

Road traffic crash related injured and fatal victims in Karachi from 2007 to 2014: A time-series analysis

Rashid Jooma,¹ Masood Ali Shaikh²

Abstract

Injuries and deaths due to road traffic crashes (RTC) are major public health concern. The Road Traffic Injury Research and Prevention Center, collects RTC data on injuries and fatalities from five emergency departments in Karachi. Data generating process (DGP) for RTC from 2007 to 2014, for monthly number of fatal and injured victims were analyzed using autoregressive integrated moving average and vector auto regression, time series models. Results provide strong evidence that the DGP for the current levels of the number of fatalities and injured owing to RTCs are significantly influenced by the own past history of the two series. The analysis with the impulse-response function also indicated that there is a slight seasonality pattern in the number of injured and fatalities. The similar behaviour and association of the two variables suggest that certain conditions e.g. road conditions, weather, volume of vehicles, and accidents might be persistent in time in Karachi.

Keywords: Road Traffic Fatalities, Injuries, Karachi, Pakistan.

Introduction

Deaths and injuries due to road traffic crashes (RTCs) are a serious global public health concern. In 2013, RTCs killed 1.25 million people globally, with ninety percent of RTC mortality brunt was borne by low and middle countries like Pakistan.¹ Globally, RTC fatality rate has been reported as 17.4 per 100,000 population¹. Pakistan is categorized as part of the Eastern Mediterranean region, in which RTC fatality rate was reported as 19.9 per 100,000 population, which is second highest, after the African region's rate of 26.6, and compares unfavourably with the rate of 17.4, globally.¹ Among all age groups, RTCs are the ninth leading cause of death, while among the 15-29 age group it is the leading cause of death globally.¹

The World Health Organization (WHO) reported 7,636 RTC

fatalities in the year 2013 for Pakistan. However, for Pakistan, WHO's estimated road traffic fatalities for the same year were 25,781, with a 95% confidence interval of 20,979 to 30,582. And estimated rate per 100,000 population to be 14.2, for the year 2013. Numerous studies have been published using Karachi's Road Traffic Injury Research and Prevention Center' (RTIRPC) data from the years 2007 to 2014, describing the correlates, and morbidity and mortality burden of RTCs.²⁻⁹ However, time series analysis has not been conducted using RTIRPC's unique dataset. Which entails deciphering trend, seasonal variation and structure of data in terms of autocorrelation.

In this study we provide results from the time series analysis of RTIRPC's data on the monthly numbers of fatalities and injured individuals for the eight-year period from 2007 to 2014.

Methodology and Results

Established in September 2006, Road Traffic Injury Research and Prevention Center (RTIRPC) has been collecting road traffic crash (RTC) injuries and fatalities data in Karachi. RTIRPC is a public-private partnership in Karachi, financed by community and corporate donors to study road traffic crashes in the most populous city of the country. This RTC surveillance project is a collaboration between Jinnah Post Graduate Medical Centre (JPMC), Aga Khan University Hospital, and NED University of Engineering & Technology, and was notified by the Ministry of Health, Government of Pakistan.

RTIRPC collects RTC data from the emergency departments of five major public and private hospitals in Karachi: Jinnah Postgraduate Medical Center, Civil Hospital Karachi, Liaquat National Hospital, Abbasi

Table-1: Results of Dickey-Fuller tests for the presence of unit root in fatal and non-fatal time series.

	Test Statistic	Interpolated Dickey-Fuller		
		1% Critical Value	5% Critical Value	10% Critical Value
Fatal	-6.155	-3.517	-2.894	-2.582
Non-Fatal	-6.093	-3.517	-2.894	-2.582

MacKinnon approximate p-value for both fatal as well as non-fatal time series = <0.0001.

.....
¹Section of Neurosurgery, Aga Khan University & Principal Investigator - Road Traffic Injury Research & Prevention Center, Karachi, ²Independent Consultant, Karachi, Pakistan.

Correspondence: Masood Ali Shaikh. Email: masoodali1@yahoo.com

Table-2: Coefficients and their statistical significance for the fatal ARIMA, non-fatal/injured ARIMA, and the VAR models.

Variable	Fatal_ARIMA	Non-Fatal_ARIMA	VAR
Fatal			
NonFatal	0.02***		
Fatal			
L1.			0.42***
L12.			0.05
NonFatal			
L1.			0.00
L12.			0.02***
_cons	42.41***		
ARMA			
AR			
L1.	1.34***	0.49***	
L2.	-0.38**		
MA			
L1.	-0.83***		
Non Fatal			
Fatal			
--.		7.10***	
L1.			-1.29
L12.			3.30*
Non Fatal			
L1.			0.55***
L12.			0.36***
_cons		1949.94***	

Asterisk (*) denotes the statistical significance of the coefficients:

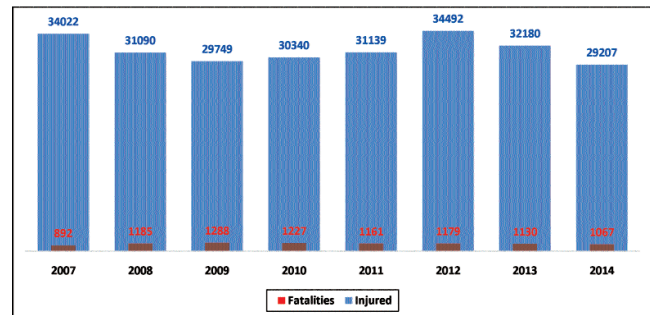
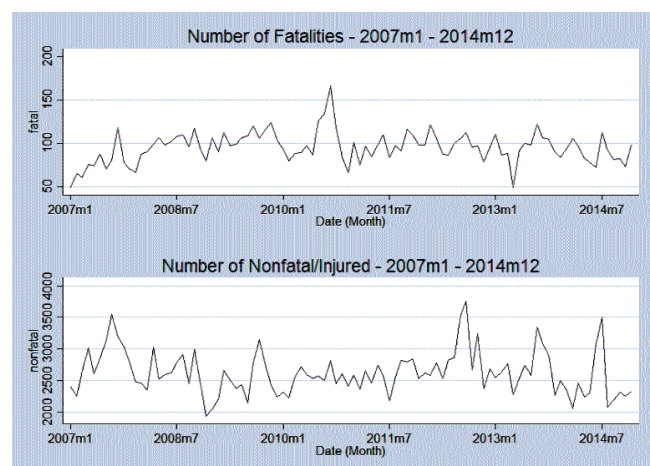
* p<0.1; ** p<0.05; *** p<0.01.

L= Lag ('L1' implies Lag number 1 and 'L12' implies Lag number 12).

AR= Autoregressive.

MA= Moving Average.

Shaheed Hospital and the Aga Khan University Hospital. Data pertaining to all injured and fatal victims of RTCs brought/admitted to these hospitals are recorded either directly from the victims and/or accompanying persons and then collated at the RTIRPC project office located at Jinnah Postgraduate Medical Center. Additional information is also obtained — where needed — from ambulance drivers, police and eye-witnesses to ensure accurate recording of all pertinent information. Data collectors are available at these five sites on 24/7 basis and work in three shifts to ensure that information pertaining to all RTCs victims gets properly collected and recorded. We analyzed the data generating process for RTC data in terms of monthly number of individuals who were injured or had died from January 2007 to December 2014, using STATA 14. The two monthly series of 95 numbers of both, fatal and non-fatal victims were analyzed for time trends using Autoregressive integrated moving average model (ARIMA), and Vector auto

**Figure-1:** Yearly number of fatalities and injured victims of Road Traffic Crashes from 2007 to 2014 based on Road Traffic Injury Research and Prevention Center data in Karachi.

*'2007m1' implies January 2007, '2008m7' implies June 2008, and so on.

Figure-2: Plot of fatalities and injured by month from 2007 to 2014.

regression (VAR), time series analysis models.

The RTC data in terms of yearly number of fatalities and injured individuals from January 2007 to December 2014, are shown in Figure-1. For years 2007, 2008, and 2009 there were 693, 222, and 14 records respectively, for which codes for disposal status from emergency departments were not assigned; hence, were not included in the analysis. The 95 monthly numbers of both fatal and non-fatal i.e. injured victims were used for time series analysis, as two series of time. Based on the results from the Dickey-Fuller test, we could not accept the null hypothesis of the presence of unit roots for any of the two variables. Table 1 shows the results from these tests along with the p-values for both time series of the number of fatal and nonfatal/injured. Therefore, we concluded that both series were stationary. Figure-2 shows the plot for the two variables as a visual support for our conclusion on stationarity. No lags were added to the Dickey-Fuller tests

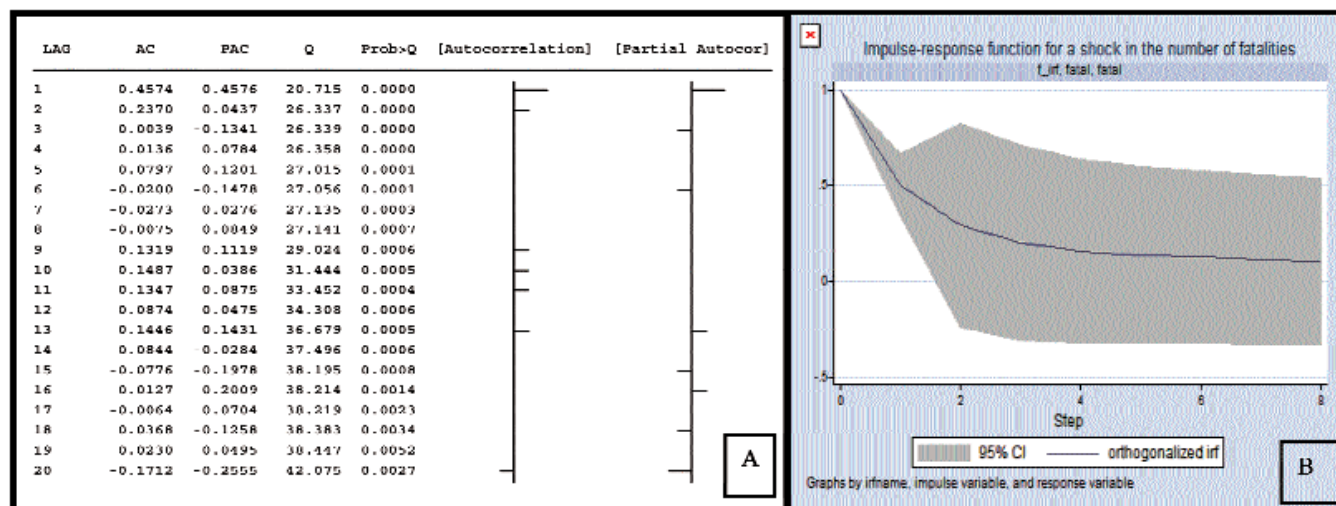


Figure-3: Correlogram for fatalities time series (A), and the impulse response function graph (B).

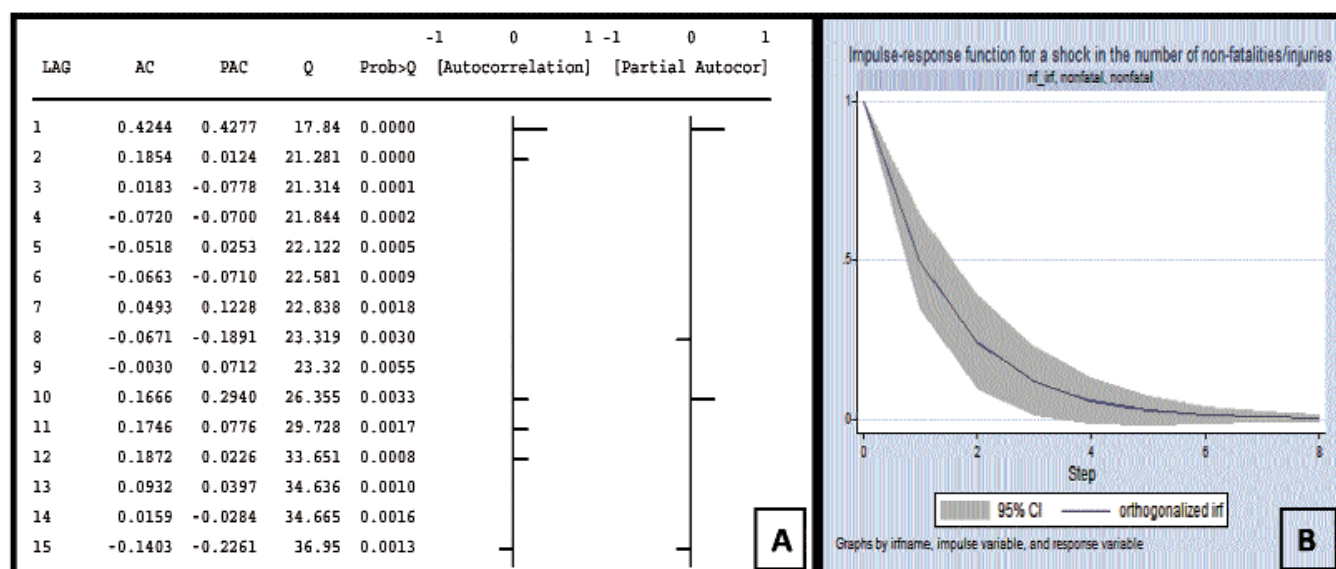


Figure-4: Correlogram for non-fatalities time series (A), and the impulse response function graph (B).

because the error terms for the auxiliary regressions for the test did not have serial correlation.

Autoregressive integrated moving average (ARIMA) models: We first modeled fatalities time series, and the correlogram for the series is presented as Figure 3-A. The correlogram indicates that we could fit an ARMA (1,1) model. However, that model did not fit well because there seemed to be a common factor (for AR1 and MA1) in the model. We decided to fit a series of ARIMA models considering the first and second lags. We also included the same models with nonfatal/injured as an explanatory

variable to detect whether there could be association on the occurrence of the two kinds of RTA victims. All the models were fitted for the full sample to compute the Akaike (AIC) and Schwarz (BIC) information criteria. As an additional tool for deciding about the best model specification, we also fitted the models for a subsample (2007m1, 2013m12) to compute the minimum square errors (MSE) for the in and out of sample forecasts. The lower the information criteria and the MSE the better the model. Based on the AIC and BIC the three best models include nonfatal as an explanatory variable and correspond to the AR1, MA2, and ARMA (2,1) structures.

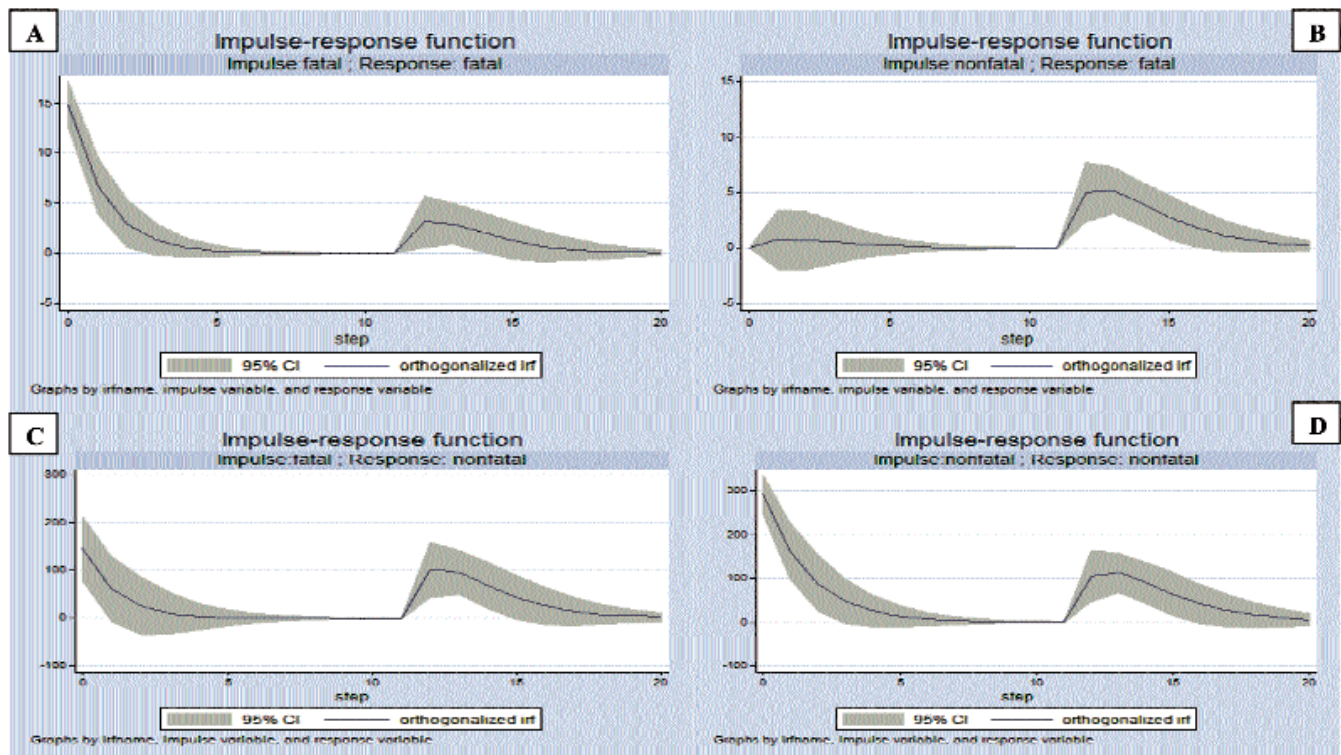


Figure-5: Impulse-response functions for the analysis of impact of shocks in the system: Figures A & B - impact of a shock in the number of fatalities or injured on the number of fatalities. Figures C & D - impact on the number of non-fatalities.

By looking at the MSE we observed that the 'latest' performs better in terms of in and out of sample forecasts, so we decided to select that model because it also had more structure. The results above have two main implications: The ARMA (2,1) structure indicates that the DGP for the number of fatalities tend to be strongly related to the level of the series in previous periods. The main impact occurs for the first two periods and then decays close to zero. This can be seen in the graph for the impulse-response function in Figure 3-B. The second implication is that the positive significance of 'nonfatal' implies that there seems to be a contemporaneous association between the number of fatalities and injured. To better address this association, a Vector autoregression (VAR) model would be appropriate. We next modeled number of non-fatalities/injuries time series, and the correlogram for the series is presented as Figure 4-A. The correlogram, which also indicates an ARMA (1,1) structure, but similar to the case for fatalities, the coefficient estimates for the AR1 and MA1 terms are not both individually significantly different than zero; only AR1 is. After an inspection of the potential ARIMA and ARIMAX (when an ARIMA model includes explanatory variables) models for nonfatal we were down to three specifications

that we decided to evaluate using the information criteria and the MSE. The information criteria and the MSE favour the AR1 model with fatal as an explanatory variable. The model's impulse-response function is presented as Figure-4-B. The results indicate that the association with the previous period number of nonfatal/injured is smaller compared to the model for 'fatal', and the effect of a shock in the current period decays quicker in this case.

Vector autoregression (VAR) model: the observed association between the two series was next analyzed using the VAR model. We selected the VAR model with lags 1 and 12. This model was selected after a comparison with a couple of other models i.e. VAR (1) and VAR (2). The output in table 2 - coefficients along with their statistical significance specified by asterisks - shows that the series for the number of fatalities is associated to its first lag and the twelfth lag of the number of non-fatal accidents. On the other hand, the number of non-fatal accidents is mainly related to its own past. Figure-5 A & B show the impulse-response functions to analyze the impact of shocks in the system. This plot shows the impact of a shock in the number of fatalities or injured on the number of fatalities. And confirms that the number of fatalities is mainly associated to the levels of the same kind of

fatalities in previous periods. This persistence in the DGP for the series could be indicating the presence of certain conditions that remain similar in time (e.g. road conditions and/or volume of vehicles). The slight move after twelve periods show the presence of some seasonality that could be related to weather conditions or seasonal traffic patterns. Figure-5 C & D show the impulse-response functions for the impact on the number of injured. The major association of the series is also in this case with its own past values, although, shocks in the number of fatalities seems to be slightly more relevant in the DGP for the number of injured compared to the crossed effect in the previous plot. The slight seasonality effect is also present in this case.

Discussion

Injuries and deaths due to RTCs in Karachi, based on RTIRPC data have been extensively reported in several descriptive and analytical epidemiological studies.²⁻⁹ However, this is the first study in Pakistan on time series analysis of the number of injured and fatalities due to RTCs. This study presents the models used to analyze the data generating process for the two, time series of the monthly number of fatalities and injuries from 2007 to 2014. All the models fitted in this analysis required checking for stationarity of the series; which was checked by performing Dickey-Fuller unit root tests. We then worked with ARIMA models to describe each of the two series in terms of their own history, but including number of injured as a repressor for number of fatalities and vice versa. We then finished the analysis by using multivariate VAR model, considering potential contemporary effects between the DGP for the two series. The different approaches used in this analysis provide strong evidence that the DGP for the current levels of the number of fatalities and injured owing to RTCs are significantly influenced by the own past history of the two series. The analysis with the impulse-response function also indicated that there is a slight seasonality pattern in the number of injuries and fatalities. The similar behaviour and association of the two variables seem to indicate that certain conditions e.g. road conditions, weather, volume of vehicles, and accidents might be persistent in time for the traffic areas where the data were collected. Hence the past levels of two series is relevant to explain the current level on the number of fatalities and injured individuals

generated by RTCs. It is fair to assume that population in Karachi has grown over the study period; we may conclude that the relative number of fatalities and injured may be lower. Since our results determined that the two series are stationary, so there is about the same average of number of fatalities and injured throughout the whole sample. Which would support the conclusion that the relative number of fatalities and injured — with respect to population — may have decreased

The RTIRPC, in its decennial year now, warrants extending the time series analysis of the number of RTCs morbidity and mortality data by forecasting future number of fatalities and injured to more effectively utilize the collected data for better informed road safety and public health planning in the largest metropolitan city of Pakistan.

Disclaimer: None to declare.

Conflict of Interest: The first author being incharge of the Karachi's Road Traffic Injury Research and Prevention Center (RTIRPC) has signed the IRB statement declaring that ethical approval has been granted for the conduct of time series analysis on RTIRPC's data.

Funding Disclosure: None to declare.

References

1. World Health Organization. Global status report on road safety 2015. WHO, Geneva: 2015.
2. Jooma R, Shaikh MA. Descriptive epidemiology of Karachi road traffic crash mortality from 2007 to 2014. *J Pak Med Assoc* 2016; 66: 1475-80.
3. Jooma R, Shaikh MA. Epidemiology of Karachi road traffic crash mortality in 2013. *J Pak Med Assoc* 2015; 65:548-51.
4. Mehmood A, Razzak JA, Mir MU, Jooma R. Differences in risk-adjusted outcome of road traffic injuries in urban tertiary care centers of Pakistan. *J Pak Med Assoc* 2015; 65: 984-9.
5. Mehmood A, Khan IQ, Mir MU, Moin A, Jooma R. Vulnerable road users are at greater risk during Ramadan -- results from road traffic surveillance data. *J Pak Med Assoc* 2015; 65: 287-91.
6. Khan UR, Bhatti JA, Shamim MS, Zia N, Razzak JA, Jooma R. Clothing-related motorcycle injuries in Pakistan: findings from a surveillance study. *Int J Inj Contr Saf Promot* 2015; 22: 308-13.
7. Razzak JA, Shamim MS, Mehmood A, Hussain SA, Ali MS, Jooma R. A successful model of Road Traffic Injury surveillance in a developing country: process and lessons learnt. *BMC Public Health* 2012; 12: 357.
8. Shamim S, Razzak JA, Jooma R, Khan U. Initial results of Pakistan's first road traffic injury surveillance project. *Int J Inj Contr Saf Promot* 2011; 18: 213-7.
9. Lateef MU. Spatial patterns monitoring of road traffic injuries in Karachi metropolis. *Int J Inj Contr Saf Promot* 2011; 18: 97-105.