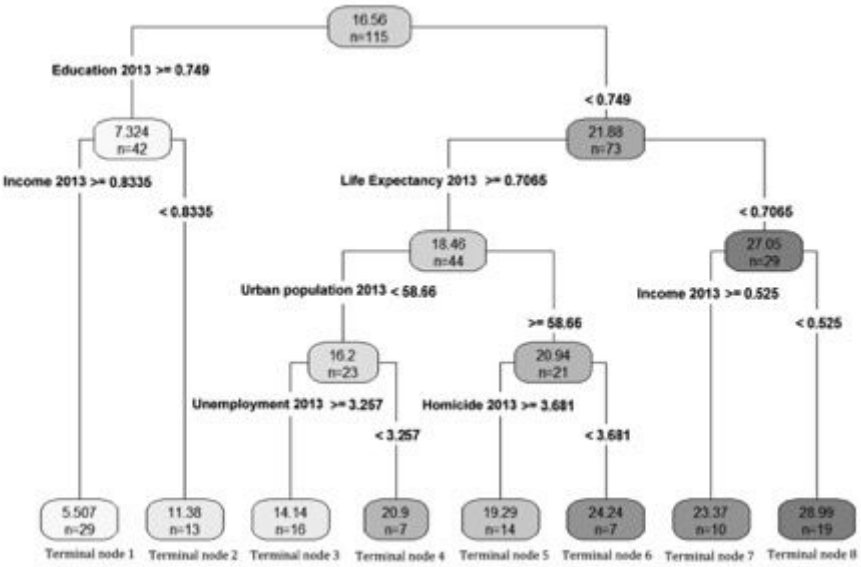

Figure 1

Optimal tree created by CART (2013).

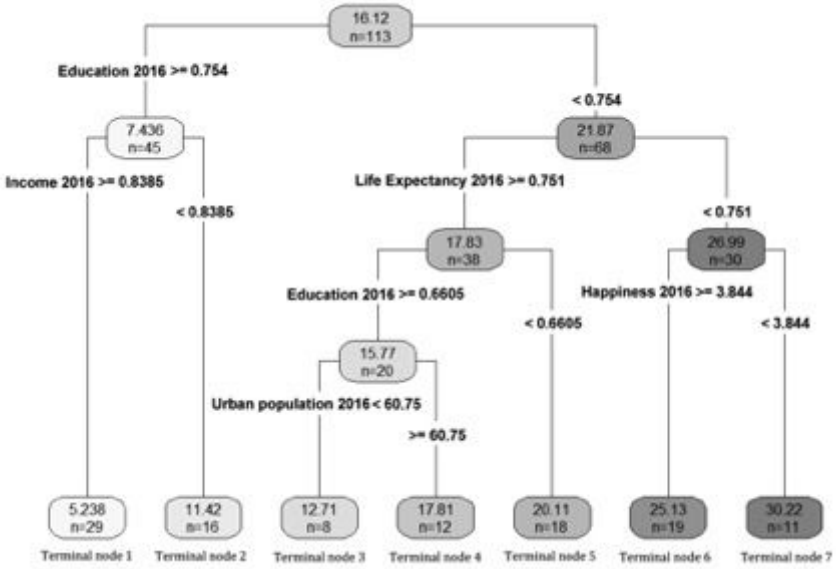

Figure 2

Optimal tree created by CART (2016).