Study on the effect of road infrastructure, socio-economic and demographic features on road crashes in Bangladesh

A Dissertation Submitted in Partial Fulfillment of the Requirements

For the Bachelor of Science Degree in

Civil & Environmental Engineering

By

Md. Hishamur Rahman (085432)

Mohammed (085437)

September, 2012

Islamic University of Technology (IUT)

A Subsidiary Organ of OIC.Dhaka,Bangladesh

The dissertation entitled "Study on the effect of road infrastructure, socio-economic and demographic features on road crashes in Bangladesh", by Mohammed, Md. Hishamur Rahman has been approved in partial fulfillment of the requirements for the Bachelor of Science Degree in Civil & Environmental Engineering.

Supervisor, Dr. Shakil Mohammad Rifaat

Department of Civil & Environmental Engineering

Author, Md.Hishamur Rahman

Department of CEE

Author, Mohammed

Department of CEE

Abstract

A statistical analysis was conducted for road traffic accidents and associated casualties in Bangladesh. This was undertaken in order to assist the policy-makers to take appropriate steps to reduce the road traffic accidents and the associated casualties. District wise data (collected from Bangladesh government publications) were explored, analyzed and modeled statistically. The modeling was done by negative binomial model. Our study considered the effect of road length (both paved and unpaved), road infrastructure and several socio economic characteristics on crash frequency. A model of accident prediction was developed with the collected data.

There are several significant findings in our research work. We founded that crash frequency increases due to the increase in paved road length. On the contrary, due to increase in unpaved road length the crash frequency decreases. Again it was seen that crash frequency increases due to the increase of road infrastructure. The increase in population also affects the crash frequency. It was seen that due to increase in total male population, the number of crashes increase. But for the increase in female population, crash frequency decreases. From economic point of view it was seen that, due to the increase of active male population, number of accident reduces. While for the increase in active female population, number of accident increases. Again for several economic stages (temporary insolvency, permanent insolvency, equivalent expenses etc) it was seen that crash frequency increases due to the increase of these economic stages.

Acknowledgements

First and foremost, we would like to state our indebtedness and genuine appreciation to our thesis supervisor, Dr. Shakil Mohammad Rifaat for his constant support and guidance to execute this research work. It is a pleasure to work with him and his consistent inspiration increases our productivity and the quality of research. Without his continuous and active attention, it would not be possible to complete this dissertation.

We would also like to acknowledge Roads & Highways Department (RHD), Bangladesh Statistical bureau (BSB), Bangladesh Road Transport Authority (BRTA) and Local Government Engineering Department (LGED) for providing important information to conduct this research. We also like to acknowledge Mr. Ziauddin, joint secretary in the Ministry of Commerce, who acted as a liaison between us and the government offices.

We are indebted to our families and friends for their understanding, patience and encouragement during this journey. Their support and cooperation were essential for the completion of this thesis work.

Finally we would like to express my sincere gratefulness to the Department of Civil Engineering, IUT and our respected teachers of our department. And once again we would like to thank our supervisor Dr. Shakil Mohammad Rifaat for giving us an opportunity to work with him.

Table of Contents

List of Tables

List of Figures

CHAPTER ONE: INTRODUCTION

1.1 Background

According to world report on traffic injury prevention-2004, road traffic accidents as the 6th place (was the 9th in 1990) of a major cause of death worldwide, will rise to become the 3rd leading cause of DALYs (Disability Adjusted Life Years) lost by 2020; the 2nd leading cause of DALYs lost for low and middle income countries; fatalities will increase worldwide from 0.99-2.34 million (representing 3.4% of all deaths); fatalities will increase on average by over 80% in low-income and middle-income countries and decline by almost 30% in high-income countries. Road accidents are increasing at an alarming rate. Every year more than 1.17 million people die in road crashes around the world. The majority of these deaths, about 70% occur in developing countries.

As a developing country, Bangladesh is not out of this situation. Road traffic accidents with casualties are causing great concern regarding communications within Bangladesh. The road safety situation in Bangladesh has been deteriorating with rapid growth in population, motorization, urbanization and lack of investment in road safety. The combination of rapid urbanization and motorization has made the problem more severe. Roads, highways and streets are fundamental to transportation system and over 70% of passenger travel and much of goods mover is over the highways (Sheikh, 2009).

The statistics reveal that the fatality rate of road accidents in Bangladesh is very high, with about 160 deaths per 10,000 motor vehicles as compared with the rates of 2 in the USA and 1.4 in the

UK. Together with social impact in terms of pain, grief and suffering, road accidents appear to impose a serious economic burden on country (Hoque et al., 2003). The 9th International Conference on Safe Community revealed that in Bangladesh, more than 2,000 people are killed in road accidents every year, which are about 6 persons every day. According to BRTA, Bangladesh has a fatality rate of 55 persons per 10,000 vehicles (60/10,000; another study at The Financial Express, July 2005). Recent figures from Transport Research Laboratory (U.K) showed that Bangladesh's death rate for traffic accidents is twice of the rate that of India (another study-33 per 1000 vehicles according to the Financial Express, July 2005) and 30 times that of developed countries like Japan. A survey conducted by the Accident Research Centre of Bangladesh University of Engineering and Technology in 2004 shows that more than 2,000 people die each year in road accidents (Sheikh, 2009).

According to the official statistics, there were at least 3334 fatalities and 3740 injuries in 4114 reported accidents in Bangladesh during 2011. It is estimated that the actual fatalities could well be 10000-12000 each year. Significant fluctuations in the numbers of fatalities and injuries as reported by police clearly reflect the problems of reporting and recording inconsistencies (BRTA). The number of fatalities has been increasing from 1009 in 1982 to 3334 in 2003, nearly 3.5 times in 22 years period (Hoque, 2008). The statistics revealed that Bangladesh has one of the highest fatality rates in road accidents, over 100 deaths per 10,000 motor vehicles (Haque & Meraz, 2005). About 70% of road accident fatalities occurred in rural areas including rural sections of national highways; here vehicles on road excluding motorcycle and non-motorized vehicles (Hoque, 1998).

Road collisions are the leading cause of permanent disability for children in Bangladesh accounting for about 1360 children being permanently disabled each year. Out of 30,000 children (aged 0-17) killed each year from injury and 3400 children (aged 1-17) are killed in road accidents, the majority of whom are from poor families (Zhou & Sisiopiku, 1997, Jonah, 2004). Although no information is available regarding the total number of disability and its relation to road accidents the proportion would be in the same range (Alam et al., 2012).

According to a study in Bangladesh carried out by TRL of UK it is observed that about two thirds of road deaths occur at the scene of the crash, one quarter take place in hospital within 30 days, and the remaining deaths occur after 30 days. Most of the seriously injured are treated at emergency units in hospitals (74%) with another 16% visiting doctors, and the remaining 10% seeing allopathic quack (TRL & Silcock, 1996).

The current road crash costing of the Roads and Highways Department is based on the assumption that 49% are fatal, 19% grievous, and 7% of simple crashes are reported (24% not reported.). Average property damage cost (includes average vehicle damage cost and average cost of the damaged goods) is Taka 55,430 for fatal crash, Taka 73,210 for grievous injury, and Taka 60,620 for simple injury. Medical costs include the any at-scene treatment through to rehabilitation; and the discounted value of funeral costs (average cost of funeral is Taka 10,600). For all those apart from vehicle owners, medical costs is the largest direct cost and have the most immediate impact on the family. Average medical cost is Taka 11,800 for road death, Taka 18,800 for serious injury, and Taka 1,400 for slightly injured (Jahid, 2009; Amin, 2011).

One way of improving road safety is to reduce crash occurrences by implementing crash reduction countermeasures. Another way is to reduce the frequency of crashes with safer vehicle and roadway designs as well as road user behaviour modification programs. However, these two methods can only be successfully applied if the relevant factors that contribute towards the occurrence or increase the severity of crashes are known. The injury risks of individuals in traffic crashes are influenced by a multitude of factors, including vehicle features, roadway designs and operations, driver characteristics, type of collisions and environmental conditions. In addition to identifying the significant factors, it is essential to quantify the relative magnitudes of the impact of these factors on collision frequency and severity so that countermeasures to prevent collisions can be prioritized and implemented.

1.2 Objective of the Study

From the above discussions, it is clear that factors related to road infrastructure, land use and socioeconomic and demographic features have significant impact on crash occurrence of those areas. The key aims are to:

- Understand the composition of traffic accidents and casualties data.
- Identify the magnitude of traffic accidents and casualties.
- To examine the effect of district wise road infrastructure, land use & demographic features on crash occurrence.
- Make recommendations to improve road safety and reduction of traffic accidents and casualties.

.

1.3 Scope & Significance of the Research

The study will examine the effects of district wise road infrastructure, land use & demographic features on the frequency of crashes using statistical models. District level data has been taken as the unit of analysis. It is expected that this study will answer the following queries:

- How the increasing or decreasing of roadway length will affect the number of accidents.
- Increasing or decreasing the length of the paved section may affect the frequency of crashes
- Increasing or decreasing the number of structures (bridge and culverts) may also affect the likelihood of crashes.
- How the increase or decrease of number of people living in any district will affect the incidence of road casualties.
- How change in literacy rate will affect the frequency of crashes.
- Increase or decrease in household size including male-female ratio may affect the probability of accidents.
- Increase or decrease in economic activities and income for may affect the number of crashes.

1.4 Outline of Thesis

The thesis is organized into five chapters. After the introduction in first chapter, the other four chapters will cover the following topics:

Chapter 2- Literature Review:

Factors associated with crash frequency explored in previous studies are reviewed in this chapter. Important information and finding from these studies are also documented.

Chapter 3- Data and Methodology:

Chapter three describes the sources of the database used in this study as well as methodology followed in statistical analysis. This chapter also justifies the methodology used in this study comparing them with the methodologies used in similar studies.

Chapter 4- Crash Frequency Models:

In this chapter, a crash frequency model will be developed by using the available data. Description of model development and evaluation and explanation of significant variables for each case study are given as well.

Chapter 5- Conclusion:

This chapter draws final conclusion based on the findings of different case studies in crash frequency analysis. Some directions for future exploration of research in this area are also mentioned.

CHAPTER TWO: LITERATURE REVIEW

2.1: Introduction

This Chapter reviews studies related to road traffic accidents and casualties mainly in Bangladesh and also in some other countries. The review identifies and reports on, the findings from studies carried out in this country as well as studies done in abroad. The Literature review of the thesis is divided into two main parts. First, the key findings of various published papers on the studies related to crash occurrences will be summarized to help readers gain some insights about the variables and contributing factors to be used. The discussion will provide some set of facts for the results in this study later on. Finally, a critical review is presented on the assumptions and limitations of different statistical models that are commonly used in crash occurrence to provide the foundation for the methodology chosen in this study.

2.2 Crash Contributing Factors

Road crashes not only claim lives and inflict injuries but also create economic burden to the society due to loss productivity. Recently, a number of studies have been done where these contributing factors are identified as playing important roles in affecting the frequency of road crashes. These studies render a deep perception on how crashes are affected by the contributing factors as well as render the fundamentals for developing empirical models to be used in this study. Hoque et al., (2003) found that, in Bangladesh the principal contribution factors of accidents are adverse roadside environment, poor detailed design of junctions and road sections, excessive speeding, overloading, dangerous overtaking, reckless driving, carelessness of road users, failure to obey mandatory traffic regulations, variety of vehicle characteristics and defects in vehicles. Others include a low level of awareness of the safety problem by policy makers, inadequate and unsatisfactory education, safety rules and regulations and inadequate and unsatisfactory traffic law enforcement and sanction (Sheikh, 2009).

Several factors related to driver, vehicle, road, crash and pedestrian characteristics have been widely explored in many studies to determine their influence on crashes. Some of these factors have been found to significantly increase crash severity whereas others have been found to reduce crash severity. Therefore, this section will provide a summary of the various factor commonly used in past studies on crash severity to get a clear idea about the possible outcomes for our empirical models.

2.2.1 Factors Related to Road Infrastructure

Various road infrastructures related factors had been explored in studies to determine their influence on crash frequency. For example, in a study Hoque(2006) analyzed the accident distribution of the national highways of Bangladesh and found that of the total reported accidents nearly 39% occurred in national highway .Regional feeder highways & feeder roads accounted for about 12% & 15% of accidents respectively. Of the accidents and fatalities of known locations, the share of national highways is much higher 56% and 61% respectively.

In another study Rab(2006) found that about 60% fatal accidents in Bangladesh occur on the national and regional highways and, of the remaining 15% occur on the feeder roads,7% on the rural roads and 19% on the city roads. The highest number of accidents is 'hit pedestrian', both in the urban and the rural areas, followed by head on and rear end collisions, and then the overturning of vehicles and often falling in side ditches. Persons who are killed as the result of road accidents are predominantly pedestrians and the passengers and drivers. The fatality rates are 48% pedestrians, 37% passengers and 15% drivers. It is encouraging to note that despite improved roads, higher speeds of vehicles and increase in the volume of traffic during last several years, the country has a somewhat reducing trend in road accident death rates.

Vogt (1999) provides an extensive review of the factors, which have been considered in past research studies in U.S.A. These factors include channelization (right and left turn lane), sight distance, intersection angle, median width, surface width, shoulder width, signal characteristics, lighting, roadside condition, truck percentage in the traffic volume, posted speed etc. Beside these factors, researchers have also considered other minor details such as surface bumps, potholes, pavement roughness, pavement edge drop-off etc. (Graves et al., 2005).

Mountain, Fawaz & Jarrett (1996) showed in a British study that the road design features- link length relates to accident rate, especially in dual carriageway. Retting, et at. (2001) studied the affect of roundabout on the traffic accidents; and found that replacing signals or stop signs with roundabouts could reduce traffic accidents. Road design factors like, the curve radii, spiral lengths, lane width, shoulder width, and tangent lengths are shown to be related the collision frequency (Easa and Mehmood, 2008).

From the above discussions, it is clear that road infrastructure and traffic related factors in a country have significant impact on crash occurrence. Road related factors that are commonly explored in those studies are length of different types of roads, roadside condition, right and left turn lane, intersection angle, median width, lane width, shoulder width, tangent lengths, spiral lengths, curve radii, signal characteristics, lighting, bumps, potholes, pavement roughness, pavement edge drop-off etc.

2.2.2 Factors Related to Vehicle Characteristics

The types of vehicle involved in accidents were considered in many studies as an important contributing factor affecting crash frequency. For Example, Hoque et al., (2006) in a study of road accidents revealed that heavy vehicles such as trucks and buses including minibuses are major contributors to road accidents (bus/minibus 33%, trucks 27%) and in fatal accidents their shares are 35% and 29% respectively. This group of vehicles is particularly over involved in pedestrian accidents accounting for about 68% (bus/minibus 38%, trucks 30%). For the case of road death, the share of buses and trucks are nearly 70% (bus/minibus 36%, trucks 24%) and for pedestrian about 72% (bus/minibus 40%, trucks 32%).

Hoque (2006) examined major fatal road accidents in Bangladesh and found that from 2002 to 2005 at least 126 major accidents occurred causing at least 612 deaths & 1857 injuries on the highways. Fatal accidents which occurred on highways are likely to involve heavy vehicles, especially buses, and are likely to involve multi vehicle accidents. Excessive speeding & road and roadsides hazardous conditions multiply the severity of accident (Hoque et al. 2006).

Hoque et al., (2006) found that, over involvement of truck and buses are particularly prevalent in Bangladesh. Earlier studies of police reported accidents on a section of Asian highways revealed that trucks and buses (including minibuses) accounted for about 72% (truck 34%, buses 20% and minibuses 18%) of the vehicles involved in accidents. This group of vehicles is particularly over involved in pedestrian accidents accounting for about 79% (trucks 37%, buses 20% and minibuses 22%). At some locations, trucks involvement was found to 43-50%.

Hoque et al., (2006) showed in one of its study that, in Bangladesh on an average, of 1.44 vehicles are involved in each major fatal accident. Of the total vehicles involved, 46% are buses, 25% are trucks and 10% are Baby Taxi/ Tempos. Heavy vehicles are involved in 71% of all major fatal accidents. Each accident in which buses are involved claims 23.9 casualties (5.2 fatalities and 8.7 injuries) whereas for trucks each accident claims 18.0 casualties (5.0 fatalities and 13.0 injuries). The study also reveals that around 38% accidents are single vehicle and 62% are multi vehicle types for all the highways. Of the total multi vehicle involvement in accidents, 35% are occurred between bus and others except truck. From critical observation it can be seen that most of the single vehicle accidents are hit object on/off road and lost control. The severity level of single accident is high as at least 5 persons are killed and 16 persons are injured per accident whereas in multi vehicle accident at least 4.6 persons are killed and 14 persons are injured.

Hossain (2006) found in a research work in Thailand that, motorcycles are involved in 51.22% accidents, pickup van in 22.05% accidents, personal car in 6.75% accidents and heavy truck in 3.51% accidents. A farther analysis reveals that children have tendency of being struck by a motorcycle (57.50%). The percentage is higher for pickup van also (24.26%). Almost 82% of the pedestrian children who become victims of road accident are struck either by a motorcycle or by a pickup/van.

From the above discussion, it is clear that the vehicle related factors affecting crashes include type of vehicle, special class of vehicle, number of vehicle involved etc.

2.2.3 Factors Related to Driver Characteristics

A number of studies have identified the effect of driver characteristics on crash frequency. For example, Hoque et al. (2006) found that, in Bangladesh drivers in 26-45 years of age constitute the most dominant group involved in accidents as revealed in the analyses of 8,500 drivers. The overall age distributions of drivers' involvement in accidents: Age group 26-35, 51.8% (26-30, 24% and 31-35, 27.8%); 36-45, 27.3%; 16-25, 16.3%; 46-55, 3.4%. Drivers in 26-35 years of age from the dominant cohort in driver casualties with 48.2 percent of total driver casualties (26-30 age group, 23.6% and 31-35 age group, 24.6%). Drivers of age below 16 years (55 drivers) are about 1.9 percent of total driver casualties.

The study also showed that, out of 8,500 drivers with recorded information, about 4,480 were heavy vehicles' drivers which accounted for nearly 52.8% of total drivers' involvement in the casualties with Bus drivers (minibus inclusive)-28%, Truck drivers (heavy trucks, articulated trucks and oil tankers inclusive)-24.6%, Car drivers (microbuses, jeeps and pickups inclusive)-14.5%, Baby taxi drivers (tempos inclusive)-11.5%, Rickshaw pullers (push carts inclusive)-7.7%, Motor cyclists-5.8% and Bicyclists-4.6%.

A study in India (Mishra et al., 1984) found that motorcyclists benefited from any type of helmet with padding, whereas a study in USA (Peek-Asa et al., 1999) found that the non-standard helmets used by half of all motorcyclists produced more frequent head injuries than not wearing a helmet at all. Among moped and motorcycle riders, head injuries account for about 75% of deaths in Europe (COST 327, 2001) and 55-88% in Malaysia (Radin, 2002).One study (Kulanthayan et al., 2000) found that riders without helmets were 3 times more likely to sustain head injuries than those with helmet. Another (Servadei et al., 2003) found that helmets reduced fatal and serious head injuries by 20-45%. Fell et al. (2008) found that 52% of the offenders were considered problem drinks, 46 were repeat offenders, 57% were classified as hardcore drinking drivers, 51% were drinking at a bar or restaurant before their arrest, and 72% were drinking beer before their arrest, analyzing data for drinking characteristics of drivers arrested for driving while intoxicated, of 1027 drivers arrested by police for alcohol consumption in USA (Sheikh, 2009).

From the above discussions, it is clear that the common driver related factors using in crash frequency studies include the age of the driver, type of vehicle driving, safety adoption, whether the driver of vehicle was drinking or intoxicated, driver's condition etc.

2.2.4 Pedestrian Related Factors

In many studies, pedestrian characteristics were found to have an effect on accident frequency. For example, Hoque (2006) found that, pedestrians alone are involved in more than 47 percent of road accidents and 49 percent of all fatalities. In urban areas pedestrians accounted for 52 percent of fatalities and in Dhaka city are nearly 70 percent. About one third of the total pedestrian fatalities are children under age of 16 years of age are accounting for nearly 22 percent of all fatalities in Bangladesh. Fatalities of female child pedestrians are disproportionately higher than those of the male child pedestrians (45% vs. 29%). It was found that road traffic accident is the leading cause of death in 10-14 years age cohort. It was also revealed in the study that most economically active age group (16-45) is the hardest hit in the road traffic accidents.

Hoque et al. (2003) mentioned that for Bangladesh about 50% of the reported road accident deaths and 17% of reported road accident injuries are pedestrians. Among them, the overall child involvement in road accident fatalities in Bangladesh is found to be very high, accounting for about 22%. The children pedestrian fatality rate is 32% and injury rate is 25%. It is important to note that compared to industrialized countries, the proportion of fatalities to under 15 years of age in developing countries is approximately 2.5 times higher. Of the total child fatalities of road accidents, nearly 82% are involved as pedestrians with the dominant age group of 5-10 years. Indeed, about 1/3 of total pedestrian fatalities are children under the age of 15 years. The female child pedestrians are disproportionately higher than the male child pedestrians (44.6% vs. 28.9%).

According to Aeron-Thomas et al. (2004) that girls accounted for a larger share of total female deaths and injuries (32%) than boys did for total males (12%). This was consistent with the draft findings of the Bangladesh Health and Injury Survey that found the mortality rate for transport injuries was much higher for girls under the age of 10 than for boys (Rahman, 2003). According to White (2002), Child passenger casualty numbers in Britain are particularly high in conditions where it is raining or snowing. Over 18% of child passenger casualties occurred while it was raining compared to 7% of cyclists and 9% of pedestrians.

Hossain (2006) found that, in Thailand the most vulnerable group of pedestrians is the children younger than 10 years old represented by 26.12% of all pedestrian victims. In case of children pedestrians younger than 10 year old, more than 57% of them face accident due to motorcycles.

From the above discussion, it is clear that the pedestrian characteristics effect crashes. Age of the pedestrian is a very important factor which affects crash frequency.

2.3 Different Types of Crash

'Accident type' is regarded as one of the primary importance among the variables readily studied. According to Hoque et al. (2006), in Bangladesh most common accident types are hit pedestrian(43.7%), rear end collisions (16.4%), head on collisions (13.3%), and overturning (9.4%). These four accident types account for nearly 83% of the accidents, hit pedestrian accidents account for 52.9%, followed by rear end 11.9%, head on 11.9% and overturning 9.8 percent. In Metropolitan Dhaka , dominant accident types are hit pedestrians 44.4%, rear end collisions 29.1%, sideswipes 7.1%, and head on collisions 6%. As regards fatal accidents hit pedestrian accidents account for 68.9% followed by rear end collisions 17.2%.

Australian model guidelines (Andreassen, 1994) are used to describe the movements of the road users and vehicles involved in the accident event where different types of accidents are expressed by different codes. It is found that head on (201,205,501) type accidents which accounts for 39% of the total accidents dominate the others. This is followed by lost control (701,702,703,705,502) 34%, rear end (301,304,305) 16%, hit object on/off road (601,605,607,608) 7%. These four accident type accounted for 96% of the total accidents. Hit object on/off road type of accident represents the highest rate of about 5.0 fatalities and 23.5 injuries per accident. This is followed by the lost control (about 5.34 fatalities and 19.6 injuries per accident. This is followed by lost control (about 5.34 fatalities and 19.6 injuries per accident) and head on (5.0 fatalities and 16.1 injuries per accident).

2.4 Review of Statistical Models and Methodologies

The road accidents are influenced by several of factors, including vehicle characteristics, driver behavior, pedestrian characteristics, type of collisions, road and traffic characteristics etc. It is essential to quantify the relative magnitudes of the impacts of these factors on collision occurrence. This will help in developing and implementing countermeasures to prevent collisions or reduce collision severity.

Various studies have been done to identify those factors related to crash frequency as well as crash severity. These studies use statistical models either to determine collision frequencies, collision involvement rates, collision probabilities and collision severities. In this section, the assumptions as well as the limitations of commonly used statistical models applied in crash frequency as well as crash severity studies will be discussed. However, the details and mathematical formulations of these models will only be described in Chapter Three.

2.4.1 Crash Frequency Models

The most commonly adopted approach to correlate annual accident frequencies with geometric and traffic factors is to use econometric or statistical models. These methods are popular because of the strong association between risk factors and crashes. Researchers have been using various statistical techniques to model accident frequencies, ranging from the use of multiple linear regression (MLR) to methods involving exponential distribution families such as Poisson and negative binomial (NB) regression models.

Initially multiple linear regression was used for model formulation. However, as pointed out by Joshua and Garber (1990), linear regression models do not describe the nature of the crash frequency data adequately. There are several statistical properties that are considered undesirable to describe adequately the random, discrete, nonnegative and sporadic accident data.

Chung, (2005) discovered that Generalized Linear Models is widely used, is that elements affecting in accident are categorized data. Moreover, because accidents are discrete essentially, expression of difference about accident reaction is most efficient in expressed data system by categorized style. And this method makes it possible to test significance of categorized data as fixed quantity.

Compare to the Multiple Linear Regression, Chung (2005) the model has various shortcomings to use for predicting the number of accident. First, dependent variables are assumed to follow normal distribution in this model, but the number of accident is not so. And it is assumed that there is no relation between error and independent variable, but this assumption is not always true in case of accident in actuality. In addition, this model can deduce the negative number that could not appear as the number of accident. Moreover, when accident did not happen in any spot, this method always predicts zero as the number of accident, and this result strains the truth that zero number means that spot absolutely safe (Manan, 2011). As a result, these linear regression models are not appropriate to make probabilistic statements about vehicle accidents on the road, and test statistics derived from these models are thus questionable.

The unsatisfactory property of linear regression models has led to the investigation of the Poisson model. Poisson or Negative Binomial (NB) regression models, instead, are better suited for defining the random, discrete, and nonnegative nature of crash occurrence (Milton & Mannering, 1998). The log-linear model is the best known example of Poisson regression. It essentially is a generalized linear model (GLM) for Poisson-distributed data and specifies how the size of a cell frequency depends on the levels of categorical variables for that cell. The nature of this specification relates to the association and interaction structure among the categorical variables (Agresti, 2002).

It should be noted that the Poisson model formulation requires the mean and variance of the crash data to be equal. While, the assumption is a good approximation in some cases, it is not generally appropriate. If this assumption is not valid, the standard errors, usually estimated by the maximum likelihood (ML) method, will be biased and the test statistics derived from the models will be incorrect. Therefore, the NB model, which has all the desirable statistical properties and also relaxes this constraint, is the most popular model formulation for crash frequency estimation. A detailed comparison between Poisson and Negative Binomial crash frequency models may be found in Miaou (1994). The findings suggested that since crash data tend to be over dispersed (i.e., variance $\gg$ mean), negative binomial modeling is the more appropriate technique of the two. Negative binomial (NB) distribution can allow greater variance than the Poisson distribution. There are two basic assumptions with the negative binomial model:

(1) Accident frequencies at any given site are Poisson distributed.

(2) Individual site means among the similar sites are gamma distributed.

The combination of the two assumptions results in variance function that is greater than that of the Poisson distribution and this overcomes or at least reduces the over-dispersion problem encountered in the Poisson model.

CHAPTER THREE: DATA AND METHODOLOGY

This chapter contains the data collection method, formulation of data and the methodology used in the analysis of crash data in this study. The key steps in the methodology are presented first, followed by the theoretical framework of several crash frequency models, such as multiple linear regression (MLR), Poisson regression and negative binomial(NB) model, used in this study. The development of these models will assist us to understand how these models can be employed to fulfill the major objective of the study; that is, to identify the effects of street pattern on crash occurrence as well as crash severity. However, these models need to be calibrated with a reliable and representative dataset. For this reason, the sources of database used in this study are discussed before describing the mathematical formulation of the statistical models used in this study.

3.1 Main Steps in Methodology

To achieve the objectives of the study, suitable statistical models need to be selected. The models will be developed using actual accident data that relates crashes with road number, road infrastructure, socioeconomic and demographic factors. The methodology is divide into several steps:

1. Statistical model selection to correlates number crashes as a function of different factors
2. Surveying, collection and processing of data ,both accident and different factors to establish the model
3. Analysis and interpretation of model result.

3.2 Description of Data

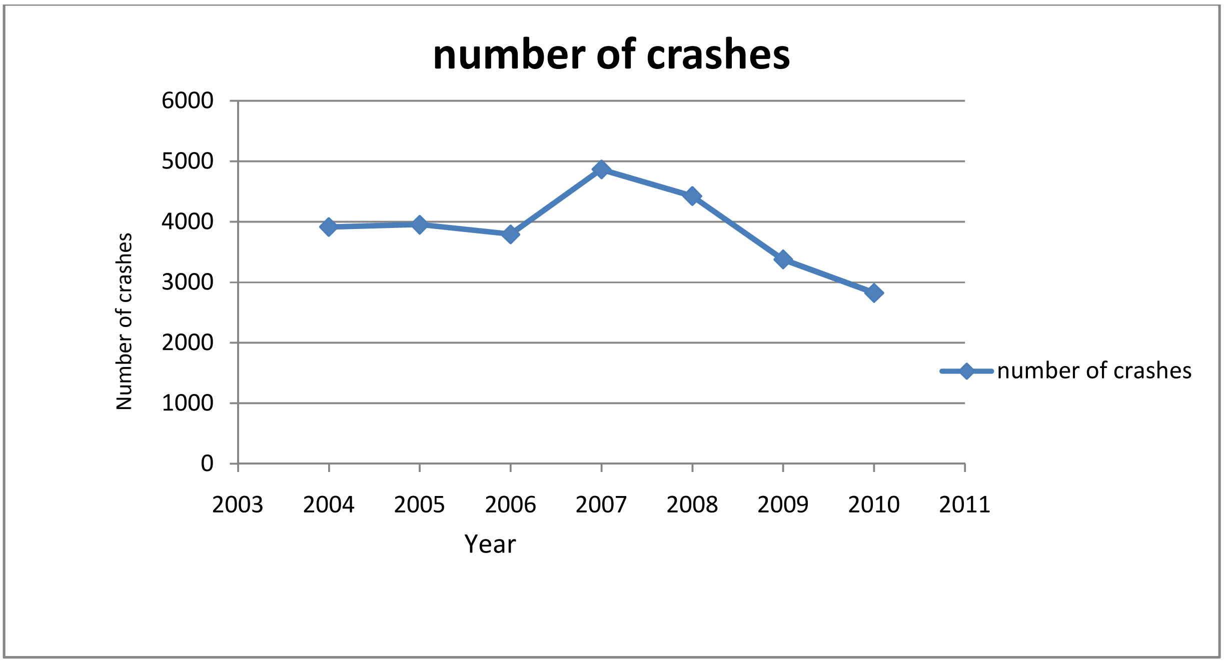
3.2.1 National Road Traffic Accident Database

The crash data for this study was obtained from the official crash database maintained by Bangladesh Road Transport Authority (BRTA). BRTA and the Police have been working closely together to safeguard the integrity of the database and to promote its statistical accuracy. The information of crashes first obtains by the police department. Therefore, the accuracy of the crash data would largely be dependent on the policeman who recorded the information about the crash. In this study, district wise accident data from the year 2004 to 2010 for Bangladesh were collected from the official crash database.

3.2.2 Crash Trends in Bangladesh

To recognize the accident condition in Bangladesh, a concise review of the accident statistics is provided. The crash statistics in Bangladesh for the year 1994-2011 is shown in Figure 3.1

Fig 3.1: crash trends in Bangladesh



On average, there are close to 3882 traffic accidents on Bangladesh roads annually.

3.2.3 Socio-Economic, Road number and Road Infrastructure Data

Demographic and socioeconomic characteristics for each district were collected from the 2001 census data published by Bangladesh Statistical bureau (BSB). From there we got average dwelling household size for year 2008. Percentage distribution of household living structure for year 2008, Percentage distribution of household economic condition for year 2008, Number of population for year 2008, household, birth & death by sex, Literacy rate & Estimated population for year 2001, Economically active population by labour force participation rates for year 2005-06. Several important socioeconomic data were missing from the census data such as: education level, income level, age distribution etc.

From Road and highways department (RHD), the total length of district wise national, regional, Zilla road with paved & unpaved length for each of the classification was found. For each type of total and categorical data were considered such as, for national paved the categorical data were 0-40, 41-80, 81-120 and above 120.

The total district wise number & span of structure (bridge, culvert etc), number & span of existing gap, these data were acquired from Local Government Engineering Department (LGED). Like road number, continuous and categorical data were considered for road infrastructure such as 0-1500, 1501-3000, 3001-4500 and above 4500 for total number road structure (bridge, culvert etc).

3.3 Statistical Models for Crash Frequency Analysis

The principal objective of a statistical model is to identify a probabilistic system of the form:

$$Y = f(x) \tag{3.1}$$

Where the dependent variable Y is a function of a set of independent variables X. The regression analysis of the above-mentioned form examines: a) whether the observed patterns in the data are consistent with theoretical prediction; and b) the relationship between a quantitative dependent variable and one or more quantitative or qualitative independent variables. In many accident frequency studies, the dependent variable Y usually represents the total number of annual traffic accidents as a count outcome and the independent variables X represent the associated roadway geometrics, traffic and regulatory controls, and other relevant characteristics.

This section gives a detail description of the statistical models which will be used for accident frequency analysis, their mathematical form, assumptions and limitations. These models include the conventional multiple linear regression model (MLR), poisson and negative binomial regression models. A detail description of model evaluation that includes tests based on goodness of fit will then be discussed. Finally, a method to find the elasticity of independent variables will be described.

3.3.1 Multiple Linear Regression Model

The most widely used regression model to analyze collision frequency data is the multiple linear regression model (Hakkim, 1984) which is given by:

$$y = X\beta + \varepsilon \tag{3.2}$$

Where y is the number of collisions, X is a vector of explanatory variables; β is a vector of parameters to be estimated. ε is the error term.

The assumptions of the OLS model are:

a) Expectation of the error term is zero.

b) Homoskedasticity: variance of the errors is the same regardless of the value of X.

c) Normality: the error is normally distributed.

d) Independence: the observations are free from autocorrelation.

The above model is usually estimated using the ordinary least square method (OLS). If the assumptions are valid, then the OLS estimator is the best linear unbiased estimator.

However, since accident data are usually random, discrete, nonnegative, sporadic and count data, there are a lot of undesirable properties in MLR such as assumption of normality and common variance as well as the possibility of negative outcomes that results in misinterpretation of count data (Joshua and Graber, 1990). To overcome the problems associated with MLR models, Miaou (1994) proposed that the Poisson regression should be instead used for modeling accident frequencies.

3.3.2 Poisson Regression Model

Poisson models deal with discrete data so that they have most of desirable characteristics to describe vehicle collisions of positive number and random attribution. Joshua and Garber (1990), Milton & Mannering (1998) studied the relationship between highway geometric factors and accidents and all of them came to the conclusion that Poisson regression model is superior to MLR to describe discrete, random, non-negative, sporadic accident data.

The basic assumptions of Poisson distribution are:

a) Probability of more than one event occurring in a short period of time is zero;

b) Probability of one count in a subinterval is the same for all subintervals and proportional to the length of the subinterval; and

c) Count in each subinterval is independent of other subintervals.

If event ‘n’ occurs according to a Poisson process with parameter μ, then the Poisson distribution can be written as:

$$\Pr(n_{it} \mid \mu_{it}) = \exp(-\mu_{it})_{*}\, \mu_{it\,/}\, n_{it!} \qquad (3.3)$$

where (n_{it}) is the probability of ‘n’ accidents occurring on roadway section ‘i’ in time ‘t’. μ_{it} is the expected number of accidents on roadway section ‘i’ in time ‘t’.

The Poisson distribution has the limitation that the variance and mean should be approximately equal i.e.

$$\mathrm{Var}(n_{it}) = E(n_{it}) = \mu_{it} \qquad (3.4)$$

In the case of accident frequencies, the variance is generally much larger than the mean (described as the over-dispersion phenomenon) at which point the Poisson model becomes inappropriate.

3.3.3 Negative Binomial Model

The negative binomial model has been used intensively by recent studies as the most appropriate methodological technique for modeling accident frequencies (for example: Walmsley et al., 1999 and Lee and Mannering, 2002). As mentioned above, the limitation of the Poisson model is that the variance and mean must be approximately equal but, in general, accident data have a variance exceeding the mean. To deal with the limitations of the Poisson model, a negative binomial based on a gamma-distributed error term is commonly used. Therefore, equation 2 can be rewritten as follows (Shankar et al., 1995, Walmsley et al., 1999 and Lee and Mannering, 2002):

$$E\,(ni) = \lambda i = \exp\,(\beta Xi) + \varepsilon i \qquad (3.5)$$

Where εi is a gamma distributed error term. This addition will make it possible that the variance is different from the mean following the next equation:

$$Var(n) = E(n) + \theta E(n)^2 \qquad (3.6)$$

The negative binomial model is described by the following equation:

$$P(ni) = \Gamma((1/\theta)+n\,)/\;\Gamma(1/\theta)n!*((\,1/\theta)/(\,1/\theta+\lambda i))\;1/\varphi*(\lambda i/(1/\theta+\lambda i))\;ni \qquad (3.7)$$

From equation 3.7, it can be seen that if parameter θ is equal to zero, then the negative binomial model becomes a Poisson model; therefore the Poisson model can be described as an absolute of the negative binomial model.

3.3.4 Model Evaluation

The statistical models will be evaluated to select the best model from the competitive set of models. The evaluation will be done with the help of two statistics:

1. Likelihood Ratio Test.
2. Log-likelihood Ratio Index (ρ^2).

Likelihood ratio test is a common test used to assess two competing models. It provides evidence in support of one model, usually a full or complete model, over another competing model that is restricted by having a reduced number of model parameters. The models are also tested for an overall goodness-of-fit with the help of log-likelihood ratio index. Finally, the elasticity values are calculated for the final model to find the relative effects of different independent variables. These are described in detail in the following.

3.3.4.1 Likelihood Ratio Test

The likelihood ratio test statistic is:

$$X^2 = -2\ LL\ \beta_{\boldsymbol{R}} - LL\ \beta \qquad (3.8)$$

Where, $LL\ \beta_{\boldsymbol{R}}$ is the log likelihood at convergence of the restricted model.

$LL\ \beta_{\boldsymbol{U}}$ is the log likelihood at convergence of the unrestricted model.

The test statistic is X^2 distributed with the degrees of freedom equal to the difference in the numbers of parameters in the restricted and unrestricted model.

3.3.4.2 Log-Likelihood Ratio Index

To measure the overall goodness of the models, the log-likelihood ratio index will be calculated. The R^2 as a goodness-of-fit measure for OLS estimator has been used by traffic safety engineers and researchers for many years. However, since count data models are non linear, there is no R^2

instead, the common practice is to use a pseudo R^2 statistic, which is often known as log-likelihood ratio index which is given by:

$$\rho^2 = 1- L(\beta)/L(0) \qquad (3.9)$$

Where, $L(\beta)$ is the log-likelihood value of the fitted model.

L(0) is log-likelihood value of the model only with constant term.

Everything else being equal, a specification with a higher maximum value of log-likelihood function is considered to be better. The lowest value of log-likelihood function corresponds to the model with constant term only and is considered the worse case. The value of R^2 is between 0 and 1, the better models approaching the latter. Like the R^2 statistic, it has the undesirable characteristic that for same data set, it will increase whenever new variables are added to the model. To overcome this disadvantage a correction for the number of covariates, p, to give the adjusted log-likelihood ratio index as:

$$\rho^2 = 1- (L(\beta)-p) / L(0) \qquad (3.10)$$

3.3.5 Elasticity of Variables

In order to check the relative significance of independent variables from the final model, the elasticity of the variables was calculated. Elasticity is defined as the percent change in dependent variable due to one percent change in the independent variable. The elasticity of accident frequency μit with respect to x_{itk} is defined as:

$$E = (d\mu it / \mu it)*(x_{itk} / dx_{itk}) \quad (3.11)$$

Where, xitk is the k th independent variable for section i in year t.

Using this equation, the sensitivity of accident occurrence to any variable can be obtained. With the increase in elasticity value the sensitivity of accident occurrence to a change in the specific variable also increases. The elasticity in Equation 3.12 is only appropriate for continuous variables. It is not valid for non continuous variables such as indicator variables that take on values of 0 or 1. For an indicator variable, pseudo- elasticity is computed to estimate an approximate elasticity of the variables. The pseudo-elasticity gives the incremental change in frequency caused by discrete (0-1) change inthe indicator variables. The pseudo-elasticity for indicator variable is computed as:

$$E = \{\exp(\beta)-1\} / \exp(\beta) \quad (3.12)$$

CHAPTER FOUR: CRASH FREQUENCY ANALYSIS

4.1 Overview

The main objective of crash frequency analysis to find out the factors related to road characteristics (type wise length), road structures (culvert, bridge etc, their length, span), existing gap and socioeconomic and demographic characteristics and their effect on number of accident. The second objective is to examine the sensitivity of our findings and policy recommendations to different specifications of performance functions with respect to combining collision frequency . To fulfill the objective, this study will develop a crash frequency model by using statistical analysis to relevant factors that affects the numbers of crashes. The unit for the model development will be taken as individual district which has not been explored much in the past.

4.2 Data

Several sources of data were used in this study. The organizations were considered for data collections are: Roads & Highways Department (RHD), Bangladesh Statistical bureau (BSB), Bangladesh Road Transport Authority (BRTA) and Local Government Engineering Department (LGED).Bangladesh Road Transport Authority (BRTA) provided yearly national road traffic accident reports for 64 districts in Bangladesh. The distribution of the number of reported crashes for the years 2004-2010 across Bangladesh. During the period analyzed, 21448 crashes occurred in the 64 districts in Bangladesh, giving an average of 48 crashes per district each year.

From Road and highways department, the total length of district wise national, regional, Zilla road, UpaZilla road & Union road with paved & unpaved length for each of the classification was found. The map shows different road network on Bangladesh. From GPS survey, RHD calculated the length of each class of road. The survey was done from 2002-2004.

The total district wise number & span of road structure (bridge, culvert etc) & of existing gap were acquired from Local Government Engineering Department (LGED).

Bangladesh Statistical bureau (BSB) provided us with several district wise demographic data. BSB provided Average dwelling household size for year 2008. Percentage distribution of household living structure for year 2008, Percentage distribution of household economic condition for year 2008, Number of population for year 2008, household, birth & death by sex, Literacy rate & Estimated population for year 2001, Economically active population by labour force participation rates for year 2005-06. Several important socioeconomic data were missing from the census data such as: education level, income level, age distribution etc.

Figure 4.2: Road Network in Bangladesh

4.3 Model development

As discussed in previous Chapter, count data models such as Poisson and Negative Binomial Model are considered in crash frequency analysis. The reasons for choosing these models, their assumptions and mathematical formulations of both models are detail presented and discussed in those chapters. As discussed in the previous chapter, negative binomial model have been chosen for crash frequency analysis. By using this model, co efficient can be developed for accident causing factors e.g. culvert, bridge etc, their length, span, type wise road length, socioeconomic characteristics etc. These coefficients will be use accident prediction model. This prediction model will be able to predict increase or decrease in the number accident due to the increase or decrease of one or several factors.

After data collection from several organizations, 107 variables were found which include both continuous and categorical variables. The collected data consist of district wise road characteristics (type wise length), road structures (culvert, bridge etc, their length, span), existing gap and socioeconomic characteristics. For insufficient data, different variables such as number bus, train station, licensing information etc are excluded from the model. The mean and the standard deviation of the variables used in the final models:

Table 4.1: Summary Statistics

Variable	Mean	Std. dev
Road length		
National paved road		
Total	103.2929	412.5381
0-40	0.4688	0.5029
41-80	0.4688	0.5029
81-120	0.2344	0.4269
Above 120	0.1094	0.3146
National road unpaved		
Total	0.3742	1.8244
0-5	0.9688	0.1754
Above 5	0.0312	0.1754
Regional paved road		
Total	124.5991	496.6572
0-60	0.4844	0.5037
61-120	0.4219	0.4978
121-180	0.4219	0.4978
Above 180	0.0312	0.1754
Regional unpaved road		
Total	1.9142	8.2469
0-5	0.9375	0.2439
6-10	0.0156	0.1250
11-15	0.0313	0.1754
Above 15	0.0156	0.1250
Zilla paved road		
Total	322.5458	1282.1940
0-100	0.2344	0.4269
101-200	0.4688	0.5029
201-300	0.2656	0.4452
Above 300	0.0313	0.1754
Zilla unpaved road		
Total	51.3859	205.7684
0-30	0.6875	0.4672
31-60	0.2031	0.4055
61-90	0.0469	0.2130
Above 90	0.0625	0.2439
Upazilla paved road		
Total	1593.68	6333.6890
200-700	0.4688	0.5029
701-1200	0.4219	0.4978
1201-1700	0.0938	0.2938

Above 1700	0.0156	0.1250
Upazilla unpaved road		
Total	913.4745	3637.9090
0-400	0.515625	0.5037
401-800	0.328125	0.4732
801-1200	0.1250	0.3333
Above 1200	0.0312	0.1754
Total number of road	3146.744	12504.05
Road structure & existing gap		
Number of Structure(bridge, culvert etc)		
Total	3574.8310	14226.60
0-1500	0.4531	0.5017
1501-3000	0.4219	0.4978
3001-4500	0.0938	0.2938
Above 4500	0.0313	0.1754
Structural span		
Total	20604.59	81935.87
0-700	0.3125	0.4672
701-1400	0.4531	0.5017
1401-2100	0.2031	0.4055
Above 2100	0.0313	0.1754
Existing gap		
Number of existing gap		
Total	381.5692	1531.689
0-300	0.8281	0.3803
301-600	0.1406	0.3504
601-900	0.0156	0.1250
Above 900	0.0156	0.1250
Span of gap		
Total	6797.934	27189.12
0-3000	0.5781	0.4978
3001-6000	0.2813	0.4532
6001-9000	0.0625	0.2439
Above 9000	0.0781	0.2705
Socio & demographic characteristics		
Household number	785119.5	3126425
Total house hold	3787491	1.51e+07
Male	1962816	7820955
Female	1824676	7267342
Male %	51.9034	5.1301
Female %	48.0966	5.1301
Household size		
Total	4.8828	0.4203
4-5	0.7188	0.4532
Above 5	0.2813	0.4532

Sex ratio		
Total ratio	105.5	4.6050
98-110	0.8906	0.3146
111-120	0.0938	0.2938
Above 120	0.0156	0.1250
Both sex	43.7813	8.5159
Male sex	47.8688	8.1040
Female sex	39.8813	9.2131
Economically active population 15+		
Total 15+	1322.484	1009.552
Male+	671.1719	508.0074
Female+	648.75	501.8555
Male% 15+	50.9717	1.2618
Female % 15+	49.0283	1.2618
Active total 15+	824.5625	664.4481
Active male 15+	586.1094	421.5046
Active female 15+	195.9375	162.5081
Active male% 15+	75.2161	13.5699
Active female% 15+	24.78394	13.5699
Economic condition		
Permanent insolvency %		
Total%	303.4508	2287.58
0-15%	0.4219	0.4978
16-30%	0.4531	0.5017
30-45%	0.0938	0.2938
Above 45%	0.03125	0.1754
Temporary insolvency %		
Total %	22.8527	7.55076
0-18%	0.28125	0.45316
19%-36%	0.625	0.4880
Above 36%	0.0938	0.2938
Equal expense		
Total	30.8008	9.3711
0-20	0.0781	0.2705
20-40	0.7656	0.4270
Above 40	0.1563	0.3660
Solvent		
Total	18.1459	6.4179
0-14	0.2344	0.4269
15-28	0.7031	0.4605
Above 28	0.0625	0.2440

Saving		
Total	10.7692	5.0441
0-14	0.8125	0.3934
15-28	0.1719	0.3803
Above 28	0.01563	0.125

Since some of the factors are recorded in a categorical manner, several discontinuous variables are used to find out their effects. Several of the contributing factors are consider in terms of percentage shares of the different categories. Since categorical data always sum to one or 100%, one of the categories has to be omitted from the model and used as a reference or base case by which the estimates of other categorical variables are compared.

4.4 Estimation Results

4.4.1 Aggregated Crash Frequency Models

In the aggregated crash frequency models, the estimated results from eight models are reported in table. All the eight models comply with the data fairly well. The chi square values were large when the goodness of fit of the models was checked. The over-dispersion parameters (a) in all three models were statistically significant (p-value < 0.0001) which showed the validity of using the negative binomial model instead of the Poisson model. The models also showed very small p values. The over dispersion parameter in all the models were very much statistically significant, this justifies the suitability of using negative binomial model.

Table 4.2: Estimation Results of Aggregated Crash Frequency Models

	Total accident	Year 2004	Year 2005	Year 2006	Year 2007	Year 2008	Year 2009	Year 2010
Number of observation	64	64	64	64	64	64	64	64
Chi-square	90.98	73.72	96.63	71.96	112.83	58.36	79.89	65.89
Alpha	0.1482	0.3527	0.2021	0.1822	0.1448	0.1715	0.1364	0.1455
P-value (Alpha)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Road length (kilometers)								
National paved road								
total	-	-	0.0094**	-	-	-	-	0.0059**
0-40	-	-	-	-0.3513*	-0.4973**	-0.2165	-	-
41-80	-	-	-	-	-	-	0.4790**	-
81-120	0.6104**	-	-	0.3921**	0.5592**	-	0.8454**	-
120>	0.9032**	-	-	0.6207**	-	0.5555**	0.8569**	-
National unpaved road								
total	-	-1.2925**	-	-	-	-	-	-
0-5	-0.6916 *	-	-	-	-	-	-	-
Above five	-	7.0511**	-	-	-	-	-	-
Regional paved road								
total	-	-	-	-	0.0043**	-	-	-
0-60	-	1.1026**	-	-	-	-	-	-0.2382*
61-120	0.3484**	1.0388**	-	0.3160**	-	0.3798**	0.5786**	-
121-180	0.4944**	1.5516**	-	0.7646**	-	-	0.8473**	-
Above 180	-	-	-	-	-	1.0599*	0.7288**	-
Regional unpaved road								
total	-	-	-	-	-	0.0334*	0.0478**	-
5-10	-	-	-0.9711*	-	-	-	-	-
10-15	-	-	-	-	-	-	-	0.7145**
Above 15	-	-	-	-	0.9345 **	-	-	-
Zilla paved road								
total	-	-	-	-	1.0054**	-	0.0014	-
200-300	-	-	-	-	0.2257	-	-	-
Above 300	-	-	-1.3596**	-	-	-	-	-

Zilla unpaved road								
total	-	-	-	-	-	-	-	0.0067**
0-30	-0.5084**	-	-0.5458**	-0.6871**	-	-	-0.2113*	-
31-60	-0.6222**	-	-0.610**	-0.8808**	-	-	-	-
Above 90	-	-	-	-	0.5407**	-	-	-
Union zilla paved road								
total	-	-	0.0006*	-	-	-	-	-
700-1200	0.2426**	-	-	0.2394*	-	-	-	-
Union zilla unpaved road								
total	-	-	-	-	-	-	-	-0.0008**
0-400	-	-	-	-	-	-	0.2493*	-
400-800	-	-	-0.2584*	-	-	-	-	-
Total number of road	-	-	-	-	-	-	-	0.0003
Road structure and existing gap								
Structure number								
0-1500	-	-	-	-	-	0.5202**	-	-
1500-3000	-	-	-	-	-	0.4851**	-	-
Structure span								
total	-	-	-	-	-	-	0.00003**	-
7000-14000	-	0.4269**	-	-	-	-	-	-
Number of gap								
total	-0.0007**	-	-0.00140**	-0.0006**	-	-	-0.0006**	-
0-300	-	-	-	-	-	0.6258*	-	-
301-600	-	-	-	-	-	0.8821**	-	-
Span of gap	-	-	-	-	-	-	-	-
0-3000	-	-	-	-	-	-	-	-0.3248**
Socio-economic characteristics								
Total household population								
Household number	-	-	-	-	-	2.24e-06**	-	-
Male%	-	-0.0661**	-	-	-0.0218*	-	-	-

Female%	-	-	-.03527**	-	-	-	-	-
Female 15+	-	-	-	-	-	-0.0010**	-	-
Male% 15+	-0.0733*	-	-	-0.0872	-	-	-	-0.1557**
Female% 15+	-	0.1644**	-	-	-	-	-	-
Economically active male % 15+	0.0074*	-0.0666**	-	-	0.0096	-	-	0.0105**
Economically active female % 15+	-	-0.0584**	0.0035**	-	-	-	-	-
Total active population	-	-	-	-	-	0.0003**	-	-
Permanent insolvency %								
Total %	-	-	-	-	-	-0.00004	-	-
15%-30%	-	-	-	-	-	-	-0.2266*	-
30-45%	-	-	-	-	-0.4091**	-	-	-
Above 45%	-	-	-	-	-0.6583**	-	-	-
Temporary insolvency %								
Total%	-0.0218**	-0.0318**	-0.04334**	-0.0238**	-	-	-	-
19%-36%	-0.3314**	-	-0.6571**	-0.4236**	-0.3661**	-	-	-
0%-20%	-	-	-	-	0.6438**	-	-	-
Equivalent expense %								
20%-40%	-	-	-	-	0.3932**	-	-	-
Solvent%								
total	-	-	-0.0324**	-	-	-	-	-
15-28	-	-	-	-	-	-	-	-0.2914*
Above 28%	-	-	-	-0.5092*	-	-	-	-
Saving%								
Above 28%	-1.1479**	-1.8820**	-	-2.4416**	-	-	-	-
Household size								
Above 5	-0.2889**	-	-	-0.3105**	-	-	-0.2374*	-0.4693**

Sex ratio								
Total	-	0.0565**	-	-	-	-	-	-
90-110	-	-	-	-	-2.2578**	-	-	-
110-120	-	-	-	-	-1.8626**	-	-	-
Above 120	0.8050*	-	-	-	-	-	-	-
Male sex	-	-	0.0994**	-	-	-	-	-
Female sex	-	-	-0.0927*	-	-	-	-	-
Constant	10.3126**	-7.1441	8.0782**	10.0732**	5.9081**	2.07385**	2.4982**	10.3413**
Note: * & ** denote statistically significant at α = 0.10 & 0.05 respectively; - denotes insignificant variable								

Also to the model estimates, the elasticity of crash frequency with respect to each contributing factor was also computed. The calculated elasticities are reported in Table.

Table 4.3: Elasticity Estimates for Crash Frequency Models

	Total accident	Year 2004	Year 2005	Year 2006	Year 2007	Year 2008	Year 2009	Year 2010
Road length (kilometers)								
National paved road								
Total			0.4947					0.3073
0-40	-	-		-.039811	-0.6444	-0.2418		
41-80	-	-	-	-	-	-	.38059	-
81-120	0.4569	-	-	0.3244	0.4284	-	0.5707	-
120>	0.5947	-	-	0.4625	-	0.4262	0.5755	-
National unpaved road								
Total	-	-0.2456	-	-	-	-	-	-
0-5	-0.9970	-		-	-	-	-	-
Above five		-5.5668	-	-	-	-	-	-
Regional paved road								
Total	-	-	-	-	0.0043	-	-	-
0-60	-	0.6680	-	-	-	-	-	-0.2690
61-120	0.2942	0.6461	-	0.2710	-	0.3160	0.4393	-
121-180	0.3901	0.7881	-	0.53450	-	-	0.5714	-
Above 180	-	-	-	-	-	0.6535	0.5175	-
Regional unpaved road								
Total	-	-	-	-	-	0.0325	0.0466	-
5-10	-	-	-0.1641	-	-	-	-	-
10-15	-	-	-	-	-	-	-	0.5106
Above 15	-	-		-	0.6072	-	-	-

Zilla paved road								
Total	-	-	-	-	1.0054	-	0.2216	-
200-300	-	-	-	-	0.2020	-	-	-
Above 300	-	-	-2.8947	-	-	-	-	-
Zilla unpaved road								
Total	-	-	-	-	-	-	-	0.1736
0-30	-0.6623	-	-0.7260	-0.9880	-	-	-0.2353	-
31-60	-0.8630	-	-0.8405	-1.4128	-	-	-	-
Above 90	-	-	-	-	0.4176	-	-	-
Union zilla paved road								
Total	-	-	0.4781	-	-	-	-	-
700-1200	0.2154	-	-	0.2130	-	-	-	-
Union zilla unpaved road								
Total	-	-	-	-	-	-	-	-0.3733
0-400	-	-	-	-	-	-	0.2207	-
400-800	-	-	-0.2948	-	-	-	-	-
Total number of road	-	-	-	-	-	-	-	0.5273
Road structure and existing gap								
Structure number								
0-1500	-	-	-	-	-	0.4056	-	-
1500-3000	-	-	-	-	-	0.3844	-	-
Structure span								
Total	-	-	-	-	-	-	0.3421	-
7000-14000	-	0 .3475	-	-	-	-	-	-
Number of gap								
Total	-0.1249	-	-0.2721	-0.1207	-	-	-0.1169	-
0-300	-	-	-	-	-	0.4652	-	-
301-600	-	-	-	-	-	0.5861	-	-
Span of gap	-	-	-	-	-	-	-	-
0-3000	-	-	-	-	-	-	-	-0.3839
Socio-economic characteristics								
Total household population								
Household number	-	-	-	-	-	0.8931	-	-
Male%	-	-3.4292		-	-0.0217	-	-	-
Female%	-	-	-1.69652	-	-	-	-	-
Female 15+	-	-	-	-	-	-0.6338	-	-
Male% 15+	-0.3737	-	-	-4.44568	-	-	-	-7.9381
Female% 15+	-	8.0593	-	-	-	-	-	-
Economically active male % 15+	0.5551	-5.0057	-	-	0.0096	-	-	0.7921
Economically active female	-	-1.4475	0.0861	-	-	-	-	-

% 15+								
Total active population	-	-	-	-	-	0.2276	-	-
Permanent insolvency %								
Total %	-	-	-	-	-	-0.01305	-	-
15%-30%	-	-	-	-	-	-	-0.2544	-
30-45%	-	-	-	-	-0.5054	-	-	-
Above 45%	-	-	-	-	-0.9315	-	-	-
Temporary insolvency %								
Total%	-0.4979	-0.7210	-0.9905	-0.5439	-	-	-	-
19%-36%	-.39294	-	-0.91908	-0.5274	-0.4420	-	-	-
0%-20%	-	-	-	-	0.4747	-	-	-
Equivalent expense %								
20%-40%	-	-	-	-	0.3251	-	-	-
Solvent%-								
Total	-	-	-0.5884	-	-	-	-	-
15-28	-	-	-	-	-	-	-	-0.3384
Above 28%	-	-	-	-0.6640	-	-	-	-
Saving %								
Above 28%	-2.1516	-2.6418	-	-10.4912	-	-	-	-
Household size								
Above 5	-0.33493	-	-	-0.3605	-	-	-0.2679	-0.5988
Sex ratio								
Total	-	5.9557	-	-	-	-	-	-
90-110	-	-	-	-	-8.5624	-	-	-
110-120	-	-	-	-	-5.4408	-	-	-
Above 120	0.5529	-	-	-	-	-	-	-
Male sex	-	-	4.6908	-	-	-	-	-
Female sex	-	-	-3.2442	-	-	-	-	-

4.5 Discussion

To find out our significant variables we have included district wise road characteristics (type wise length), road structures (culvert, bridge etc, their length, span), existing gap and socioeconomic characteristics in our model in order to determine their complex effects. Discussions of the significant control variables in the final models are presented using three categories: road length, road structure and existing gap and socio-economic characteristics.

4.5.1 Road characteristics

Several road characteristics were considered as significant variables and their effects would be briefly discussed. A number of studies (Amoros et al., 2003; Henning-Hager, 1986; Graham and Glaister, 2003) have examined the effect of the length of road (total or functional class wise) on the expected crash rate. The expected effect of length of a road on crashes may vary depending on the type of road. For example, higher functional classes of roadways such as national and regional road have better design standard and thus may have fewer number of crashes. On the contrary, these roads are designed for higher speed which may be associated with more casualties.

In the model we consider two types roads for all classes of roads (national, regional, zilla and union zilla road) .they are paved road and unpaved road. Also consistent with other studies(Amoros et al. 2003; Hager, 1986; Graham and Glaister, 2003), our study has found that increasing the length of paved road per year has an adverse effect on the number of crashes. These result is almost true for most the year for both continues and categorical variables. The highest elasticity in the road accident due to higher road length is 1.0054 which means one percent increase in road length causes 1.005% increase in road accident. There were some exceptional cases, where increases in road length cause decreases in road causalities. The variables which reduce number of accident for their increase value are national paved 0-40km, Zilla paved (2006, 2007, 2008) and road above 300km (2005). This may happened due better road management plan during road construction which may reduce the possibility of road accident. The highest elasticity in the road accident which decreases road accident due to increase in road length is -2.8974; which indicates that increase in one percentage in road length causes 2.8974% decrease in crashes.

On the other hand it is seen that increased unpaved road length associated with decreased road accident. It is almost true for all classes of road for all years. The highest elasticity in the road accident due to higher road length is -2.89 which means one percent increase in road length causes 2.89% decrease in road accident.

4.5.2 Road Structure

The presence of road structure in each district has been included in the model to estimate the effect of different levels of road structure supply on the expected crash rate. Here number of road structure and their span size has been considered. In here road structure means bridge, culvert etc. Also the number of existing gap and their span also included in the model to find out their effect on crash. For both case categorical and continuous variables are included in the analysis. From the analysis it is found that increase number of road structure enhances the probability of number of accident. Increase road structure, increases vehicle transportation which may result increase in crashes. The highest elasticity in the accident due to increase in road structure is 0.4056 which shows that one percent increase in road structure increase the likelihood of crashes by 0.4056%. This effect is also same for their span. The highest elasticity in the accident due to increase in span size is 0.3475.

From the analysis different effects due to the variation gap number on crashes is found. Gap in a continuous road reduce traffic flow, which may decrease number of crash. The highest elasticity in the accident due to enhance gap number is -0.0014 which indicates that one percentage increase of gap number result in a 0.0014% decrease in crash. It is also seen that increase number of gap, increase the chances of road accident. The highest elasticity in the accident due to increase in gap number is -0.2703 which means one percent increase in road structure increase road accident by 0.2721%.

4.5.3 Demographic and Socioeconomic Characteristics

In parallel with road characteristics, various socio-economic and demographic factors are also examined in this study. Here different aspect of socio demographic feature like total number of house hold , percentage of male and female , sex ratio etc were considered. For economic features like economically active population, percentage of economically active male and female, different earning group (permanent solvency, temporary solvency, saving, solvent etc) were examined. In the model development both continuous and categorical variables are considered.

Number of house hold population has important role in crash number. From the model it is found out that in the case of crashes in 2008, the elasticity of this variable is 0.8931 which indicates that one percentage of increase in household number will increase the probability of crashes by 0.8931%. Total percentage of male and female have significant impact on accident numbers. For male percentage the highest elasticity in the case of crashes in 2004 is -3.4292. It means that one percentage of increase in accident number will reduce the number of accident by 3.4292%. While for female population the elasticity in the case of crashes in 2005 is -1.6965 which means that one percentage of increase in accident number will decrease 1.6965 percentage of accident number.

Here we also examined the total sex ratio of the population. Sex ratio has both positive and negative effects on crash frequency. The highest elasticity of this variable is 5.9564 which represents that one percentage of increase in sex ratio will result in 5.9564% increase in crash frequency. Also for the year 2007 the elasticity was -8.5624 which indicates that one percentage of increase in sex ratio will cause 8.5624% decrease in the number of accidents. We also found out the elasticity for male and female. For male, the elasticity was 4.6908, which shows that one percent increase in male will increase the crash frequency by 4.6908%. For female sex, the elasticity was -3.2442,which means one percent increase in female sex will cause decrease the crash frequency by 3.2442%. This may be because of more risk taking behavior of the males than the females.

For economic condition, population of age 15+ and economically active population of age 15+ were considered. In the case of population of age 15+ it was seen that, increase in male population reduces the number of crashes. On the other hand, the situation is vice versa for the female population. The highest elasticity for male% 15+ is -7.9381 which indicates that one percent increase in male% 15+ will reduces the likelihood of the crash frequency by 7.9381. The highest elasticity for female% 15+ is 8.0593 which means one percent increase in male% 15+ will cause increase the crash frequency by 8.0593. For economically active population of age 15+ it was seen for almost all the cases that increase in male population enhances the number of crashes. On the other hand, the situation is totally opposite for the female population. The highest elasticity for active male% 15+ is which represents that one percent increase in male% 15+ will increase the probability crash frequency by 0.7921. The highest elasticity for female% 15+ is -1.4475 which means one percent increase in male% 15+ will decrease the crash number by 1.4475.

In our Study, four economic stages were considered. They are permanent insolvency, temporary insolvency, equivalent expense and solvent. For all the cases of permanent insolvency it was seen that, increase in permanent solvency decreases the number of accident. The highest elasticity for percentage of permanent insolvency was -0.9315 which means one percent increase in permanent insolvency will decrease the crash frequency by 0.9315.

For all the cases of temporary insolvency it was seen that, increase in temporary insolvency decreases the number of accident. The highest elasticity for percentage of temporary insolvency was -0.9905 which indicates that one percent increase in temporary insolvency will result in a 0.9905% in number of crashes. An exceptional case was seen in the year 2007. In 2007, for temporary insolvency the elasticity was 0.4747 which means one percent increase in temporary insolvency will increase the crash frequency by 0.4747%.

For all the cases of equivalent expense it was seen that, increase in equivalent expense declines the number of accident. The highest elasticity for percentage of equivalent expense was 0.3251 which represents one percent increase in permanent insolvency will increase probability the crash frequency by 0.3251%.

For solvent, in all situations it was seen that, increase in percentage of solvent population reduces the likelihood of road accidents. The highest elasticity for percentage of solvent population was -0.6640. It represents that one percent increase in permanent insolvency will reduce the chances of crash frequency by 0.6640%.

4.6 Summary

The foremost objective of the crash frequency analyses presented in this chapter is to investigate the outcome of different district wise factors on the frequency of crashes and for this purpose individual district taken as unit which has not been explored much in the past. To developed statistical model, negative binomial method was considered. Year 2004, 2005, 2006, 2007, 2008, 2009 and 2010 data were considered to develop the model. Here variables were calculated based on total number of crashes as well as the crashes for each individual year.

Mainly road characteristics (type wise length), road structures (culvert, bridge etc, their length, span), existing gap and socioeconomic and demographic characteristics were considered to find their effect on number of accident. The effects of these control variables are less consistent across the models. After the analysis, different kinds of outcome were found. Our result indicates that increase in paved road length enlarge the probability of road crashes for most of the cases.

While increase in unpaved road reduces the likelihood of crashes. This outcome may be because of the increase in road traffic due to increased number of paved road, which may introduce amplification in road crashes.

For road structure, it was found that rise in bridge, culvert etc number increases vehicle transportation which may causes increase in crash frequency. While gap number has both positive and negative effects on number of crashes.

Socioeconomic and demographic characteristics have important role influencing the number of crashes. From the analysis it was found that increase in household population, result in increase accident number. We also observe that male population enhances the probability of number of crash, while female population decrease the accident number. This situation is vice versa for economical condition. From the analysis, for most of the cases, it was observed that increase in active male population lessen the likelihood of number of crashes, while increase in active female population increases the chances of crashes. Considering different economical stage such as permanent insolvency, temporary insolvency, solvent etc, it was analyzed that increase population reduce the accident number.

CHAPTER FIVE: CONCLUSION

5.1 Background and Rationale

Traffic accidents are a ‘global tragedy’ with ever-rising trends in fatalities and injuries in the developing countries. Road trauma has now been recognised as one of the significant diseases of industrial societies and is an increasing public health economic issue in developing countries. According to the ‘World Report on Road Traffic Injury Prevention-2004’, worldwide an estimated 1.2 million people are killed in traffic accidents each year and as many as 50 million are injured. Projections indicate that these figures will increase by about 65% over the next 20 years unless there is new commitment to prevention. Furthermore, traffic deaths are predicted to increase by 83% in low income and middle income countries and to decrease by 27% in high income countries. Of total 1.2 million deaths, over 80% of traffic fatalities occur in so called developing and emerging countries, even though these countries account only about ⅓ of the total motor vehicle fleet. Accident rates in developing countries are often 10-70 times higher than in developed countries. Whereas traffic accident situation is slowly improved in the industrialised societies (e.g. Australia, USA, UK etc.), most developing countries face a worsening situation. It is expected that over the next ten years developing countries will experience the alarming increase in traffic injuries. Thus, as a developing country, Bangladesh has to give emphasis to adopt countermeasures for reducing road traffic crashes and for that a deep analysis on crashes and related factor is very much necessary.

It is clear that factors related to road infrastructure, land use and socioeconomic and demographic features have significant impact on crash occurrence. That's why the objectives of the study is to examine the effect of district wise road infrastructure , land use & demographic features on crash occurrence .The unit will be taken as individual district which has not been explored much in the past.

5.3 Main Findings of the Study

Three categories of factors were considered: road characteristics (type wise length), road structures (culvert, bridge etc, their length, span), existing gap and socioeconomic and demographic characteristics. These factors were used to develop statistical models. The model out comes will tell which factors increase/decrease the district level crash frequency.

Our result indicates that when the length of the paved road increases, it enhances the chances of crashes. On the contrary, when the length of the unpaved road increases, it reduces the likelihood of crashes.

For road structure (i.e., bridge and culvert), it was found that when the bridge and culvert number increase, the chances of crashes will increase. On the other hand, gap number has both positive and negative effects on crash occurrence.

Socioeconomic and demographic characteristics have significant role influencing the quantity of accident. From the study it was found that increase in household population result in the increase of accident number. It is also obtained that male population increases the crash number, while female population shows vice versa. While incorporating the economic condition on the effect of gender, this situation changes. It is observed that increase in active male population decreases the likelihood of crashes, while increase in active female population increases the chances of crashes From the model results it is seen that different economical stages such as permanent insolvency, temporary insolvency, solvent etc. reduces the accident number.

5.4 Limitation

Conducting this thesis, it was found out that it had quite a number of limitations. The study could not include some important factors which might increase the accident numbers. This study did not include the effects of the registered vehicles and licensing information in each district due to lack of official information. We were also unable to include district wise number of bus stations and rail stations. This lack of information was mainly because of the inadequate government data base.

5.5 Recommendation

The study will provide policy makers as well as transportation engineers and planners with important information for adopting countermeasures to reduce crash frequency. The information is particularly useful for developing countries like Bangladesh, which has experienced great growth in road accidents over the past decade because of increased population, increased traffic, less development in roads and highways sector etc.

The results of this study are used in developing the future forecasting model for predicting crash frequency for future years. From this model we can predict the approximate number of accidents that may happen subsequent years. This information will be very useful in taking proper steps and countermeasures to reduce crashes in Bangladesh.

REFERENCES

Agresti, A. 2002. Categorical data analysis. New York, NY: John Wiley and Sons, Inc.

Alam, M. J. B., Nury, A. H., Farzana, S. Z., Zafor, M. A. 2012. Study on frequency analysis of Sylhet city's road accident. *International Journal of Engineering & Technology*, 4(2), 557-683.

Amin, S. 2011. The relationship between speed and accidents: a literature review, *Proceedings of the Conference on Engineering Research, Innovation and Education, Bangladesh*.

Amoros, E., Martin, J. L., Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. Accident Analysis and Prevention, 35, 537–547.

Atini, C., & Maisey, G. 1990. Effectiveness of public posting of speeding information on an urban road: Accident analysis, Western Australian Police Department, Research and Statistics Section, Perth, WA, 17 p.

Bailey, P. K. (1987). Vehicle deployment options for traffic enforcement operations: The effect on road user behavior. *Journal of Behavioral Sciences*, 46(2), 111-123.

Bangladesh Road Transport Authority (BRTA).

Easa, S. M., and Mehmood, A. 2008. Optimizing design of highway horizontal alignments: New substantive safety approach. *Computer-Aided Civil and Infrastructure Engineering*. 23 (7), 560-573.

Graham, D. J., and Glaister, S. 2003. Spatial variation in road pedestrian casualties: The role of urban scale, density and land use mix. *Urban Stud*, 40(8), 1591–1607.

Hager, H. U. 1986. Urban development and road safety. *Accident Analysis and Prevention*, 18(2), 135-145.

Haque, M. S., Meraz, K. G. 2005. Study of hazard index on Amberkhana to Bandar under Sylhet city corporation. *Journal of Civil Engineering, The Institution of Engineers Bangladesh*, CE36, 15-23.

Hossain, M. 2006. A study on pedestrian accidents based on the injury surveillance (IS) data: Thailand's case. *Proceedings of the Confernece on Road Safety in Developing Countries*, 46-51.

Hoque. M. M .(2008), The prevention of road traffic accidents: An overview of behavioural modification techniques. *Journal of Scientific Research*, 22(2), 85-97.

Hoque. M. M. 1998. The accident problem in rural highways. *Journal of Civil Engineering, The Institution of Engineers, Bangladesh*, 26(1), 51-59.

Hoque, M. M., Sarkar, S., Rahman, M. F., Ashrafuzzaman, M. 2006. Hazardous road locations (HRL) on national highways in Bnagladesh. *Proceedings of the Conference on Road Safety in Developing Countries, Bangladesh*, p 229-237

Hoque, M. M., Sarkar, S., Rahman, M. F., Talukdar, M. A. 2006. Heavy vehicles safety in Bangladesh: Accident invovlement, driversbehavior & understanding road safety. *Proceedings of the Conference on Road Safety in Developing Countries, Bangladesh*.

Hoque, M. M., Azad, A. K., Mahmud, H. M. I., Mahmud, S. M. S., Sarkar, S. 2006. The risk of children in road traffic accidents in Bangladesh. *Proceedings of the Confernece on Road Safety in Developing Countries, Bangladesh*.

Hoque, M. M. 2003. Injuries from road traffic accidents: A serious health threat to the children. *Proceedings Published on the World Health Day*.

Jahid, G. H. 2009. Incentive programs for promoting safer driving. *Journal of Civil Engineering, The Institution of Engineers, Bangladesh*, 35(1), 51-59.

Jonah, B. A. 2004. Accident risk analysis and prevention. *Proceedings of the conference on Road Safety, IRC, India*.

Manan, W. N. B. W., 2011. Accident prediction model at un-signalized intersections using multiple regression method.

Miaou, S. P. 1994. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis and Prevention*, 26(4), 471−482.

Joshua, S., and Garber, N. 1990. Estimating truck accident rate and involvement using linear and Poisson regression model. *Transportation Planning and Technology*, 15,41−58.

Milton, J., and Mannering, F. 1998. The relationship among highway geometrics, traffic related elements and motor-vehicle accident frequencies. *Transportation*, 25(4),395−413.

Rab, M. A. 2006. Rural road safety problem & prospects in Bangladesh context. *Proceedings of the Confernece on Road Safety in Developing Countries, Bangladesh*, 211-216.

Sheikh, M. M. R., 2009. A Statistical Analysis of Road Traffic Accidents and Casualties in Bangladesh.

Swedish Road and Traffic Research Institute. *Speeds as a function of tolerance limit, penalties and surveillance intensity, Linköping*, VTI 1989, 41, VTI rapport 337.

Vogt, A. 1999. Crash models for rural intersections: four-lane by two-lane stop controlled and two-lane by two-lane signalized. US Department of Transportation, Federal Highway Administration Report #FHWA-RD-99-128.

World Report on Traffic Injury Prevention. 2004.

Zhou and Sisiopiku. 1997. Injuries from road traffic accidents: A serious health threat to the children. *Proceeding Published on the World Health Day*, 2003.