Significant Hazard Factors Causing the Pervasiveness of Hepatitis B and C in Punjab: using Statistical Analysis

By

Azeem Iqbal

Reg. No. 31-FBAS/MSST/S13

Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan

2015

1- Multivariate analysis
2 - Statistical analysis
3 - Hepatitis B
4 - Hepatitis C
5) [illegible]

Significant Hazard Factors Causing the Pervasiveness of Hepatitis B and C in Punjab: using Statistical Analysis

By

Azeem Iqbal

Supervised by

Dr. Irshad Ahmad Arshad

Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan

2015

Significant Hazard Factors Causing the Pervasiveness of Hepatitis B and C in Punjab: using Statistical Analysis

By

Azeem Iqbal

Reg. # 31-FBAS/MSST/S13

A Dissertation
Submitted in the Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE
IN
STATISTICS

Supervised by

Dr. Irshad Ahmad Arshad

Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan

2015

بسم الله الرحمن الرحيم

Certificate

Significant Hazard Factors Causing the Pervasiveness of Hepatitis B and C in Punjab: using Statistical Analysis

Submitted By

Azeem Iqbal

A DISSERTATION SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF THE ***MASTER OF SCIENCE in STATISTICS***

We accept this dissertation as conforming to the required standard.

1. Prof. Dr. Rana Abdul Wajid
(External Examiner)

2. Dr. Muhammad Akbar
(Internal Examiner)

3. Prof. Dr. Irshad Ahmad Arshad
(Supervisor)

4. Dr. Rehmat Ellahi 26/6/15
(Chairman)

Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan

2015

Dedication

To my Father (Perveiz Iqbal late),

To my Mother & Wife,

For the endless support and patience.

Forwarding Sheet by Research Supervisor

The thesis entitled **"Significant Hazard Factors Causing the Pervasiveness of Hepatitis B and C in Punjab: using Statistical Analysis"** submitted by **Azeem Iqbal** (Registration # 31-FBAS/MSST/S13) in partial fulfillment of M.S. degree in Statistics has been completed under my guidance and supervision. I am satisfied with the quality of his research work and allow him to submit this thesis for further process to graduate with Master of Science degree from Department of Mathematics and Statistics, as per IIU Islamabad rules and regulations.

Dated 11/05/2015

Dr. Irshad Ahmad Arshad,
Associate Professor
Department of Mathematics & Statistics,
International Islamic University, Islamabad.

Acknowledgements

All the praises to Almighty **ALLAH** the most gracious, the most merciful and the creator of all the creatures. Thanks to **ALLAH** and his **Prophet (P.B.U.H)**. There are few people who made this journey easier for me with their encouragement and ideas.

I express my gratitude to my supervisor **Dr. IRSHAD AHMAD ARSHAD** for his regardless and inspirational efforts and moral support throughout my research carrier. His sound advices and lots of good ideas were very helpful to me. May **ALLAH** bless him with all kinds of happiness and success in his life and may all his wishes come true. I also want to express my unbounded thanks to all the faculty members of Department of Mathematics and Statistics. I also acknowledge **Pakistan Medical Research Council** for the provision of the Data about the risk factors of Hepatitis B & C in Punjab.

My Deepest gratitude to my family who are the real pillars of my life. They always encouraged me and showed their everlasting love, care and support throughout my life. The continuous encouragement and humble prayers, support from my mother and my wife is unforgettable. I am also thankful to my friend **Mr. Saqib Munawar.** I am grateful for all the emotional support and care he provided.

Finally, I express my sincere appreciation to all the helpful staff of the Mathematics and Statistics Department IIUI. I would like to thanks to everybody who was important in the successful realization of this thesis as well as expressing my apology to those that I could not mention.

Azeem Iqbal

Declaration

I hereby declare that this thesis, neither as a whole nor a part of it, has been copied out from any source. It is further declared that I have prepared this dissertation entirely on the basis of my personal efforts made under the supervision of my supervisor **Dr. Irshad Ahmad Arshad.** No portion of the work, presented in this dissertation, has been submitted in the support of any application for any degree or qualification of this or any other learning institute.

Signature: ____________________

Azeem Iqbal
M.S. (Statistics)
Reg. No. 31-FBAS/MSST/S13
Department of Mathematics and Statistics,
Faculty of Basic and Applied Sciences,
International Islamic University Islamabad, Pakistan.

Contents

List of Tables

List of Figures

Acronyms

AUC	Area under the curve
CI	Confidence Interval
Df	Degrees of freedom
FBS	Federal Bureau of Statistics
GLM	Generalized Linear Model
Ln	Natural Logarithm
MLE	Maximum Likelihood Estimation
OR	Odds Ratio
PSUs	Primary Sampling Units
PMRC	Pakistan Medical Research Council
ROC	Receiver Operating Characteristic
SSUs	Secondary Sampling Units

Abstract

Hepatitis B & C viral infections are one of the most prevalent health hazards in Pakistan. Pakistan has been avowed "Cirrhotic State" in universal health circles. These viruses have appalled the developing countries where illiteracy and poverty may chip in to the increase the risk and cause of Hepatitis viral infections. The major objective of the research study is to determine the significant hazard factors which are related to the pervasiveness of Hepatitis B & C among the people of Punjab, Pakistan by using multivariate logistic regression. Data on the Hazard factors of hepatitis B & C in Punjab were obtained from PMRC, Islamabad. In this research study, the factors were divided as Demographic, Socioeconomic & Medical hazard factors. Wald test statistic was used to assess the individual worth of parameters with the response variable. Hosmer-Lemeshow and R.O.C. curve were used to estimate the fit of logit regression model. The Multivariate model of Hepatitis C (for females) demonstrated that age, re-use syringes, ear and nose piercing, jaundice's history, ever-married and family history of hepatitis were the significant hazard factors associated with the Hepatitis C (for females). However, the dominant hazard factors related to Hepatitis C (for males) were age, barber shaving, re-use syringes, jaundice's history, share cigarettes/hookah, ever-married, family history of hepatitis and tattooing/acupuncture. The Multivariate analysis of Hepatitis B (for females) revealed that age, ear and nose piercing, jaundice's history, ever-married and family history of hepatitis were the momentous hazard factors concomitant with the Hepatitis B (for females). In spite of this,

the substantial hazard factors interrelated with Hepatitis B (for males) were age, barber shaving, jaundice's history, ever-married and family history of hepatitis. Mutual hazard factors of hepatitis B & C for both genders were age, jaundice's history, ever-married and family history of hepatitis. Barber shaving in males played a crucial role in the pervasiveness of both Hepatitis B & C, whereas, ear and nose piercing was the major cause of extensiveness of Hepatitis in females. Lack of awareness about the Hepatitis played a key role in the extensiveness of liver cancers. The need of an hour is to educate people about the factors that are responsible for the commonness of Hepatitis.

Chapter 1

Introduction

Viral hepatitis is the inflammation or infection of the liver. Hepatitis is documented as a disease causing liver infection all over the world. It is regarded as one of the most prevalent health hazards. Hepatitis may be acute causing liver infection for less than 6 months and chronic for a longer period of time. The rate of viral hepatitis is increasing due to various reasons. Scientific knowledge has enabled us identify different types of viral hepatitis, such viruses are accountable sources of acute or chronic infectivity and swelling of the liver leading to a severe problem of the public health worldwide. Hepatitis-B and Hepatitis-C viral infections are not only the principal reasons of chronic hepatitis but also the prominent cause of cirrhosis. These viral infections are also the main causes of heavy sickness and mortality. (Hafeez et al).

Pakistan consumes 2.4 billion syringes annually, which is the highest rate among the syringes consuming countries. Most of the Pakistani people have been affected by viral hepatitis, due to suspicious quality or reused syringes. Consequently, Pakistan has been avowed "Cirrhotic State" in universal health circles. Therefore, the main origin of Hepatitis in Pakistan is commonly due to reuse and the poor quality syringes. (W.H.O. 2012, ISLAMABAD)

Hepatitis has emerged as a foremost health hazard among the developing countries like Pakistan due to which it is one of the worst distressed and tormented place. Both Hepatitis-B & Hepatitis-C are widespread infectious viruses that affect great number of individuals and are major reasons of Chronic Liver infection in

Pakistan as well as worldwide. The significant factors linked with the Hepatitis B & C viruses may be prevented to reduce the risk factors. The vulnerability factors for the pervasiveness of Hepatitis B & C are alike. Thus the perseverance or objective of our research study is to determine the significant hazard factors associated with the pervasiveness or transmission of the Hepatitis B & C in Punjab, Pakistan.

1.1 Hepatitis B Virus

Hepatitis-B virus is possibly a life hazardous liver infectious disease. It is significant health threat for the world. It is a viral harmful disease that assaults the liver. The Hepatitis-B virus is passed on through blood contact, semen or other body liquids of the septic one. An infected woman can transmit Hepatitis-B virus to her infant at the time of birth. Concrete precautionary measures are requisite to build up a policy to educate the people regarding the risk factors of Hepatitis B virus. Hepatitis-B is avoidable with available effective & save vaccine. It was 1st ever vaccine against the lethal human cancer, which had been developed in 1982.

It is expected that approximately two billion or one out of three (two thousand million) persons globally have been effected with the Hepatitis-B virus and about 400 million individuals have chronic liver infections related to Hepatitis-B virus. About one million persons expire per year due to Hepatitis-B virus infection. In Pakistan, more than 6 Million people are infected with Hepatitis-B virus which is approximately 3% of the Pakistan's population. Thousands of people in Pakistan pass away due to Hepatitis-B per annum. (W.H.O., 2013)

Hepatitis B is hazardous for the reason that it is "silent viral infection". It can aggravate the liver of patients without their awareness about it, which may result in the chronic liver infection that leads to cirrhosis or liver cancer. Some people are in good health even in the presence of chronic Hepatitis-B, but these people never strive for any medical care. This can be perilous because such people may be the risk factor for unknowingly transmission of virus to others.

1.2 Hepatitis C Virus

Hepatitis C is the universal, infectious and transmittable viral disease which is caused by Hepatitis C virus and a possible reason of prevalence of disease and death in future. The Hepatitis-C virus is generally prevalent when infects a person's blood who comes in contact with another person at risk. It is gradually mounting liver disease that rigorously or ruthlessly influences the cells of the liver that can result in cirrhosis or liver cancer. Hepatitis-C has turned out to be a foremost health suffering in developing countries like Pakistan.

According to the report of World Health Organization (W.H.O., 2013), Pakistan has been ranked 2nd having greater rates of liver infective diseases in the world. Hepatitis-C is the fastest emergent cause of liver cancer in Pakistan. We are unable to understand the underlying aspects of treatment of Hepatitis-C's correctly, which is highly disastrous. The number of people contaminated with Hepatitis-C virus in Pakistan are approximately 10 Million, which is 6% of the Pakistan's population. The principal means of widespread of Hepatitis in Pakistan is associated with unsafe injection and unhygienic medical treatment. Approximately 3% (i.e. 200 Million) of total population of the world has been influenced with Hepatitis-C virus, nearly 0.5 Million people pass away due to Hepatitis-C per annum.

1.3 Research Objectives

The main objectives of this study are:

- To determine the significant hazard factors which are related to the pervasiveness of Hepatitis B & C among the people of Punjab, Pakistan.

- To measure an association between Hepatitis B & C and hazard factors independently i.e. odds ratio (Bivariate Analysis)

- To construct predictive model for both males and females separately for Hepatitis B & C using Multiple Logistic Regression model on the basis of those significant hazard factors. (Multivariate Analysis)

- To determine and discuss certain worthwhile socioeconomic, demographic and medical hazard factors of Hepatitis B & C.

Chapter 2

Review of Literature

Machado *et al* (2013) estimated the pervasiveness of hazard factors related to the Hepatitis B viral infection in senior inhabitants having age 60 years or greater of Tubarao city located in Brazil. This cross sectional study involved 820 persons which were selected by simple random sampling technique. Fisher exact & Pearson's Chi square tests were applied to compare proportions. Bivariate analysis revealed that the Hepatitis B was related with acupuncture therapy, ≤ 4 schooling year and age greater than 67 years. In the Multivariate analysis, hazard factors that were associated with Hepatitis B were male gender, ≤ 4 schooling year, acupuncture and marital status. To recognize the factors which were independently associated with Hepatitis B, the variables in bivariate-analysis were assessed by using logistic regression. Hosmer & Lameshow test was used to check the adequacy of the final multivariate logit regression model.

Abbasi *et al* (2013) investigated the hazard influences of Hepatitis B & C in Muzaffarabad, a city of Azad Jammu & Kashmir. The objective of this study was to determine the pervasiveness of Hepatitis in the existence of the possible hazard influences. And to examine the association between gender, education, area with Hepatitis. The questionnaire was organized to accumulate data from 400 persons visiting Abbas Institute of Medical Sciences Muzaffarabad and Sheikh Khalifa Bin Zaid al nabyan Hospital Muzaffarabad. It was analyzed that reuse of syringes, ear piercing, age group 15-20, low level of education and sexual relations with the influenced persons were major hazard factors of Hepatitis B and C. Odds ratio were calculated to check the association between Hepatitis C and the risk factors. The relationship of Hazard factors with the blood transfusion is assessed by using chi-square test. The multivariable Logistic Regression, Kruskal-wallis H-test & Mann Whitney U-test were used to find the major factors of Hepatitis B and C.

Qureshi *et al* (2010) identified the significant peril influences that were related to the pervasiveness of Hepatitis B & C. This study was based on the national survey of Pakistan, which was conducted in 2007-2008. The intention of the study was to check the national predictions regarding Hepatitis B & C and to assess the important risk factors which were significantly related to the Hepatitis B & C. According to the survey, the prevalence rates of Hepatitis B & C in Pakistan were 2.5% and 4.9% respectively. The probable hazard factors of Hepatitis B & C for the dominance of disease were reuse of syringes, shaving from barber, sharing tooth brush, cigarette & hookah, ear or nose piercing and tattooing. Simple logit regression model (i.e. Odds Ratio) was implemented to determine the degree of association of Hepatitis with each hazard factor.

Qureshi *et al* (2008) analyzed the hazard factors which were concomitant with the widespread of Hepatitis B & C in male patients visiting Gastroenterology/ Hepatology unit of PMRC (Pakistan Medical Research Council) and Jinnah Medical Centre, Islamabad. This was a case-control study consisted of 1773 male individuals in which 1050 patients were cases and 723 were controls. A questionnaire was developed to collect data from 1050 male patients suffering from Hepatitis B and C visiting Gastroenterology/Hepatology unit of PMRC and Jinnah Medical Centre, Islamabad. 723 controls were also selected for the data collection from blood bank of

Jinnah Medical Centre, Islamabad. It was found that member of family suffering from Hepatitis, Treatment by Dentist, Blood Transfusion and the Use-of-Injections were some important hazard factors those were related to the transmission of Hepatitis B & C in the male patients. Multivariable Logit Regression model was applied to analyze substantial hazard factors of Hepatitis B and C. Odds ratio were calculated to investigate the association between Hepatitis and hazard factors.

Ghias *et al* (2010) determined the risk factors for Hepatitis C in urban/rural patients independently using Logistic Regression Analysis. This was a hospital based study in which three main hospitals of city Lahore were selected i.e. Sheikh Zayed, Mayo and Jinnah hospitals. Numbers of urban patients were 185 in which controls=59 & cases=126. Numbers of rural patients were 215 in which controls=61 & cases=154. Urban logit regression model showed that married patients, surgical operations, uneducated mother, member of family suffering from Hepatitis & road accidents were significant hazard factors of Hepatitis-C. Whereas rural logit regression model showed that age of patients, ever had jaundice and Barber's shaving were significant hazard factors. Ever Married patients & member(s) of family suffering from Hepatitis were common hazard factors in urban & rural populations. Odds ratio was used to check the association between Hepatitis and hazard factors. Wald test statistic was applied to analyze the individual significance of logit coefficients. Hosmer & Lameshow test was used to test the satisfactoriness of overall fitted model.

Abbas *et al* (2008) identified the significant hazard factors which were related to the widespread of Hepatitis B & C in the rural regions of Sindh, Pakistan. A questionnaire was organized to accumulate data by using systematic sampling technique from 873 persons of Jarwar constituency, a small agrarian town which is

100 k.m. away from Sukkur (a City of Sindh Province). The questionnaire discovered the likely demographic, clinical and community hazard influences related to Hepatitis B & C. In this study, the hazard factors of Hepatitis-B identified were jaundice history, males, no history of vaccination, household history of liver infection and patients having age $\geq$ 16 years. The significant hazard factors of Hepatitis-C were patients having age $\geq$ 16 years, dental treatment, deficiency of vaccination, more than 10 injections in a year, barber shaves, reuse of syringes and ever had liver disease. Multivariable Logit Regression model was used to analyze important hazard influences of Hepatitis B and C. Odds ratio were calculated to investigate the association between Hepatitis and hazard factors.

Akhtar *et al* (2004) assessed the risk factors that are connected with the HCV infection in the male volunteers of blood donors in city Karachi. This was a case control study of 240 patients, from which 160 were Hepatitis-C Negative, while 80 were Hepatitis-C Positive. The data were collected through pretested questionnaire consisted on socioeconomic and demographic characteristics. This study revealed significant autonomous relationships between Hepatitis C virus among donors of blood & past history of hospitalization, amount of injections received in the former 5 years and the type of syringe used when injections received in the past. To determine the uni-variate relations between Hepatitis C virus and assumed hazard factors, odds ratio (O.R.) were evaluated by using simple logit regression method. The final multivariate logistic regression model was selected through backward stepwise logistic regression analysis. Pearson chi square test was used to check the goodness of fit. It was concluded in the final model of multivariate logistic regression that there were greater number of cases than controls who had reported previous hospitalization or had injected multiple injections.

Bari *et al* (2001) identified the risk factors related to the Hepatitis-C in males in the cities of Rawalpindi & Islamabad, Pakistan. This was a case-control study consisted of 237 patients, out of which 57 cases and 180 controls were interviewed from nine hospitals of Rawalpindi-Islamabad. A well thought-out questionnaire was prepared to collect the data, which was further examined by logistic regression. The results showed that history of therapeutic injections and daily face shaves & armpit shaves by Barbers were the significant hazard factors for Hepatitis C virus by using Logistic Regression Analysis. It was concluded that men are at greater risk of getting Hepatitis C virus infection in this world due to non-sterile or contaminated razors used by the barbers. Odds ratio were calculated to check the association between Hepatitis C and the risk factors. Multivariable Logistic Regression Method was used to find out the independent relationship between Hepatitis C & each risk factor.

Ghias & Pervaiz (2009) revealed that History of blood transfusions, History of hospitalization, Tattooing, Family history of Hepatitis, Surgical operation were significant hazard factors of Hepatitis C virus by using Logit Regression Model. This was a case-control study of 400 patients, out of which 119 were Hepatitis-C Negative (Controls), while 281 were Hepatitis-C Positive (Cases). A questionnaire was developed for the data collection from three main hospitals of city Lahore i.e. Sheikh Zayed, Mayo and Jinnah hospitals. Odds ratio were used to check the association between Hepatitis and hazard factors. Wald test statistic was applied to analyze the individual significance of logit coefficients. Hosmer & Lameshow test was used to test the satisfactoriness of overall fitted model. Multiple linear logit regression model was used to envisage the risk of Hepatitis-C in the absence or presence of significant hazard factors. To check the degree of association, the values of Cramer's V and Phi Statistics were acknowledged as significant associated factors of hepatitis C.

Sypsa *et al* (2001) identified the hazard factors that were directly concomitant with the Hepatitis B and C of Greek company employees. That cross sectional study was conducted in seventeen Greek companies. The purpose is to identify the pervasiveness of Hepatitis B and C viruses and to analyze the prognostic hazard factors that were linked with the Hepatitis B and C. A well thought-out questionnaire was prepared to collect the data from 1000 company employees. It was investigated that the rate of pervasiveness of Hepatitis B and C in Greek company employees was 19.9% and 2.6% respectively. It was shown that the blood transfusion, age, family ever suffered from chronic hepatitis and job category were significant threats of Hepatitis-C. Whereas age factor, blood transfusion and weakness/lethargy for extended period of time were important hazard factors for the Hepatitis-B. Multivariate Logistic Regression model was selected to analyze predictive hazard factors that were related to Hepatitis B and C. Associations between the response variable and all explanatory variables were determined using Chi-square test, t-test and Fisher's exact test.

Shazi & Abbas (2006) found that the significant hazard factors of Hepatitis-C were less education, more contact with syringes and blood, barber shaves, intravenous drips, blood transfusions in the past and therapeutic injections. The Hepatitis B and C infected patients visiting the Liver Stomach Clinic, Karachi were interviewed. A structured questionnaire was used to collect data. This was a case control study of 148 patients, from which 63 were Hepatitis-C Positive, while 41 were Hepatitis-B Positive & 44 were Negative (Controls). It was concluded that proper awareness should be given to the people about the hazard factors related to the commonness of both Hepatitis B & C.

Abbasi *et al* (2002) determined the associated hazard factors which were blamed for the dominance of Hepatitis B & C. The questionnaire was organized to collect data from 108 persons. This study was carried out in Department of Medicine, Federal Government Services Hospital, Islamabad. It was revealed that the Hepatitis C was more prevalent than Hepatitis B. The key hazard factors were dental surgery from infected instruments, barber shaves, unsafe sexual relation and re-use of infected syringes. The blood transmission due to deficiency of the screening facilities in the past was responsible hazard influence for the commonness of Hepatitis C.

Chapter 3

Research Methodology

3.1 About the Data

Data on the Hazard factors of hepatitis B & C in Punjab are obtained from Pakistan Medical Research Council (PMRC), Islamabad. PMRC had conducted this survey in cooperation with the Ministry of Health, Statistics division & Federal Bureau of Statistics. In Punjab, approximately 3500 houses consist of 23450 households of all genders & ages had been interviewed. PMRC collected the data from all urban and rural areas of all the four provinces of Pakistan. A stratified two stage sample design was adopted for the survey. Enumeration Blocks in urban domain and villages in rural domain were taken as PSUs. Households within sample PSUs have been taken as SSUs. A specified number of households i.e. 20 from each urban and rural sample PSU have been selected with equal probability using systematic sampling technique with a random start.

3.2 Data Analysis

The data were analyzed by STATA (version 12) and SPSS (version 21). The data analysis was done by using both descriptive and analytical methods. In descriptive section, presentation of data and percentage comparisons were calculated. In inferential section, Bivariate and Multivariate analysis were implemented to acquire the significantly associated hazard factors.

3.3 Response Variable

The outcome of Hepatitis-B is represented by the Dummy variable which accepts the value "1" if Hepatitis-B is positive and "0" elsewhere. In the same way, the outcome of Hepatitis-C is represented by another Dummy variable which accepts the value "1" if Hepatitis-C is positive and "0" elsewhere.

3.4 Description of Hazard Factors under Study

In this research study, the factors were divided as Demographic, Socioeconomic & Medical hazard factors. The description of such factors is stated below:

- **Demographic Factors**
    1. Age
    2. Gender
    3. Marital Status
- **Socioeconomic Factors**
    1. Level of Education
    2. Awareness of Hepatitis
    3. Divisions
    4. Type of House Material
    5. Drinking Source
- **Clinical Factors (Medical Risk Factors)**
    1. Family History of Hepatitis
    2. Use of Intramuscular Injections
    3. Type of Syringe Used
    4. Shaving
    5. Sharing Tooth Brush/Miswak
    6. Tattooing
    7. Ear/Nose Piercing

8. Sharing Cigarettes/Hookah
9. History of Jaundice
10. History of Hospitalization

Table 3.1: Coding Scheme of Variables

Sr. #	Variables	Code
1	Gender	0=Female, 1=Male
2	Marital Status	0=Never Married, 1=Ever Married
3	Level of Education	0=Illiterate, 1=Below Matric, 2=Matric & Above, 3=Graduate & Above
4	Awareness of Hepatitis	0=No, 1=Yes
5	Type of House Material	0=Kacha, 1=Pacca, 2= Semi-pacca/kacha, 3= Well furnished
6	Drinking Source	0=Others, 1= piped in Dwelling, 2= Public Tap, 3= Spring/Pound, 4= Tanker,Vendor, 5= Well
8	Family History of Hepatitis	0=No, 1=Yes, 2= Don't Know
9	History of Intramuscular Injections	0=No, 1=Yes
10	Type of Syringe used	0=Don't know, 1= New/disposable, 2= Re-use
11	Shaving	0=None, 1= Home, 2= Barber, 3= Both
12	Sharing Toothbrush/Miswak	0=No, 1=Yes
13	Tattooing	0=No, 1=Yes
14	Ear/Nose Piercing	0=No, 1=Yes
15	Sharing Cigarettes/Hookah	0=No, 1=Yes
16	History of Jaundice	0=No, 1=Yes
17	History of Hospitalization	0=No, 1=Yes

3.5 Statistical Analysis

3.5.1 Odds

Odds is defined as a ratio of the likelihood of the incidence of an event (i.e. π) to the likelihood of non-incidence of that event (i.e. $1-\pi$).

$$\text{ODDS} = \frac{\pi}{1-\pi} \qquad \textbf{(3.1)}$$

π is also termed as the probability of success.

3.5.2 Odds Ratio

The extent of relationship between the response variable and a regressor or the ratio of two odds is also called an "ODDS RATIO". When a comparison between two sets of dichotomous variables are to be made, assume that π_1 & π_2 are success probabilities in these two sets, then an Odd Ratio can be calculated as

$$\text{ODDS RATIO (O.R.)} = \frac{ODDS_1}{ODDS_2} = \frac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)} \qquad \textbf{(3.2)}$$

3.5.3 Applications of Odds Ratio (O.R.)

The principal use of odds ratio is to investigate if an exposure is linked with the disease i.e. response variable. Odds ratios are commonly used in case-control studies. The exposures with the higher Odds Ratio are considered to be greater magnitude. In other words, they increase the risk more.

3.5.4 Interpretation of Odds Ratio (O.R.)

1) O.R. = 1, Exposure does not influence odds of the disease.
2) O.R. > 1, Exposure related to greater odds of the disease.
3) O.R. < 1, Exposure related to lesser odds of the disease.

3.5.5 Calculation of Odds Ratio in Case-Control Study

An Odds Ratio compares the odds of cases exposed to the odds of controls exposed.

Table 3.2: Contingency Table

	DISEASE (CASES)	NO DISEASE (CONTROLS)
EXPOSED	A	B
UNEXPOSED	C	D

$$Odds\ of\ Cases\ Exposed = \frac{amount\ of\ Cases\ Exposed}{amount\ of\ Cases\ not\ Exposed} = \frac{A}{C}$$

$$Odds\ of\ Controls\ Exposed = \frac{amount\ of\ Controls\ Exposed}{amount\ of\ Controls\ not\ Exposed} = \frac{B}{D}$$

$$Odds\ Ratio = \frac{Odds\ of\ Cases\ Exposed}{Odds\ of\ Controls\ Exposed}$$

$$\text{Odds Ratio} = \frac{A/C}{B/D}$$

$$\text{Odds Ratio} = \frac{AD}{BC} \qquad (3.3)$$

3.6 Binary Logistic Regression Model

In a binary logistic regression model, the response variable is binary i.e. "success" & "failure", and the independent variables (i.e. Predictors) are used to model the probability of that response. Special types of regression models have been established to overcome these circumstances. The commonest type of categorical data is the binary data and the illustrious regression model for such type of data is the "*logistic regression model*". Logistic regression is also named as Logistic Classification which is widely used in medicines and bio statistics.

Logistic regression model is a type of G.L.M. (Generalized Linear Model) for response variables where regular multiple regression does not work very well. The distribution of response variable is specified by the probability of success (i.e. π) and failure (i.e. $1-\pi$). It follows binomial distribution with parameters n & π. A logistic regression model produces S shaped curve.

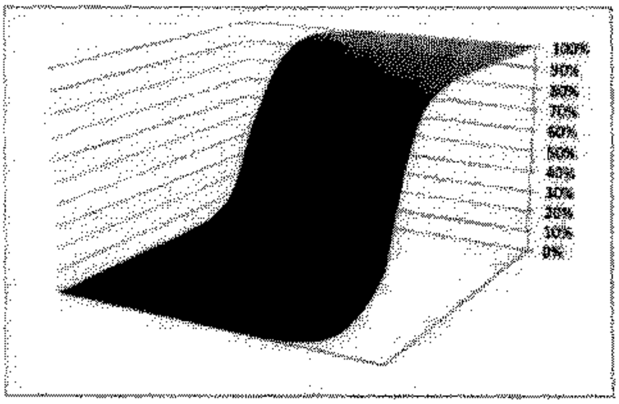


Figure 3.1: Logistic Regression Curve

3.6.1 Logit Model for Single Explanatory Variable

Consider a logistic regression model for one explanatory variable x. "**Logit Model**" is

$$\text{Logit} = \ln\text{ (ODDS)} = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x \qquad \textbf{(3.4)}$$

Where:

"π" is the probability of success, i.e. P(Y=1).

"1- π" is the probability of failure, i.e. 1 – P(Y=1)

"$\frac{\pi}{1-\pi}$" is the odds.

$\ln[\frac{\pi}{1-\pi}]$ is the log odds, or "logit".

The equation (3.4) demonstrates the logit model (i.e., log-odds) is equal to the linear regression. The logit model is linear. Theoretically, the logistic and logit are the designations for transformations. The logit transformation assumes a value π between 0 to 1 and then transforms it to log $[\pi/(1-\pi)]$.

From equation (3.4)

$$\text{Logit} = \ln\text{ (ODDS)} = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x$$

By taking exponential, we have

$$\frac{\pi}{1-\pi} = e^{(\beta_0+\beta_1 x)}$$

$$\pi = e^{(\beta_0+\beta_1 x)}.(1-\pi)$$

$$\pi = e^{(\beta_0+\beta_1 x)} - \pi.e^{(\beta_0+\beta_1 x)}$$

$$\pi + \pi.e^{(\beta_0+\beta_1 x)} = e^{(\beta_0+\beta_1 x)}$$

$$\pi(1+e^{(\beta_0+\beta_1 x)}) = e^{(\beta_0+\beta_1 x)}$$

$$\pi = \frac{e^{(\beta_0+\beta_1 x)}}{(1+e^{(\beta_0+\beta_1 x)})} \qquad \textbf{(3.5)}$$

Or

$$\pi = \frac{e^{(\beta_0+\beta_1 x)}}{(1+e^{(\beta_0+\beta_1 x)})} \cdot \frac{e^{-(\beta_0+\beta_1 x)}}{e^{-(\beta_0+\beta_1 x)}}$$

$$\pi = \frac{1}{(1+e^{-(\beta_0+\beta_1 x)})} \qquad \textbf{(3.6)}$$

The parameter β_1 in equation (3.4) interprets the rate of ascending or descending. According to figure 3.2, $\beta_1 > 0$ means that as x increases, $\pi(x)$ increases and $\beta_1 < 0$ interprets that as x increases, $\pi(x)$ decreases. The rate of change will be steeper (vertical) as $|\beta_1|$ increases. The curve levels the straight line horizontally when $\beta_1 = 0$.

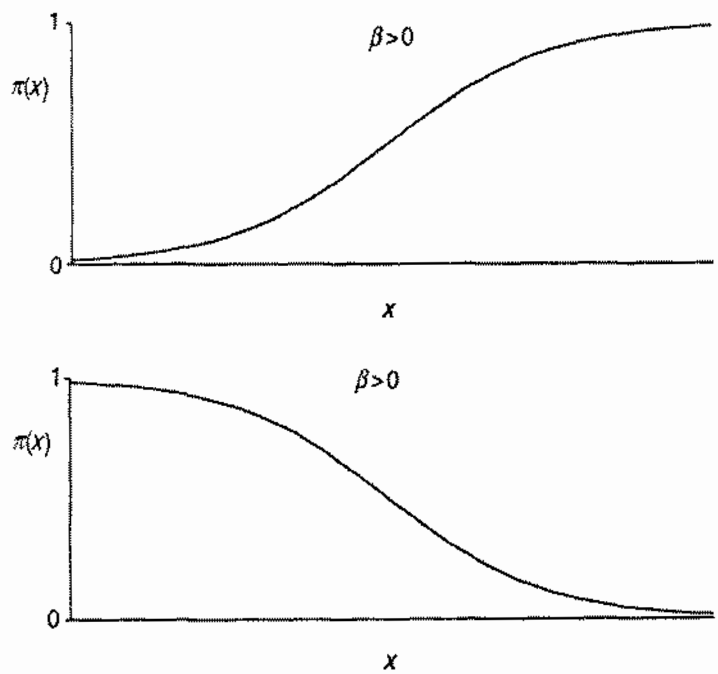


Figure 3.2: Logistic Regression Functions.

3.6.2 Logit Model for Multiple Explanatory Variables

The multivariate logistic regression model is used when we have at least two explanatory variables. Consider a logistic regression model with multiple regressors (i.e. $x_1, x_2, x_3 \ldots x_k$) for a binary response (i.e. Y). "**Logit Model**" for multiple regressors is

$$\text{Logit}(\pi(x)) = \ln(\text{ODDS}) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \qquad \textbf{(3.7)}$$

Where:

"π" is the probability of success, i.e. P(Y=1).

"1- π" is the probability of failure, i.e. 1 – P(Y=1)

"$\frac{\pi}{1-\pi}$" is the odds.

$\ln[\frac{\pi}{1-\pi}]$ is the log odds, or "logit".

Consider the equation (3.7) in matrix form

$$\text{Logit}(\pi(x)) = \ln(\text{ODDS}) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \mathbf{X\beta} \qquad \textbf{(3.8)}$$

By taking exponential, we have

$$\frac{\pi}{1-\pi} = e^{X\beta}$$

$$\pi = e^{X\beta}.(1-\pi)$$

$$\pi = e^{X\beta} - \pi.e^{X\beta}$$

$$\pi + \pi.e^{X\beta} = e^{X\beta}$$

$$\pi(1+e^{X\beta}) = e^{X\beta}$$

$$\pi = \frac{e^{X\beta}}{(1+e^{X\beta})} \qquad (3.9)$$

Or

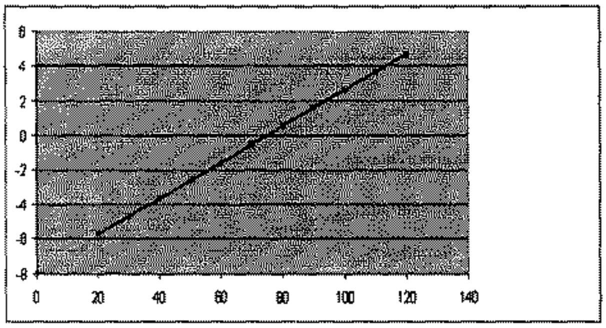
$$\pi = \frac{e^{X\beta}}{(1+e^{X\beta})} \cdot \frac{e^{-X\beta}}{e^{-X\beta}}$$

$$\pi = \frac{1}{(1+e^{-X\beta})}$$

Or it can be written as

$$\pi = \frac{1}{(1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_m x_m)})} \qquad (3.10)$$

$\ln(\frac{\pi}{1-\pi})$



Predictor (X)

Figure 3.3: Logit Curve (LOG ODDS)

3.6.3 Log Odds Transformation (Difference between logits)

Suppose, a binary response variable (Y) has a binary regressor (X). Let x assumes the values 0 and 1 to indicate the two classifications. The logit model for $P(Y = 1)$ is

$$\text{logit}\,[P(Y=1)] = \beta_0 + \beta_1 x \qquad (3.11)$$

The variable x is called *dummy variable*. Table 3.3 illustrates the values of logit at two categories of a predictor.

Table 3.3: Logit values Implied by Dummy Variable in Model, logit [$P(Y=1)$] = $\beta_0 + \beta_1 x$

X	Logit
0	β_0
1	$\beta_0 + \beta_1 x$

The influence on the logit model of varying from $x = 0$ to 1 is

$$\{\beta_0 + \beta_1(1)\} - \{\beta_0 + \beta_1(0)\} = \beta_1 \qquad \textbf{(3.12)}$$

It is apparent from the equation (3.12) that the difference between two logits equals β_1. Also the difference between the two logits is equal to the log odds difference, and as a result, that difference is equivalent to the log of odds ratio (O.R.) between the response variable Y and the predictor X. Hence, exp $(\beta 1)$ is equal to the odds ratio.

From equation (3.12)

$$\text{Logit}(\pi_1) - \text{Logit}(\pi_2) = \beta_1$$

$$\ln(\text{ODDS}_1) - \ln(\text{ODDS}_2) = \beta_1$$

$$\ln\left\{\frac{\pi_1}{1-\pi_1}\right\} - \ln\left\{\frac{\pi_2}{1-\pi_2}\right\} = \beta_1$$

$$\ln\left[\frac{\left\{\frac{\pi_1}{1-\pi_1}\right\}}{\left\{\frac{\pi_2}{1-\pi_2}\right\}}\right] = \beta_1$$

$$\ln[\mathit{ODDS\ RATIO}] = \beta_1$$

or

$$\mathit{ODDS\ RATIO} = \text{Exp}(\beta_1) \qquad \textbf{(3.13)}$$

3.6.4 Interpretation of Logit Coefficient

The logit coefficient (β_1) is the probable increase in the ln(Odds) due to one unit increase in exposure.

Or it can be interpreted as "The exponential function of the logit coefficient (e^{β_1}) is the odds ratio allied with per unit increase in exposure".

3.6.5 Difference between Linear & Binary Logistic Regression.

In Logistic Regression, the response variable is categorical, i.e. binary. After applying the logit transformation, logistic regression implement same general rule practiced in the linear regression. Therefore, the procedure employed in the linear regression model will encourage our approach towards the logistic regression. But we cannot apply linear regression method directly, because in logistic regression

- The error terms do not follow normal distribution.
- The probability of success is limited to 0 to 1 interval.

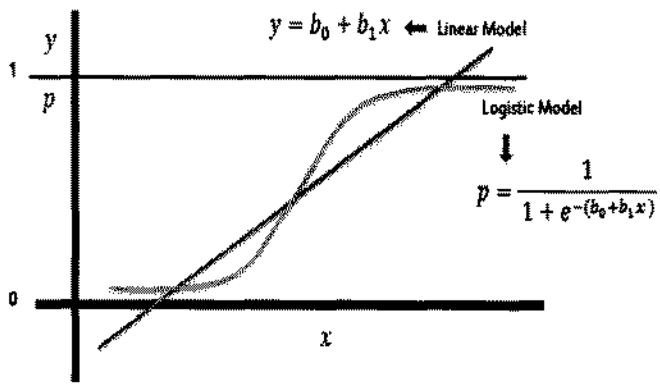


Figure 3.4: Linear and Logistic Models

3.7 Model Fitting

3.7.1 Estimation of Parameters

Consider a sample having k independent values of pair (x_j, y_j), where j varies from 1 to k, y_j represents the binary outcome variable and x_j denotes an independent variable. Furthermore, suppose that a binary variable is coded as '0' or '1', where '0' represents an absence and '1' indicates the presence of a characteristic. In order to assess the unknown values of the parameters i.e. β_0 & β_1, it requires to fit a model of logistic regression by using equation (3.5). In logistic regression, the least squares method is not appropriate for the estimation of the unknown parameters. The reason is that the error terms don't follow the normal distribution in logistic regression. Therefore, Maximum Likelihood (M.L.) estimation method is applied to find the unknown parameters which maximizes the probability of attaining an observed set of the data.

3.7.1.1 Maximum Likelihood Estimation

The 1st step in this method is to build a function, known as Likelihood function. As y_j is categorized as '0' or '1', then the statement for $\pi(x)$ in equation (3.5) offers the conditional probability P(Y=1 / x) and the quantity {1 - $\pi(x)$} offers the conditional probability P(Y=0 / x). Hence, for the pairs where y_j=1, $\pi(x_i)$ will be contributed in likelihood function and for the pairs where y_j=1, (1-$\pi(x_i)$) will be contributed in likelihood function. An easy method to express the likelihood function for pair (x_j, y_j) is from the expression

$$\{\pi(x_i)\}^{y_i}.\{1-\pi(x_i)\}^{1-y_i} \qquad \textbf{(3.14)}$$

Suppose, the observations are independent, then the likelihood function is achieved by the multiplication of terms used in expression (3.14) as follows:

$$l(\beta) = \prod_{i=1}^{n}\{\pi(x_i)\}^{y_i}.\{1-\pi(x_i)\}^{1-y_i} \qquad \textbf{(3.15)}$$

applying ln on both sides

$$\ln\{l(\beta)\} = \ln\left[\prod_{i=1}^{n}\{\pi(x_i)\}^{y_i}.\{1-\pi(x_i)\}^{1-y_i}\right]$$

$$\ln\{l(\beta)\} = \sum_{i=1}^{n}[y_i\,.\,ln\{\pi(x_i)\} + (1-y_i)\,.\,ln\{1-\pi(x_i)\}]$$

$$\ln\{l(\beta)\} = \sum_{i=1}^{n}\left[y_i\,.\,ln\left\{\frac{e^{(\beta_0+\beta_1.x)}}{(1+e^{(\beta_0+\beta_1.x)})}\right\} + (1-y_i)\,.\,ln\left\{1-\frac{e^{(\beta_0+\beta_1.x)}}{(1+e^{(\beta_0+\beta_1.x)})}\right\}\right]$$

$$\ln\{l(\beta)\} = \sum_{i=1}^{n}\left[y_i\,.\,ln\left\{\frac{e^{(\beta_0+\beta_1.x)}}{(1+e^{(\beta_0+\beta_1.x)})}\right\} + (1-y_i)\,.\,ln\left\{\frac{1}{(1+e^{(\beta_0+\beta_1.x)})}\right\}\right]$$

$$\ln\{l(\beta)\} = \sum_{i=1}^{n}\left[y_i\,.\,ln\left\{e^{(\beta_0+\beta_1.x)}\right\} - y_i.\ln\left\{1+e^{(\beta_0+\beta_1.x)}\right\} + (1-y_i)\,.\,ln\left\{(1+e^{(\beta_0+\beta_1.x)})^{-1}\right\}\right]$$

$$\ln\{l(\beta)\} = \sum_{i=1}^{n}\left[y_i\,.\,ln\left\{e^{(\beta_0+\beta_1.x)}\right\} - y_i.\ln\left\{1+e^{(\beta_0+\beta_1.x)}\right\} - (1-y_i)\,.\,\ln\left\{1+e^{(\beta_0+\beta_1.x)}\right\}\right]$$

$$\ln\{l(\beta)\} = \sum_{i=1}^{n}\Big[y_i.\;ln\left\{e^{(\beta_0+\beta_1.x)}\right\} - y_i.ln\left\{1+e^{(\beta_0+\beta_1.x)}\right\} - ln\left\{1+e^{(\beta_0+\beta_1.x)}\right\} +$$
$$y_i\,.\,ln\left\{1+e^{(\beta_0+\beta_1.x)}\right\}\Big]$$

$$\ln\{l(\beta)\} = \sum_{i=1}^{n}\left[y_i.\,(\beta_0+\beta_1.x) - ln\left\{1+e^{(\beta_0+\beta_1.x)}\right\}\right] \qquad \textbf{(3.16)}$$

Equation (3.16) is log likelihood function of logistic regression. The maximum likelihood estimators are obtained by taking the derivative of $\ln\{l(\beta)\}$ and equating it to zero, i.e. $\frac{\partial}{\partial\beta}\ln\{l(\beta)\} = 0$.

3.8 Evaluating the Significance of Coefficients

When the coefficients are estimated, the next procedure is to analyze the impact of variables of the fitted model. This commonly consists of formulation & the testing of hypothesis to conclude whether the regressors in the estimated model are significantly related to the dependent variable. Following tests are used to investigate the coefficients' significance.

3.8.1 Wald Test

Wald test statistic is used to assess the individual worth of parameters with the response variable. Let β symbolize a logit parameter. Suppose, H_0: $\beta=\beta_0$. For logistic regression, H_0: $\beta=0$ illustrates that the predictor is independent of the probability of success. The modest test statistic uses large sample normality of maximum likelihood estimator $\hat{\beta}$. Let *S.E.* indicate the standard error of an estimator $\hat{\beta}$. The test-statistic

$$Z = \frac{\hat{\beta} - \beta}{S.E.(\hat{\beta})} \qquad (3.17)$$

follows a normal distribution. Whereas, Z^2 follows a chi-squared distribution using $df=1$. Such form of test statistic is termed as a *Wald-statistic*. The chi-squared test evaluated by this statistic is known as *Wald test.*

3.8.2 Likelihood Ratio Test (The Deviance)

Although for large samples, Wald-test is better, but likelihood ratio test is also reliable. Let l_0 equals the maximum value of likelihood function under a null hypothesis & let l_1 equals the maximum value of likelihood function of saturated model i.e. unrestricted β. For example, for only parameter β, l_0 is a likelihood function estimated at β_0 & l_1 is the likelihood function estimated at the M.L. estimate $\hat{\beta}$.

The *likelihood-ratio* test statistic equals

$$D = -2\ ln\ (\ l_0\ /\ l_1\ ) \qquad \textbf{(3.18)}$$

In the equation (3.18), D denotes the deviance & ln is abbreviated as natural logarithm. The value of the test statistic $\{-2\ ln\ (\ l_0\ /\ l_1\ )\}$ will be always positive, and comparatively low values of $(\ l_0\ /\ l_1\ )$ will lead to high values of $\{-2\ ln\ (\ l_0\ /\ l_1\ )\}$ and provide a solid evidence against the Null Hypothesis H_0. The purpose of applying log transformation and doubling up is that it follows a chi-squared distribution.

3.9 Assessing the Goodness of Fit

When the procedure of model building is accomplished, the sequence of steps may well be taken to check the model fit. In linear regression, R^2 is usually used to measure the model fit. But in logistic regression, different techniques including the following test can be applied.

3.9.1 The Hosmer-Lemeshow Test

With the categorical predictors, the data set may have the type of *grouped* or *ungrouped data*. Ungrouped data set consists of raw 0 and 1 observations, whereas Grouped data are the aggregate of "Successes" and "Failures" at every arrangement of predictor values. The Hosmer-Lemeshow test is a measure of goodness of fit of a model that can be used in modelling ungrouped binary data. The Hosmer-Lemeshow test statistic $(\hat{C})$ uses the Pearson statistic to compare fitted and observed counts. A formula describing the computation of $(\hat{C})$ is as under:

$$\hat{C} = \sum_{i=1}^{g} \frac{(O_i - n_i'\bar{\pi}_i)^2}{n_i'\bar{\pi}_i(1-\bar{\pi}_i)} \qquad \textbf{(3.19)}$$

Where, n_i' = the total quantity of the subjects in i^{th} group.

O_i= number of the responses.

and $\bar{\pi}_i$ = the average predicted probability of success.

The distribution of Hosmer and Lemeshow statistic ($\hat{C}$) approaches Chi-square distribution having d.f.= (g-2). The benefit of using this test is that it makes available a simple interpretable value that will determine the model fit.

3.9.2 Receiver Operating Characteristic (R.O.C.) Curve

R.O.C. curve is used to estimate the fit of logit regression model which is based on specificity (True negative) and sensitivity (True positive) for all probable cutoff values. (1-specificity) is plotted on x-axis and sensitivity is plotted on y-axis. The curve obtained is named as R.O.C. curve. It helps to determine the accuracy of the diagnostic test. An area under curve (A.U.C.) signifies the accuracy of model. The area below the R.O.C. curve lies between 0.5 to 1.0, and greater values indicate the superior fit. The more the curve close to 1, the better the performance of R.O.C. curve. Diagonal line represents a test that has "0" Specificity and "0" Sensitivity. An A.U.C. = 1 signifies a perfect test, i.e. perfect Specificity & perfect Sensitivity. And an A.U.C. ≤ 0.5 represents a worthless test.

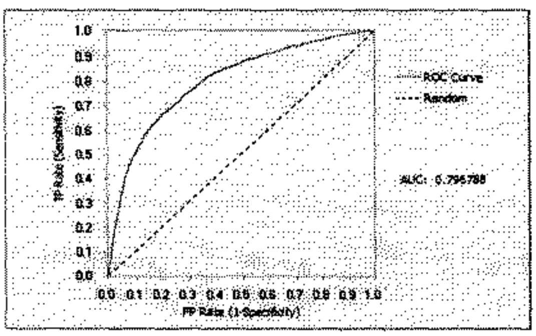


Figure 3.5: Receiver Operating Characteristic Curve

Chapter 4

Results & Discussions

4.1 Analysis of Hepatitis C Hazard Factors

The number of risk factors were divided into Demographic, socio economic and clinical (medical) factors. The data were analyzed according to Descriptive as well as Inferential Statistics. Overall possible hazard factors of Hepatitis C virus are tabulated below.

Table 4.1: Overall Possible Hazard Factors of Hepatitis C in Punjab

		Hepatitis C		Total
		Negative	Positive	
Age Group	Under 5	1954	74	**2028**
	5-19	9350	323	**9673**
	20-29	4471	269	**4740**
	30-39	2692	322	**3014**
	40-49	1952	333	**2285**
	50-59	1156	202	**1358**
	Above 60	1282	220	**1502**
Total		**22857**	1743	**24600**
Ear & Nose Piercing Females > 5 Years	No	2746	92	**2838**
	Yes	7118	658	**7776**
Total		**9864**	**750**	**10614**
Had Jaundice	No	22499	1659	**24158**
	Yes	358	84	**442**
Total		**22857**	**1743**	**24600**
History Of Hospitalization	No	21981	1610	**23591**
	Yes	876	133	**1009**
Total		**22857**	**1743**	**24600**
	No	20842	1649	**22491**

Tattooing/Acupuncture > 5 Years Age	Yes	61	20	**81**
Total		**20903**	**1669**	**22572**
Knowledge About Hepatitis	No	13481	966	**14447**
	Yes	9376	777	**10153**
Total		**22857**	**1743**	**24600**
History Of Haemodialysis	No	22827	1735	**24562**
	Yes	30	8	**38**
Total		**22857**	**1743**	**24600**
Share Smoking/Hokah Age > 19 Years	No	10909	1250	**12159**
	Yes	651	96	**747**
Total		**11560**	**1346**	**12906**
Share Toothbrush/Razor/Etc Age > 5 Years	No	20719	1644	**22363**
	Yes	184	25	**209**
Total		**20903**	**1669**	**22572**
Marital Status	Never Married	3447	100	**3547**
	Ever Married	9168	602	**9770**
Total		**12615**	**702**	**13317**
Family Suffering From Hepatitis	No	14379	991	**15370**
	Yes	508	109	**617**
	Don't Know	7970	643	**8613**
Total		**12037**	**1743**	**24600**
History Of Intramuscular Injections	No	6181	370	**6551**
	Yes	16432	1364	**17796**
	Dont Know	244	9	**253**
Total		**22857**	**1743**	**24600**
Use Of Intramuscular Injections	None	6425	379	**6804**
	Less Than 5	11369	801	**12170**
	5-10	3975	406	**4381**
	Greater Than 10	1088	157	**1245**
Total		**22857**	**1743**	**24600**
Type Of Syringe Used	Don't Know	8125	496	**8621**
	New/Disposable	7768	489	**8257**
	Re-Use	6964	758	**7722**
Total		**22857**	**1743**	**24600**
Level Of Education	Illiterate	9764	969	**10733**
	Below Matric	9985	577	**10562**
	Matric & Above	2661	171	**2832**
	Graduate & Above	447	26	**473**
Total		**22857**	**1743**	**24600**

Type Of House	Kacha	4391	327	**4718**
	Pacca	12919	937	**13856**
	Semi-Pacca/Kacha	4875	453	**5328**
	Well Furnished	672	26	**698**
Total		**22857**	**1743**	**24600**
Drinking Source	Others	638	81	**719**
	Piped In Dwelling	15705	1198	**16903**
	Public Tap	4246	298	**4544**
	Spring/Pound	151	22	**173**
	Tanker,Vendor	671	56	**727**
	Well	1446	88	**1534**
Total		**22857**	**1743**	**1743**
Source Of Shave Males Age > 19 Years	None	2012	130	**2142**
	Home	982	100	**1082**
	Barber	2110	350	**2460**
	Both	1001	245	**1246**
Total		**6105**	**875**	**6930**

4.2 Descriptive Section (Hepatitis C)

In Descriptive section, percentages and counts were evaluated for different risk factors.

4.2.1 Rate of Pervasiveness of Hepatitis C in Punjab according to Demographic Factors

The pervasiveness percentage of Hepatitis C in Punjab according to age, gender and marital status is tabulated below,

Table 4.2: Pervasiveness of Hepatitis C in Punjab

PROVINCE	NUMBER OF SAMPLES	Pervasiveness of Hepatitis C		
		Positive	%	95% Confidence Interval
PUNJAB	24600	1743	7.1	6.8 - 7.4

Table (4.2) showed that the numbers of positive cases screened were 1743 out of 24600 samples. The Pervasiveness rate of Hepatitis C in Punjab was 7.1% with the confidence interval (6.8 – 7.4).

Table 4.3: Pervasiveness of Hepatitis C according to Gender

GENDER	TOTAL	Pervasiveness of Hepatitis C		
		Positive	%	95% Confidence Interval
MALE	12964	927	7.15	7.0 – 7.3
FEMALE	11636	816	7.01	6.7 – 7.2
TOTAL	24600	1743	7.1	6.8 - 7.4

Table (4.3) illustrated the Pervasiveness of Hepatitis C according to gender. The Pervasiveness rate of Hepatitis C in Punjab for male was 7.15% with a confidence interval (7.0 – 7.3), whereas the Pervasiveness rate of Hepatitis C in Punjab for female was 7.01% with a confidence interval (6.7 – 7.2), which indicated that the pervasiveness of Hepatitis C in Punjab was relatively higher in males as compared to females.

Table 4.4: Pervasiveness of Hepatitis C according to Marital Status

MARITAL STATUS Age > 19 Years	TOTAL	Pervasiveness of Hepatitis C		
		Positive	%	95% Confidence Interval
NEVER MARRIED	3547	100	2.93	2.75 – 3.14
EVER MARRIED	9770	602	6.16	6.0 – 6.33
TOTAL	13317	702	5.27	5.01 – 5.52

Table (4.4) described the Pervasiveness of Hepatitis C according to Marital Status. The Pervasiveness rate of Hepatitis C in never-married group was 2.93%, whereas the Pervasiveness rate of Hepatitis C in ever-married group is 6.16%. Ever married persons were more likely to have Hepatitis C than Never married persons.

Table 4.5: Pervasiveness of Hepatitis C according to Age

AGE	TOTAL	Pervasiveness of Hepatitis C		
		Positive	%	95% Confidence Interval
Under 5	2028	74	3.65	3.51 – 3.83
5-19	9673	323	3.34	3.21 – 3.46
20-29	4740	269	5.68	5.45 – 5.81
30-39	3014	322	10.68	10.38 – 10.85
40-49	2285	333	14.57	14.42 – 14.78
50-59	1358	202	14.87	14.69 – 15.02
Above 60	1502	220	14.65	14.49 – 14.81
TOTAL	24600	1743	7.1	6.8 – 7.4

Table (4.5) indicated the Pervasiveness of Hepatitis C according to age. Hepatitis B was less than 4% in age groups under 5 and 5 – 19, but that rate tend to increase from 30 years of age at greater pace and achieved its maximum value at the age group of 50-59. It showed that the pervasiveness rate of Hepatitis C increased with the increase in age.

4.2.2 Pervasiveness Rate of Hepatitis C in Punjab according to Socio-economic Factors

The pervasiveness percentage of Hepatitis C in Punjab according to level of education, house material, drinking source, knowledge about hepatitis and divisions of Punjab are tabulated below.

Table 4.6: Pervasiveness of Hepatitis C according to Level of Education

Level Of Education	TOTAL	Pervasiveness Of Hepatitis C		
		Positive	%	95% Confidence Interval
Illiterate	10733	969	9.03	8.85 – 10.17
Below Matric	10562	577	5.46	5.29 – 5.63
Matric & Above	2832	171	6.04	5.90 – 6.20
Graduate & Above	473	26	5.50	5.37 – 5.63
TOTAL	24600	**1743**	7.1	6.8 – 7.4

Table (4.6) demonstrated the Pervasiveness rate of Hepatitis C according to the level of Education. The Pervasiveness of Hepatitis C was maximum (i.e. 9.03%) in the group of illiterate people whereas it was minimum (i.e. 5.46%) in the group of matric and above. The Pervasiveness rate of Hepatitis C was also less in the higher level of Education (i.e. Graduate and above).

Table 4.7: Pervasiveness of Hepatitis C according to House Material

House Material	TOTAL	Pervasiveness Of Hepatitis C		
		Positive	%	95% Confidence Interval
Kacha	4718	327	6.93	6.75 – 7.09
Pacca	13856	937	6.76	6.58 – 6.91
Semi-pacca/kacha	5328	453	8.50	8.30 – 8.70
Well furnished	698	26	3.72	3.59 – 3.80
TOTAL	24600	**1743**	7.1	6.8 – 7.4

Table (4.7) revealed the Pervasiveness rate of Hepatitis C according to the House Material. The Pervasiveness of Hepatitis C was maximum (i.e. 8.50%) in the group of people living in Semi-pacca/kacha house, whereas it was minimum (i.e. 3.72%) in the group of people living in Well-furnished house.

Table 4.8: Pervasiveness of Hepatitis C according to Drinking Source

Drinking Source	TOTAL	Pervasiveness Of Hepatitis C		
		Positive	%	95% Confidence
Others	730	81	11.09	10.92 – 11.23
Piped in Dwelling	16903	1198	7.09	6.95 – 7.22
Public Tap	4544	298	6.56	6.42 – 6.71
Spring/Pound	162	22	13.58	13.45 – 13.72
Tanker/Vendor	727	56	7.70	7.55 – 7.84
Well	1534	88	5.74	5.60 – 5.89
Total	24600	**1743**	7.1	6.8 – 7.4

Table (4.8) illuminated the Pervasiveness of Hepatitis C according to Drinking source. Hepatitis C was maximum (i.e. 13.58%) among the group of people who used water from spring/pound but that rate tend to decrease among the group of people who used water from well.

Table 4.9: Pervasiveness of Hepatitis C according to Knowledge about Hepatitis

Knowledge about Hepatitis	TOTAL	Pervasiveness Of Hepatitis C		
		Positive	%	95% Confidence Interval
No	14447	966	6.70	6.58 – 6.83
Yes	10153	777	7.65	7.55 – 7.76
Total	24600	**1743**	7.1	6.8 – 7.4

Table (4.9) illustrated the Pervasiveness rate of Hepatitis C according to the Knowledge about Hepatitis. The Pervasiveness of Hepatitis C was higher (i.e. 7.65%) in the group of people who had knowledge about Hepatitis. Therefore, it was found that the Pervasiveness rate of Hepatitis C in Punjab decreased with the increase in the awareness of Hepatitis.

Table 4.10: Pervasiveness of Hepatitis C according to Divisions of Punjab

AGE	TOTAL	Pervasiveness of Hepatitis C		
		Positive	%	95% Confidence Interval
Rawalpindi	3153	159	5.04	4.92 – 5.20
Gujranwala	5104	402	7.88	7.75 – 8.02
Sarghoda	2368	113	4.77	4.66 – 4.89
Faisalabad	1695	147	8.67	8.55 – 8.80
Lahore	3745	268	7.16	7.01 – 7.31
Multan	4495	293	6.52	6.38 – 6.68
Sahiwal	1501	148	9.86	9.75 – 10.00
Bahawalpur	2343	207	8.83	8.70 – 9.00
D.G. Khan	196	6	3.06	2.89 – 3.30
TOTAL	24600	**1743**	7.1	6.8 – 7.4

Table (4.10) showed the rate of Pervasiveness of Hepatitis C according to the Divisions of Punjab. The Pervasiveness of Hepatitis C was maximum (i.e. 9.86%) in Sahiwal Division, whereas it was minimum in Dera Ghazi Khan Division (i.e. 3.06%). The rate of Pervasiveness of Hepatitis C in Rawalpindi, Gujranwala, Sarghoda, Faisalabad, Lahore, Multan and Bahawalpur were 5.04%, 7.88%, 4.77 %, 8.67%, 7.16%, 6.52% and 8.83% respectively.

4.3 Inferential Section (Hepatitis C)

In Inferential section, bivariate analysis and multivariate analysis were applied independently to determine the significant hazard factors for both males and females. Inferential section is a major section of the research study which may be distributed as

- Bivariate Analysis
- Multivariate Analysis

4.3.1 Bivariate Analysis (for Female)

In this section, the association between each hazard factor using odds ratio was tested separately with the Hepatitis C (for females). Wald test is a statistical procedure used in bivariate analysis to test the association between Hepatitis and a Risk Factor.

4.3.1.1 Association between Hepatitis C & Hazard Factors

Table 4.11: Hepatitis C vs. Age Group

Hepatitis C (Female)		B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
							Lower	Upper
AGE GROUP	5-19	-0.08	0.168	0.25	0.621	0.92	0.64	1.31
	20-29	0.54	0.320	8.35	0.004	1.72*	1.19	2.47
	30-39	1.19	0.607	41.86	0.000	3.3*	2.29	4.73
	40-49	1.28	0.682	45.43	0.000	3.6*	2.48	5.21
	50-59	1.37	0.789	46.65	0.000	3.94*	2.65	5.82
	Above 60	1.33	0.772	42.64	0.000	3.78*	2.54	5.65

Table 4.11 illustrated the independent association between Hepatitis C (Females) and Age Group. The results showed that the pervasiveness of Hepatitis C

in females was significantly higher at age 30 or above. The binary logistic regression coefficients (Odds Ratios) of age groups 30-39, 40-49, 50-59 and above 60 were 1.19 (3.3), 1.28 (3.6), 1.37 (3.94) & 1.33 (3.78) respectively. Therefore, we reject the null hypotheses and concludes that Hepatitis C (females) and the Age Groups 30-39, 40-49, 50-59 and above 60 have significant associations at 1% level of significance.

Table 4.12: Hepatitis C vs. Ear/Nose Piercing

Hepatitis C (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Ear/Nose Piercing	1.02	0.155	40.23	.000	2.76*	2.29	3.10

Table 4.12 showed the autonomous relationship between Hepatitis C and Ear/Nose Piercing. The logistic regression coefficient for Ear/Nose Piercing was 1.02 and Odds Ratio (O.R.) = 2.76 with a confidence interval (2.29 – 3.10). The result suggested that there was 2.76 times higher risk of getting Hepatitis C in those females who had got their Ear/ Nose Pierced. The value of Wald statistic was 40.23 which was significant. Therefore, we reject the null hypotheses and determines that Hepatitis C and Ear/Nose Piercing have significant association at 1% level of significance.

Table 4.13: Hepatitis C vs. History of Jaundice

Hepatitis C (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
History of Jaundice	1.45	0.81	59.29	.000	4.28*	2.96	6.21

Table 4.13 revealed the independent association of Hepatitis C with those females which ever had jaundice in their past. The binary logistic regression coefficient for Jaundice was 1.45 and Odds Ratio (O.R.) = 4.28 with a confidence interval (2.96 – 6.21). It is concluded that the female with a history of jaundice had 4.28 times more risk of obtaining Hepatitis C as compared to those who had no jaundice ever. The value of Wald statistic was 59.29 which was statistically significant. Therefore, we reject the null hypotheses and analyzes that Hepatitis C and History of Jaundice have significant association at 1% level of significance.

Table 4.14: Hepatitis C vs. Family History of Hepatitis

Hepatitis C (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Family history of Hepatitis	1.29	0.577	51.26	.000	3.64*	2.67	4.97

Table 4.14 determined the relationship between Hepatitis C and the family history of Hepatitis independently. The binary logistic regression coefficient was 1.29 and Odds Ratio (O.R.) = 3.64 with a confidence interval (2.67 – 4.97), which expressed that the female with a family history of Hepatitis had 3.64 times greater chances of getting Hepatitis C. We assume that the base category was no family history of Hepatitis. The value of Wald statistic was 51.26 which was statistically statistically significant. Therefore, we reject the null hypothesis and states that Hepatitis C and family history of Hepatitis have significant association at 1% level of significance.

Table 4.15: Hepatitis C vs. Marital Status

Hepatitis C (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Ever Married	1.19	0.26	43.65	.000	3.27*	2.79	3.81

Table 4.15 described the independent association of Hepatitis C and the Marital Status. The binary logistic regression coefficient of females which were ever married was 1.19 and Odds Ratio (O.R.) = 3.27 with a confidence interval (2.79 – 3.81). The result suggested that there was 3.27 times greater risk of having Hepatitis C in those females who were ever married. We assume that the base category was no history of marriage. The value of Wald statistic is 43.65 which is statistically significant. Therefore, we reject the null hypothesis and interprets that Hepatitis C and Ever Married individuals have significant association at 1% level of significance.

Table 4.16: Hepatitis C vs. Syringe Type

Hepatitis C (Female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
New/Disposable	-0.04	0.093	0.144	0.703	0.96	0.80	1.17
Re-Use	0.57	0.154	41.22	0.000	1.76*	1.48	2.09

Table 4.16 investigated the relationship between Hepatitis C (Females) and Type of Syringe. The binary logistic regression coefficient of females who used New/Disposable syringe was -0.04 and Odds Ratio (O.R.) = 0.96 with a confidence interval (0.80 – 1.17). The value of Wald statistic is 0.144 which is insignificant.

Therefore, we accept the null hypothesis that Hepatitis C and the females who use New/Disposable syringe have no significant association at 1% level of significance.

The binary logistic regression coefficient of females who use Re-use syringes was 0.57 and Odds Ratio (O.R.) = 1.76 with a confidence interval (1.48 – 2.09), which concluded that the female with a history of re-use syringe had 1.76 times more risk of having Hepatitis C. We assume that the base category was no history of Injections. The value of Wald statistic was 41.22 which was statistically significant. Therefore, we reject the null hypotheses and concluded that Hepatitis C and females who used Re-use syringes have significant association at 1% level of significance.

Table 4.17: Hepatitis C vs. History of Hospitalization

Hepatitis C (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
History of Hospitalization	0.75	0.278	31.92	.000	2.11*	1.63	2.73

Table 4.17 illustrated the autonomous relation of Hepatitis C and the history of Hospitalization. The binary logistic regression coefficient is 0.75 and Odds Ratio (O.R.) = 2.11 with a confidence interval (1.63 – 2.73), which revealed that the female with the history of Hospitalization had 2.11 times higher chances of getting Hepatitis C. We assume that the base category was no history of Hospitalization. The value of Wald statistic is 31.92 which was statistically significant. So, we reject the null hypothesis and concluded that Hepatitis C and history of Hospitalization have significant association at 1% level of significance.

Table 4.18: Hazard Factors of Hepatitis C Transmission in Female

Risk Factors		Hepatitis C Negative	Hepatitis C Positive	Total	O.R.	95% C.I.
Age Group	Under 5	958	38	996	Reference	
	5-19	4409	160	4569	0.92	0.64 - 1.31
	20-29	2096	143	2239	1.72	1.19 - 2.47
	30-39	1315	172	1487	3.3*	2.29 - 4.73
	40-49	926	132	1058	3.6*	2.48 - 5.21
	50-59	576	90	666	3.94*	2.65 - 5.82
	Above 60	540	81	621	3.78*	2.54 - 5.65
				11636		
Ear & Nose Piercing (females age above 5 years)	No	2746	92	2838	Reference	
	Yes	7118	658	7776	2.76*	2.29 – 3.10
				10614		
Had Jaundice	No	9220	629	9849	Reference	
	Yes	130	38	168	4.28*	2.96 – 6.21
	Don't Know	1470	149	1619	1.48	1.23 – 1.79
				11636		
Family Suffering From Hepatitis	No	6748	407	7155	Reference	
	Yes	246	54	300	3.64*	2.67 – 4.97
	Don't Know	3826	355	4181	1.54	1.33 – 1.78
				11636		
Marital Status (age above 19 years)	Unmarried	1573	23	1596	Reference	
	Married	3882	158	4040	2.80*	1.56 - 2.53
				5636		
Type Of Syringe Used	Don't Know	3670	226	3896	Reference	
	New/ Disposable	3758	223	3981	0.96	0.80- 1.17
	Re-Use	3392	367	3759	1.76*	1.48- 2.09
				11636		
History of Hospitalization	No	10345	744	11089	Reference	
	Yes	475	72	547	2.11*	1.63 – 2.73
				11636		

O.R. = Odds Ratio, C.I. = Confidence Interval, * shows significant at 1% level

4.3.2 Multivariate Analysis of Hepatitis C (for Female)

Since the purpose of this research is to determine the significant hazard factors which are related to the pervasiveness of Hepatitis B & C for males and females independently. The question arises that how should we choose the best logistic regression model in case of a dichotomous (binary) response variable. The assortment procedure would be more difficult as the number of predictors increases. Forward selection criteria is adapted to select the model. In forward selection method, the terms are added one after the other until more addition of terms do not increase the model fit. In this section, Hepatitis C (Females) is taken as a binary response variable, whereas all other factors are taken as independent variables. Table 4.19 showed the multivariate model for Hepatitis C (for Females).

Table 4.19: Multivariate Model for Hepatitis C (for females)

VARIABLES	B	S.E.	Wald Statistic	p-value
Age	.012	.0026	22.140*	0.007
New/Disposable Syringe	-.029	.02	2.10	0.573
Re-Use Syringe	.350	.075	21.561*	0.002
Ear & Nose Piercing	.374	.103	13.309*	0.004
Jaundice's History	1.079	.200	29.233*	0.000
Ever Married	.763	.104	54.306*	0.0005
Family History Of Hepatitis	.975	.168	33.780*	0.0001
Constant	-1.913	.35	29.87*	0.000

Note: * shows significant at 0.01 level.

The equation for Multivariate Logit Model for Hepatitis C (for females) is

$$\ln\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = -1.913 + 0.012\,(Age) + 0.35\,(Reuse\ Syringe) + 0.374\,(Piercing)$$
$$+ 1.079(Jaundice) + 0.763\,(Ever\ Married)$$
$$+ 0.975\,(Family\ History\ of\ Hepatitis)$$

(4.1)

4.3.2.1 Estimated Probabilities

The equation (4.1) can be used to estimate the different probabilities of Hepatitis C in females in the absence or presence of significant hazard factors. Few examples are stated below,

- If Age=60, Reuse Syringe=1, Piercing=1, Jaundice=1, Ever Married=1 and Family history of Hepatitis=1, then the estimated probability is

$$\hat{\pi} = \frac{1}{1 + e^{(-2.348)}} = 0.9127$$

It means that the probability of getting Hepatitis C is 0.9127 in the existence of above declared hazard factors.

- If Age=60, Reuse Syringe=0, Piercing=1, Jaundice=0, Ever Married=1 and Family history of Hepatitis=0, then the estimated probability is

$$\hat{\pi} = \frac{1}{1 + e^{(0.056)}} = 0.4860$$

It means that the probability of getting Hepatitis C is 0.4860. It reveals that the less exposure to hazard factors lead to lower probability of getting Hepatitis C in females.

Table 4.20: Hosmer - Lemeshow Statistic

$\hat{C}$	P-value
7.66	.467

The H.L. (Hosmer Lemeshow) statistic has a value i.e. $\hat{C}$ = 7.66 & P-value evaluated from chi square distribution having d.f.= 8 is 0.467, that leads to accept H_0 and specifies that the model has been fitted well.

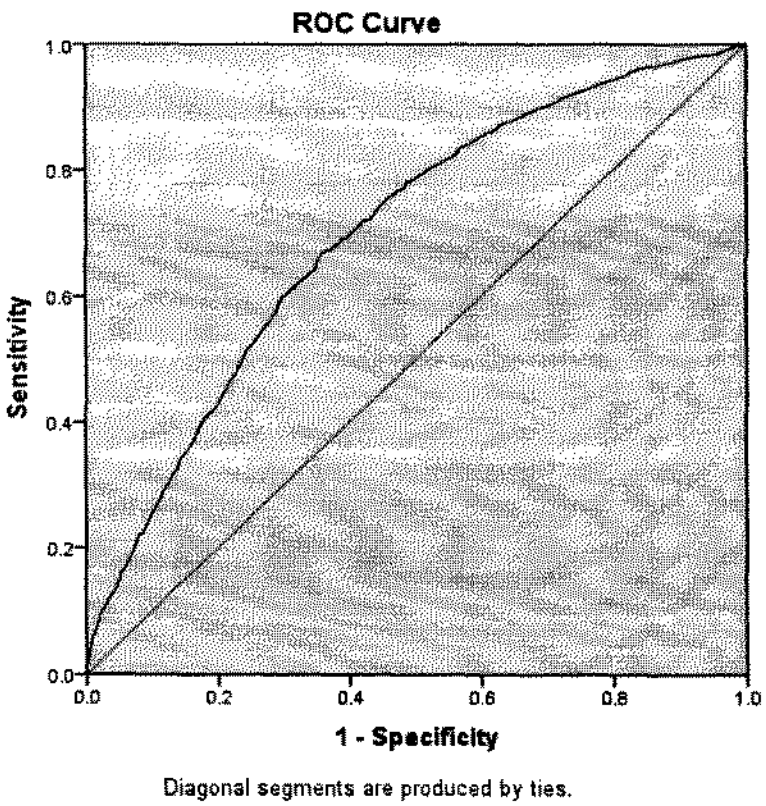


Figure 4.1: R.O.C. Curve of Hepatitis C (for females)

Table 4.21: Area under the Curve (A.U.C.)

Test Result Variable(s): Predicted probability				
A.U.C.	S.E.	Asymptotic Sig.	Asymptotic 95% C.I.	
			Lower Bound	Upper Bound
0.71	0.009	.000	0.679	0.721

Table 4.21 indicates the area under the R.O.C. curve. Since A.U.C. is 0.71, which is significantly different from 0.5, depicted an acceptable discrimination. The significance of A.U.C. (p-value=0.000) provides a strong evidence of the accuracy of fitted model. It concludes that the data follows logistic regression distribution.

4.3.3 Bivariate Analysis of Hepatitis C (for Males)

In this section, the association between each hazard factor using odds ratio was tested separately with the Hepatitis C (for Males). Wald test is the statistical procedure used in bivariate analysis to test the association between Hepatitis and a Risk Factor.

4.3.3.1 Association between Hepatitis C and Risk Factors

Table 4.22: Hepatitis C vs. Age Group

Hepatitis C (male)		B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
							Lower	Upper
AGE GROUP	5-19	-0.08	0.172	0.221	0.639	0.92	0.63	1.32
	20-29	0.385	0.285	4.08	0.043	1.47	1.01	2.15
	30-39	1.112	0.578	34.22	0.000	3.04*	2.09	4.41
	40-49	1.686	1.01	82.26	0.000	5.4*	3.76	7.81
	50-59	1.673	1.06	71.40	0.000	5.33*	3.62	7.90
	Above 60	1.637	1.001	72.42	0.000	5.14*	3.55	7.56

Table 4.22 illustrated the independent association between Hepatitis C (males) and Age Group. The results showed that the pervasiveness of Hepatitis C in males was significantly higher at age 30 or above. The binary logistic regression coefficients (Odds Ratios) of age groups 30-39, 40-49, 50-59 and above 60 are 1.112 (3.04), 1.686 (5.4), 1.673 (5.33) & 1.637 (5.14) respectively. As a result, we reject the null hypotheses and concluded that Hepatitis C (males) and the Age Groups 30-39, 40-49, 50-59 and above 60 have significant association at 1% level of significance.

Table 4.23: Hepatitis C vs. Shaving

Hepatitis C (Males)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Home	0.432	0.534	3.75	0.06	1.54	1.25	1.89
Barber	0.944	0.161	55.61	0.000	2.57*	2.28	3.04
Both	1.332	0.092	70.24	0.000	3.79*	3.36	4.18

Table 4.23 described the relationship of Hepatitis C (Males) and Shaving independently. The binary logistic regression coefficient of males who shave at Home was 0.432 and Odds Ratio (O.R.) = 1.54 with the confidence interval (1.25 – 1.89). The value of Wald statistic was 3.75 which was statistically insignificant. Therefore, we accept the null hypotheses that Hepatitis C and the males shave at Home have no significant association at 1% level of significance.

The binary logistic regression coefficient of males who got their shave from Barber was 0.944 and Odds Ratio (O.R.) = 2.57 with a confidence interval (2.28 – 3.04), which concluded that the Male with a history of Barber Shave had 2.57 times more risk of getting Hepatitis C. We assume that the base category was no History of Shaving. The value of Wald statistic was 55.61 which was statistically significant. So, we reject the null hypotheses and found that Hepatitis C and males who get their shave from Barber has significant effect at 1% level of significance.

The binary logistic regression coefficient of males who got their shaves both from Barber and at Home was 1.332 and Odds Ratio (O.R.) = 3.79 with a confidence interval (3.36 – 4.18), which concluded that the Male with a history of both Barber and Home shaves had approximately double risk of getting Hepatitis C. We assume that the base category was no History of Shaving. The value of Wald statistic was

70.24 which was statistically significant. Consequently, we reject the null hypotheses and interprets that Hepatitis C and males who get their shaves from Barber and at home have significant association at 1% level of significance.

Table 4.24: Hepatitis C vs. History of Jaundice

Hepatitis C (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
History of Jaundice	1.001	0.452	36.35	.000	2.72*	1.96	3.77

Table 4.24 indicated the autonomous association between Hepatitis C and those Males which ever had jaundice in their past. The binary logistic regression coefficient for Jaundice was 1.001 and Odds Ratio (O.R.) = 2.72 with a confidence interval (1.96 – 3.77), which concluded that the Male with a history of jaundice had 2.72 times greater risk of attaining Hepatitis C as compared to those who had no jaundice ever. The value of Wald statistic was 36.35 which is significant. For that reason, we reject the null hypothesis and analyzes that Hepatitis C and History of Jaundice has significant association at 1% level of significance.

Table 4.25: Hepatitis C vs. Marital Status

Hepatitis C (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Ever Married	1.11	0.018	61.34	.000	3.03*	2.15	4.12

Table 4.25 described the independent association between Hepatitis C and the Marital Status. The binary logistic regression coefficient of Males which were ever married was 1.11 and Odds Ratio (O.R.) = 3.03 with a confidence interval (2.15

– 4.12), which showed that the married male had more than three times risk of getting Hepatitis C as compared to those who have no history of Marriage. The value of Wald statistic was 61.34 which was highly significant. Therefore, we reject the null hypotheses and finds that Hepatitis C and Ever Married males have significant association at 1% level of significance.

Table 4.26: Hepatitis C vs. Share Cigarettes/ Hookah

Hepatitis C (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Share cigarettes/hookah	0.59	0.029	20.31	0.007	1.8*	1.25	2.62

Table 4.26 illuminated the independent relationship of Hepatitis C and the Males who Share cigarettes/ hookah. The binary logistic regression coefficient was 0.59 and Odds Ratio (O.R.) = 1.8 with a confidence interval (1.25 – 2.62). The result suggested that the risk of Hepatitis C becomes double approximately in those males who Share cigarettes/ hookah. We assume that the base category was no sharing. The value of Wald statistic was 20.31 which was statistically significant. Consequently, we reject the null hypotheses and interprets that Hepatitis C and sharing cigarettes/hookah have significant association at 1% level of significance.

Table 4.27: Hepatitis C vs. Family Suffering From Hepatitis

Hepatitis C (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Family Suffering from Hepatitis	1.01	0.154	42.77	0.000	2.74*	2.03	3.71

Table 4.27 explained the independent association between Hepatitis C and the family suffering from Hepatitis. The binary logistic regression coefficient was 1.01 and Odds Ratio (O.R.) = 2.74 with a confidence interval (2.03 – 3.71), which concluded that the Male with a family history of Hepatitis had 2.74 times higher chances of getting Hepatitis C. We assume that the base category was no family history of Hepatitis. The value of Wald statistic was 42.77 which was highly significant. Thus, we reject the null hypothesis and concludes that Hepatitis C and family history of Hepatitis has significant association at 1% level of significance.

Table 4.28: Hepatitis C vs. Use of Intramuscular Injection

Hepatitis C (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Use of Intramuscular Injections	0.31	0.081	14.44	0.000	1.36*	1.16	1.59

Table 4.28 showed the autonomous relationship of Hepatitis C and those Males which ever used Intramuscular Injections in the past. The binary logistic regression coefficient was 0.31 and Odds Ratio (O.R.) = 1.36 with a confidence interval (1.16 – 1.59), which concluded that the Male with a history of Intramuscular Injections had 1.36 times more risk of obtaining Hepatitis C as compared to those who had no history of Intramuscular Injections. The value of Wald statistic was 14.44 which was statistically significant. Therefore, we reject the null hypothesis and reveals that Hepatitis C and use of Intramuscular Injections have significant association at 1% level of significance.

Table 4.29: Hepatitis C vs. Syringe Type

Hepatitis C (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
New/Disposable	0.09	0.089	1.023	0.311	1.09	0.92	1.30
Re-Use	0.59	0.082	51.69	0.000	1.81*	1.54	2.12

Table 4.29 illustrated the independent association between Hepatitis C (Males) and Type of Syringe. The binary logistic regression coefficient of males who use New/Disposable syringe was 0.09 and Odds Ratio (O.R.) = 1.09 with a confidence interval (0.92 – 1.30). The value of Wald statistic was 1.023 which was insignificant. Therefore, we accept the null hypotheses that Hepatitis C and the males who use New/Disposable syringe have no significant effect at 1% level of significance.

The binary logistic regression coefficient of males who use Re-Use syringes was 0.59 and Odds Ratio (O.R.) = 1.81 with a confidence interval (1.54 – 2.12), which concluded that the Male with a history of re-use syringe had 1.81 times greater risk of having Hepatitis C. We assume that the base category was no history of Injections. The value of Wald statistic was 51.69 which was highly significant. Therefore, we reject the null hypothesis and finds that Hepatitis C and males who use Re-Use syringes have significant association at 1% level of significance.

Table 4.30: Hepatitis C vs. History of Hospitalization

Hepatitis C (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for OR	
						Lower	Upper
History of Hospitalization	0.715	0.142	25.40	0.000	2.04*	1.55	2.70

Table 4.30 demonstrated the independent relationship of Hepatitis C and the history of hospitalization. The binary logistic regression coefficient was 0.715 and Odds Ratio (O.R.) = 2.04 with a confidence interval (1.55 – 2.70), which concluded that the Male with a history of hospitalization had 2.04 times more chances of getting Hepatitis C. We assume that the base category was no history of hospitalization. The value of Wald statistic was 25.40 which is statistically significant. As a result, we reject the null hypothesis and states that Hepatitis C and history of hospitalization has significant association at 1% level of significance.

Table 4.31: Hepatitis C vs. Tattooing/Acupuncture

Hepatitis C (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Tattooing/ Acupuncture	0.626	0.129	23.54	0.003	1.87*	1.23	2.51

Table 4.31 investigated the association between Hepatitis C and Tattooing/ Acupuncture independently. The binary logistic regression coefficient was 0.626 and Odds Ratio (O.R.) = 1.87 with a confidence interval (1.23 – 2.51), which explained that the Male with a history of Tattooing/Acupunture had 1.87 times higher risk of obtaining Hepatitis C as compared to those who had no history of Tattooing/Acupunture. The value of Wald statistic was 23.54 which was statistically significant. Consequently, we reject the null hypothesis and explains that Hepatitis C and Tattooing/Acupunture has significant association at 1% level of significance.

Table 4.32: Hazard Factors of Hepatitis C (for Males)

Risk Factors		Hepatitis C Negative	Hepatitis C Positive	Total	O.R.	95% C.I.
Age Group	Under 5	998	36	1034	Reference	--------
	05-19	4934	163	5097	0.92	0.63–1.32
	20-29	2370	126	2496	1.47	1.01-2.15
	30-39	1370	150	1520	3.04*	2.09-4.41
	40-49	1032	201	1233	5.4*	3.76-7.81
	50-59	583	112	695	5.33*	3.62-7.90
	Above 60	750	139	889	5.14*	3.55-7.56
				12964		
Shaving (age above 19 years)	No	2012	130	2142	Reference	
	Home	982	98	1080	1.54	1.01 - 2.01
	Barber	2110	350	2460	2.57*	1.47 – 3.62
	Both	1001	245	1246	3.79*	2.35 – 4.84
				6928		
Ever had Jaundice	No	10236	759	10995	Reference	
	Yes	228	46	274	2.72*	1.96 – 3.77
	Don't Know	1573	122	1695	1.04	0.86 - 1.28
				12964		
Marital Status (age above 19 years)	Never Married	2080	106	2186	Reference	
	Ever Married	4025	622	4647	3.03*	2.15 – 4.12
				6833		
Share Cigarettes/ Hookah (age above 19 years)	No	5501	609	6110	Reference	
	Yes	604	119	723	1.8*	1.25 - 2.62
				6833		
Family Suffering From Hepatitis	No	7631	584	8215	Reference	
	Yes	262	55	317	2.74*	2.03 – 3.71
	Don't Know	4144	288	4432	0.91	0.78 – 1.05
				12964		
Use Of Injections	No	3424	211	3635	Reference	
	Yes	8467	710	9177	1.36*	1.06 – 1.71
	Don't Know	146	6	152	0.67	0.29 – 1.53
				12964		

Type Of Syringe Used	Don't Know	4455	270	4725	Reference	
	New/Disposable	4010	266	4276	1.09	0.92 – 1.31
	Re-Use	3572	391	3963	1.81*	1.54 – 2.13
				12964		
History of Hospitalization	No	11636	866	12502	Reference	
	Yes	401	61	462	2.04*	1.55 – 2.70
				12964		
Tattooing/ Acupuncture (Age ≥ 5 Years)	No	11039	891	11930	Reference	
	Yes	73	11	84	1.87*	1.23 – 2.51
				12014		

4.3.4 Multivariate Analysis of Hepatitis C (for Males)

As the perseverance of this research is to reveal the significant hazard factors related to the pervasiveness of Hepatitis B & C. Forward selection criteria is adapted to select the model. Hepatitis C is taken as a binary response variable, whereas all other factors are taken as independent variables. Table 4.33 indicates the multivariate model for Hepatitis C (for Males).

Table 4.33: Multivariate Model for Hepatitis C (for Males)

	B	S.E.	Wald	p-value
Age	0.023	.0041	30.063*	0.003
Home Shaving	0.114	.126	.825	0.364
Barber Shaving	0.280	.089	10.01*	0.002
Jaundice's History	0.635	.177	12.906*	0.0006
Ever Married	0.714	.112	40.718*	0.0001
Share Cigarettes/ Hookah	0.229	.071	11.153*	0.005
Family History of Hepatitis	0.755	.164	21.095*	0.0002
New/Disposable Syringe	0.03	.021	1.98	0.19
Reuse Syringe	0.425	.062	25.248*	0.0001
Tattooing/ Acupuncture	0.461	0.112	16.942*	0.004
Constant	-2.309	.32	45.888	.000

Note: * shows significant at 0.01 level.

The equation for overall fitted Logit Model for Hepatitis C (for Males) is

$$\ln\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = -2.309 + 0.023\,(Age) + 0.280\,(Barber\ Shaving) + 0.425\,(Reuse\ Syringe)$$
$$+ 0.635\,(Jaundice) + 0.714\,(Ever\ Married) + 0.461\,(Tattooing)$$
$$+ 0.755\,(Family\ History\ of\ Hepatitis) + 0.229\,(Share\ Cigarette/Hookah) \tag{4.2}$$

4.3.4.1 Estimated Probabilities

The equation (4.2) can be used to estimate the different probabilities of Hepatitis C in females in the absence or presence of significant hazard factors. Few examples are stated below,

- If Age=50, Reuse Syringe=1, Barber Shaving =1, Jaundice=1, EverMarried=1, Share Cigarettes/ Hookah=0, Tattooing/Acupuncture=1 and Family history of Hepatitis=0, then the estimated probability is

$$\hat{\pi} = \frac{1}{1+\ e^{(-1.356)}} = 0.7951$$

It means that the probability of getting Hepatitis C is 0.7951 in the existence of above declared hazard factors.

- If Age=50, Reuse Syringe=0, Barber Shaving=1, Jaundice=0, EverMarried=1, Share Cigarettes/ Hookah=0, Tattooing/Acupuncture=0 and Family history of Hepatitis=1, then the estimated probability is

$$\hat{\pi} = \frac{1}{1+\ e^{(-0.59)}} = 0.6433$$

It means that the probability of getting Hepatitis C is 0.6433, which interprets that the less exposure to hazard factors lead to lower probability of getting Hepatitis C in Males.

Table 4.34: Hosmer - Lemeshow Statistic

$\hat{C}$	**P-value**
7.393	.495

The H.L. (Hosmer Lemeshow) statistic has a value i.e. $\hat{C}$ = 7.393 & P-value evaluated from chi square distribution having d.f.= 8 is 0.495, that leads to accept H_0 and specifies that the model has been fitted well.

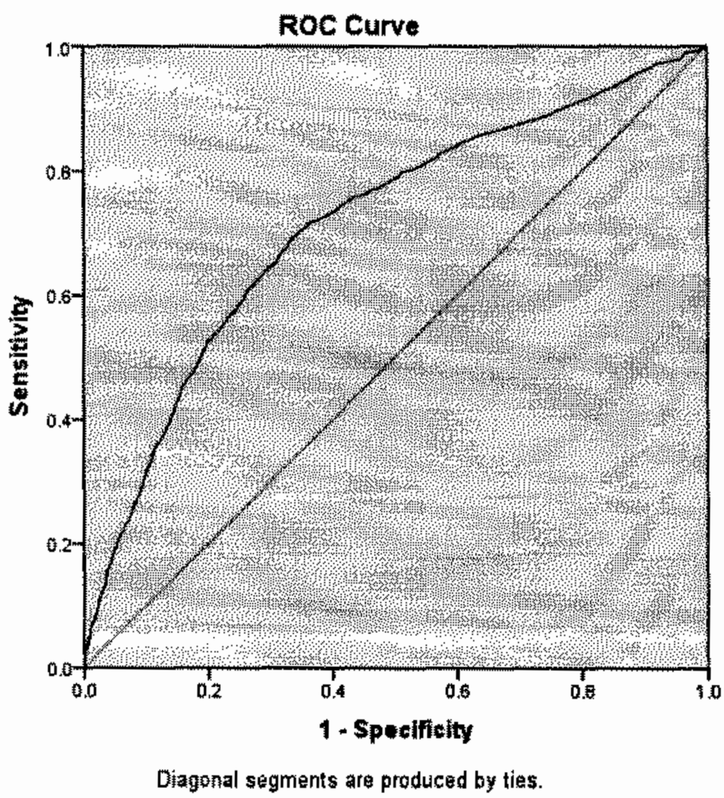


Figure 4.2: R.O.C. Curve of Hepatitis C (for males)

Table 4.35: Area under the Curve (A.U.C.)

Test Result Variable(s): Predicted probability				
A.U.C.	S.E.	Asymptotic Sig.	Asymptotic 95% C.I.	
			Lower Bound	Upper Bound
0.733	0.007	.000	0.719	0.746

Table 4.35 indicates the area under the R.O.C. curve. Since A.U.C. is 0.733, which is significantly different from 0.5, depicted an acceptable discrimination. The significance of A.U.C. (p-value=0.000) provides a strong evidence of the accuracy of fitted model. This means that the data has a specified logistic regression distribution.

4.4 Analysis of Hepatitis B Hazard Factors

The number of risk factors were divided into Demographic, socio economic and clinical factors. The data were analyzed according to Descriptive as well as Inferential Statistics.

Table 4.36: Overall Possible Hazard Factors of Hepatitis B in Punjab

		Hepatitis B Test Result		
		negative	**positive**	**Total**
Age Group	under 5	1954	34	1988
	5-19	9350	188	9538
	20-29	4471	100	4571
	30-39	2692	108	2800
	40-49	1952	106	2058
	50-59	1156	61	1217
	above 60	1282	69	1351
Total		**22857**	**666**	23523
Ear & Nose Piercing females age > 5 years	No	2746	41	2787
	Yes	7118	220	7338
Total		**9864**	**261**	10125
History of Jaundice	no	22787	657	23444
	yes	70	9	79
Total		**22857**	**666**	23523
History of Hospitalization	no	21981	631	22612
	yes	876	35	911
Total		**22857**	**666**	23523
Had Tattooing/Acupuncture > 5 years age	no	20825	631	21456
	yes	75	4	79
Total		**20900**	**635**	21535
Knowledge about Hepatitis	no	13481	357	13838
	yes	9376	309	9685
Total		**22857**	**666**	23523
History of Haemodialysis	no	22827	664	23491
	yes	30	2	32
Total		**22857**	**666**	23523
	no	10909	406	11315

Share Cigarettes/Hookah age > 19 years	yes	651	38	689
Total		**11560**	**444**	12004
Share ToothBrush/Razor/etc age > 5 years	no	20587	622	21209
	yes	316	10	326
Total		**20903**	**632**	21535
Marital Status	Unmarried	3077	55	3132
	Married	8483	389	8872
Total		**11560**	**666**	12004
Family Suffering from Hepatitis	no	14379	351	14730
	yes	508	39	547
	don't know	7970	276	8246
Total		**12037**	**391**	12428
History of IntraMuscular Injections	no	6181	157	6338
	yes	16432	507	16939
	dont know	244	2	246
Total		**22857**	**666**	23523
Use of IntraMuscular Injections	none	6425	159	6584
	less than 5	11369	336	11705
	5-10	3975	136	4111
	greater than 10	1088	35	1123
Total		**22857**	**666**	23523
Type of Syringe used	don't know	8125	210	8335
	new/disposable	7768	215	7983
	re-use	6964	241	7205
Total		**22857**	**666**	23523
Level of Education	illiterate	9764	350	10114
	below matric	9985	256	10241
	matric & above	2661	55	2716
	graduate & above	447	5	452
Total		**22857**	**666**	23523
TYPE OF HOUSE	Kacha	4391	140	4531
	Pacca	12919	338	13257
	Semi-pacca/kacha	4875	178	5053
	Well furnished	672	10	682
Total		**22857**	**666**	23523
DRINKING SOURCE	Others	638	10	648
	piped in Dwelling	15705	445	16150

	Public Tap	4246	101	4347
	Spring/Pound	151	8	159
	Tanker, Vendor	671	36	707
	Well	1446	66	1512
Total		**22857**	**666**	23523
Source of Shave Males age > 19 Years	None	2012	63	2075
	Home	982	39	1021
	Barber	2110	106	2216
	Both	1001	55	1056
Total		**6105**	**263**	**6368**

4.5 Descriptive Section (Hepatitis B)

In Descriptive section, percentages & counts were evaluated for different risk factors.

4.5.1 Rate of Pervasiveness of Hepatitis B in Punjab according to Demographic Factors

The pervasiveness percentage of Hepatitis B in Punjab according to age, gender and marital status are tabulated below,

Table 4.37: Pervasiveness of Hepatitis B in Punjab

PROVINCE	NUMBER OF SAMPLES	Pervasiveness of Hepatitis B		
		Positive	%	95% Confidence Interval
PUNJAB	23523	666	2.8	2.6 - 3.0

Table 4.37 showed that the numbers of positive cases screened were 666 out of 23523 samples. The Pervasiveness rate of Hepatitis B in Punjab was 2.8% with the confidence interval (2.6 - 3.0).

Table 4.38: Pervasiveness of Hepatitis B according to Gender

GENDER	TOTAL	Pervasiveness of Hepatitis B		
		Positive	%	95% Confidence Interval
MALE	12428	391	3.15	3.0 – 3.3
FEMALE	11095	275	2.48	2.3 – 2.7
TOTAL	23523	666	2.8	2.6 - 3.0

Table 4.38 illustrated the Pervasiveness of Hepatitis B according to gender. The Pervasiveness rate of Hepatitis B in Punjab for male was 3.15% with a confidence interval (3.0 – 3.3), whereas the Pervasiveness rate of Hepatitis B in Punjab for female was 2.48% with a confidence interval (2.3 – 2.7), which concluded that the pervasiveness of Hepatitis B in Punjab was relatively higher in males as compared to females.

Table 4.39: Pervasiveness of Hepatitis B according to Marital Status

MARITAL STATUS Age > 19 Years	TOTAL	Pervasiveness of Hepatitis B		
		Positive	%	95% Confidence Interval
NEVER MARRIED	3132	55	1.75	1.61 – 1.96
EVER MARRIED	8872	389	4.38	4.12 – 4.60
TOTAL	12004	444	3.7	3.5 – 3.9

Table 4.39 described the Pervasiveness of Hepatitis B according to Marital Status. The Pervasiveness rate of Hepatitis B in never-married group was 1.75%, whereas the Pervasiveness rate of Hepatitis B in ever-married group was 4.38%. Ever-married persons were more likely to have Hepatitis B than Never-married persons.

Table 4.40: Pervasiveness of Hepatitis B according to Age

AGE	TOTAL	Pervasiveness of Hepatitis B		
		Positive	%	95% Confidence Interval
under 5	1988	34	1.71	1.60 – 1.82
5-19	9538	188	1.97	1.85 – 1.15
20-29	4571	100	2.19	2.01 – 2.43
30-39	2800	108	3.86	3.75 – 4.10
40-49	2058	106	5.15	5.01 – 5.33

50-59	1217	61	5.01	4.9 1– 5.12
Above 60	1351	69	5.11	4.95 – 5.20
TOTAL	23523	666	2.8	2.6 - 3.0

Table 4.40 indicated the Pervasiveness of Hepatitis B according to age. Hepatitis B was less than 2% in age groups under 5 and 5 – 19, but that rate tend to increase from 30 years of age at greater pace and achieved its maximum value at the age group of 40-49. The rate of Hepatitis B was also greater than 5% at the age groups of 50-59 and above 60.

4.5.2 Pervasiveness Rate of Hepatitis B in Punjab according to Socio-economic Factors

The pervasiveness percentage of Hepatitis B in Punjab according to socioeconomic factors are tabulated below,

Table 4.41: Pervasiveness of Hepatitis B according to Level of Education

Level Of Education	TOTAL	Pervasiveness Of Hepatitis B		
		Positive	%	95% Confidence Interval
Illiterate	10114	350	3.46	2.6 – 4.8
Below Matric	10241	256	2.5	2.2 – 2.8
Matric & Above	2716	55	2.03	1.8 – 2.2
Graduate & Above	452	5	1.11	1.0 –1.3
TOTAL	23523	666	2.8	2.6 - 3.0

Table 4.41 demonstrated the Pervasiveness rate of Hepatitis B according to the level of Education. The Pervasiveness of Hepatitis B was maximum (i.e. 3.46%) in the group of illiterate people whereas it was minimum (i.e. 1.11%) in the group of

graduate and above. The Pervasiveness rate of Hepatitis B in Punjab decreased with the increase in the level of Education.

Table 4.42: Pervasiveness of Hepatitis B according to House Material

House Material	TOTAL	Pervasiveness Of Hepatitis B		
		Positive	%	95% Confidence Interval
Kacha	4531	140	3.10	2.6 – 4.8
Pacca	13257	338	2.55	2.4 – 2.8
Semi-pacca/kacha	5053	178	3.52	3.41 – 3.64
Well-furnished	682	10	1.47	1.38 –1.60
TOTAL	23523	666	2.8	2.6 - 3.0

Table 4.42 revealed the Pervasiveness rate of Hepatitis B according to the House Material. The Pervasiveness of Hepatitis B was maximum (i.e. 3.10%) in the group of people living in Kacha house, whereas it was minimum (i.e. 1.47%) in the group of people living in Well-furnished house.

Table 4.43 Pervasiveness of Hepatitis B according to Drinking Source

Drinking Source	TOTAL	Pervasiveness Of Hepatitis B		
		Positive	%	95% Confidence Interval
Others	648	10	1.54	1.4 – 1.7
piped in Dwelling	16150	445	2.76	2.5 – 3.0
Public Tap	4347	101	2.32	2.12 – 2.52
Spring/Pound	159	8	5.03	4.7 – 5.18
Tanker/Vendor	707	36	5.09	4.8 – 5.2
Well	1512	66	4.37	4.2 – 4.6
Total	23523	**666**	2.8	2.6 – 3.0

Table 4.43 determined the Pervasiveness of Hepatitis B according to drinking source. Hepatitis B was maximum (i.e. 5.09%) among the group of people who used water from tanker or taken from vendor but that rate tend to decrease among the group who used proper connected piped water in dwelling.

Table 4.44: Pervasiveness of Hepatitis B according to Knowledge about Hepatitis

Knowledge about Hepatitis	TOTAL	Pervasiveness Of Hepatitis B		
		Positive	%	95% Confidence Interval
No	13838	357	2.58	2.42 – 2.74
Yes	9685	309	3.20	3.05 – 3.35
Total	23523	**666**	2.8	2.6 – 3.0

Table 4.44 illustrated the Pervasiveness rate of Hepatitis B according to the Knowledge about Hepatitis B. The Pervasiveness of Hepatitis B was higher (i.e. 3.20%) in the group of people who had knowledge about Hepatitis B. Hence, the Pervasiveness rate of Hepatitis B in Punjab decreased with the increase in the awareness of Hepatitis B.

Table 4.45: Pervasiveness of Hepatitis B according to Divisions of Punjab

AGE	TOTAL	Pervasiveness of Hepatitis B		
		Positive	%	95% Confidence Interval
Rawalpindi	3090	96	3.11	2.91 – 3.20
Gujranwala	4852	150	3.09	2.87 – 3.18
Sargodha	2308	53	2.30	2.21 – 2.41
Faisalabad	1585	37	2.33	2.22 – 2.42
Lahore	3545	68	1.92	1.80 – 2.05

Multan	4337	139	3.21	3.11 – 3.32
Sahiwal	1387	34	2.45	2.32 – 2.57
Bahawalpur	2219	83	3.74	3.61 – 3.89
D.G. Khan	200	6	3	2.86 – 3.12
TOTAL	23523	666	2.8	2.6 - 3.0

Table 4.45 revealed the rate of Pervasiveness of Hepatitis B according to the Divisions of Punjab. The Pervasiveness of Hepatitis B was maximum (i.e. 3.74%) in Bahawalpur Division, whereas it was minimum in Lahore Division (i.e. 1.92%). The rate of Pervasiveness of Hepatitis B in Rawalpindi, Gujranwala, Sargodha, Faisalabad, Multan, Sahiwal and D.G. Khan were 3.11%, 3.09%, 2.30%, 2.33%, 3.21%, 2.45% and 3% respectively.

4.6 Inferential Section (Hepatitis B)

In Inferential section, bivariate analysis and multivariate analysis were applied independently to determine the significant risk factors. Inferential section is the major section of the research study which may be distributed as

- Bivariate Analysis
- Multivariate Analysis

4.6.1 Bivariate Analysis (for Females)

In this section, the association between each hazard factor using odds ratio was tested separately with the Hepatitis B (for females). Wald test is the statistical procedure used in bivariate analysis to test the association between Hepatitis and a Risk Factor.

4.6.1.1 Associations between Hepatitis B & Risk Factors

Table 4.46: Hepatitis B vs. Age Group

Hepatitis B (Females)		B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
							Lower	Upper
AGE GROUP	5-19	.214	.292	.539	.463	1.239	.699	2.196
	20-29	.488	.305	2.561	.110	1.629	.896	2.961
	30-39	.872	.308	7.997	.005	2.391*	1.307	4.373
	40-49	1.084	.314	11.920	.001	2.956*	1.598	5.469
	50-59	.912	.349	6.829	.009	2.490*	1.256	4.934
	Above 60	1.114	.341	10.696	.001	3.046*	1.563	5.938

Table 4.46 illuminated the independent association between Hepatitis B (females) and Age Group. The results showed that the pervasiveness of Hepatitis B in females was significantly higher at age 30 or above. The binary logistic regression coefficients (Odds Ratios) of age groups 30-39, 40-49, 50-59 and above 60 are 0.872 (2.391), 1.084 (2.956), 0.912 (2.490) & 1.114 (3.046) respectively. Therefore, we reject the null hypothesis and concludes that Hepatitis B (females) and the Age Groups 30-39, 40-49, 50-59 and above 60 have significant association at 1% level of significance.

Table 4.47: Hepatitis B vs. Ear/Nose Piercing

Hepatitis B (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Ear/Nose Piercing	0.73	0.155	22.276	.000	2.07*	1.535	2.823

Table 4.47 showed the relationship between Hepatitis B and Ear/Nose Piercing independently. The binary logistic regression coefficient for Ear/Nose Piercing was 0.73 and Odds Ratio (O.R.) = 2.07 with a confidence interval (1.535 – 2.823). The result suggested that the risk of Hepatitis B becomes double in those females who had got their Ear/ Nose Pierced. We assume that the base category was no Ear/Nose Piercing. The value of Wald statistic was 22.276 which was statistically significant. Therefore, we reject the null hypothesis and find that Hepatitis B and Ear/ Nose Piercing has significant association at 1% level of significance.

Table 4.48: Hepatitis B vs. History of Jaundice

Hepatitis B (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
History of Jaundice	0.851	0.394	21.89	.000	2.34*	1.08	5.07

Table 4.48 revealed the independent association of Hepatitis B and those females which ever had jaundice in their past. The binary logistic regression coefficient for Jaundice was 0.851 and Odds Ratio (O.R.) =2.34 with a confidence interval (1.08 – 5.07). It concluded that the female with a history of jaundice had 2.34 times higher risk of obtaining Hepatitis B as compared to those who had no jaundice ever. The value of Wald statistic was 21.89 which was statistically significant. Consequently, we reject the null hypotheses and concludes that Hepatitis B and History of Jaundice has significant association at 1% level of significance.

Table 4.49: Hepatitis B vs. Family History of Hepatitis

Hepatitis B (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Family history of Hepatitis	1.54	0.245	39.44	.000	4.65*	2.88	7.51

Table 4.49 demonstrated the autonomous relationship between Hepatitis B and the family history of Hepatitis. The binary logistic regression coefficient was 1.54 and Odds Ratio (O.R.) = 4.65 with a confidence interval (2.88 – 7.51), which concluded that the female with a family history of Hepatitis had 4.65 times more chances of getting Hepatitis B. We assume that the base category was no family history of

Hepatitis. The value of Wald statistic was 39.44 which was significant. For that reason, we reject the null hypotheses and find that Hepatitis B and family history of Hepatitis has significant relationship at 1% level of significance.

Table 4.50: Hepatitis B vs. Marital Status

Hepatitis B (female)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Ever-Married	1.03	0.102	30.80	.000	2.8*	1.56	2.53

Table 4.50 described the independent association between Hepatitis B and the Marital Status. The binary logistic regression coefficient of females which were ever married was 1.03 and Odds Ratio (O.R.) = 2.8 with a confidence interval (1.56 – 2.53). The result suggested that there was 2.8 times greater risk of Hepatitis B in those females who were ever married. We assume that the base category was no history of marriage. The value of Wald statistic was 30.80 which was highly significant. Therefore, we reject the null hypothesis and analyzes that Hepatitis B and Ever-Married females have significant association at 1% level of significance.

Table 4.51: Hazard Factors of Hepatitis B Transmission in Females

Risk Factors		Hepatitis B		Total	O.R.	95% C.I.
		Negative	Positive			
Age Group	Under 5	956	14	970	Reference	
	5-19	4409	80	4489	1.239	.699 – 2.196
	20-29	2096	50	2146	1.629	.896 – 2.961
	30-39	1314	46	1360	2.391*	1.307 – 4.373
	40-49	924	40	964	2.956*	1.598 – 5.469
	50-59	576	21	597	2.490*	1.256 – 4.934
	Above 60	538	31	569	3.046*	1.563 – 5.938
				11095		
Ear & Nose Piercing (Age Above 19 Years)	No	2746	41	2787	Reference	
	Yes	7118	220	7338	2.07*	1.535 - 2.823
				10125		
Had Jaundice	No	9220	212	9432	Reference	
	Yes	130	7	137	2.34*	1.08 – 5.07
	Don't Know	1470	56	1526		
				11095		
Family Suffering From Hepatitis	No	6748	124	6872	Reference	
	Yes	246	21	267	4.65*	2.88 - 7.51
	Don't Know	3826	130	3956		
				11095		
Marital Status (Age Above 19 Years)	Unmarried	1573	23	1596	Reference	
	Married	3882	158	4040	2.80*	1.56 - 2.53
				5636		

O.R. = Odds Ratio, C.I. = Confidence Interval, * shows significant at 1% level

4.6.2 Multivariate Analysis of Hepatitis B (for Females)

Table 4.52 showed the multivariate model for Hepatitis B (Females). Forward selection criteria is adapted to select the model. Hepatitis B is taken as a binary response variable, whereas all other factors (i.e. Risk Factors) are taken as independent variables.

Table 4.52: Multivariate Model of Hepatitis B (for Females)

Risk Factors	B	S.E.	Wald	p-value
Age	.018	0.005	12.457*	0.006
Piercing	0.470	0.167	8.102*	0.005
Jaundice's History	0.494	0.126	15.371*	0.0004
Ever Married	0.625	0.138	20.491*	0.000
Family History of Hepatitis	1.365	0.251	29.552*	0.000
Constant	-1.831	.28	39.637*	0.000

Note: * shows significant at 0.01 level.

The equation for overall fitted Logit Model for Hepatitis B (for Females) is

$$\ln\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = -1.831 + 0.018\,(Age) + 0.470\,(Piercing) + 0.494\,(Jaundice) + 0.625\,(Ever\ Married) + 1.365\,(Family\ History\ of\ Hepatitis) \qquad (4.3)$$

4.6.2.1 Estimated Probabilities

The equation (4.3) can be used to estimate the different probabilities of Hepatitis B in females in the absence or presence of significant hazard factors. Few examples are stated below,

- If Age=50, Piercing=1, Jaundice=1, Ever Married=1 and Family history of Hepatitis=1, then the estimated probability is

$$\hat{\pi} = \frac{1}{1 + e^{(-2.023)}} = 0.8832$$

It means that the probability of getting Hepatitis B in female is 0.8832 in the existence of above declared hazard factors.

- If Age=50, Piercing=1, Jaundice=0, Ever Married=1 and Family history of Hepatitis=0, then the estimated probability is

$$\hat{\pi} = \frac{1}{1 + e^{(-0.164)}} = 0.5409$$

It means that the probability of getting Hepatitis B in female is 0.5409. It reveals that the less exposure to hazard factors lead to lower probability of getting Hepatitis B in females.

Table 4.53: Hosmer - Lemeshow Statistic

$\hat{C}$	P-value
8.624	.375

The H.L. (Hosmer Lemeshow) statistic has a value i.e. $\hat{C}$ = 8.624 & P-value evaluated from chi square distribution is 0.375, that leads to accept H_0 and specifies that the model has been fitted well.

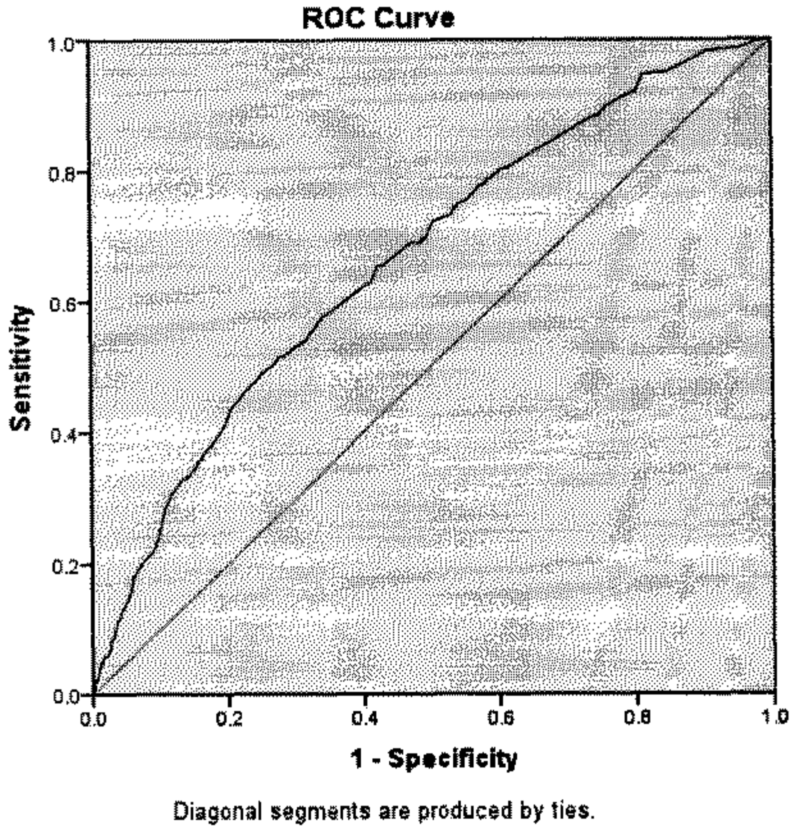


Figure 4.3: R.O.C. Curve of Hepatitis B (for females)

Table 4.54: Area under the Curve (A.U.C.)

Test Result Variable(s): Predicted probability				
A.U.C.	**S.E.**	**Asymptotic Sig.**	**Asymptotic 95% C.I.**	
			Lower Bound	**Upper Bound**
0.685	0.013	.000	0.651	0.715

Table 4.54 indicates the area under the R.O.C. curve. Since A.U.C. is 0.685, which is significantly different from 0.5, depicted an acceptable discrimination. The significance of AUC (p-value=0.000) provides a strong evidence of the accuracy of fitted model. This means that the data has a specified logistic regression distribution.

4.6.3 Bivariate Analysis of Hepatitis B (for Males)

In this section, the association between each hazard factor using odds ratio was tested separately with the Hepatitis B (for Males). Wald test is the statistical procedure used in bivariate analysis to test the association between Hepatitis and a Risk Factor.

4.6.3.1 Association between Hepatitis B & Risk Factors

Table 4.55: Hepatitis B vs. Age Group

Hepatitis B (Males)		B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
							Lower	Upper
AGE GROUP	5-19	.088	.246	.129	.720	1.092	.675	1.769
	20-29	.054	.267	.040	.841	1.055	.625	1.781
	30-39	.816	.261	9.813	.002	2.262*	1.357	3.768
	40-49	1.164	.259	20.195	.000	3.204*	1.928	5.323
	50-59	1.236	.279	19.651	.000	3.441*	1.993	5.944
	Above 60	1.105	.273	16.364	.000	3.018*	1.767	5.154

Table 4.55 described the autonomous association between Hepatitis B (Males) and Age Group. The results revealed that the pervasiveness of Hepatitis B in males was significantly higher at age 30 or above. The logistic coefficients (Odds Ratios) of age groups 30-39, 40-49, 50-59 and above 60 are 0.816 (2.262), 1.164 (3.204), 1.236 (3.441) & 1.105 (3.018) respectively. As a result, we reject the null hypothesis and concluded that Hepatitis B (males) and the Age Groups 30-39, 40-49, 50-59 and above 60 have significant association at 1% level of significance.

Table 4.56: Hepatitis B vs. Shaving

Hepatitis B (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Home	0.24	0.171	4.96	0.081	1.27	1.01	2.01
Barber	0.47	0.109	22.47	0.006	1.604*	1.47	2.26
Both	0.56	0.084	28.47	0.002	1.754*	1.61	2.35

Table 4.56 demonstrated the independent relationship of Hepatitis B (Males) and Shaving. The binary logistic regression coefficient of males who shaved at Home was 0.24 and Odds Ratio (O.R.) = 1.27 with the confidence interval (1.01 – 2.01). The value of Wald statistic was 4.96 which was statistically insignificant. Therefore, we accept the null hypothesis that Hepatitis B and the males shave at Home have no significant association at 1% level of significance.

The logistic coefficient of males who got their shaves from Barber was 0.47 and Odds Ratio (O.R.) = 1.604 with a confidence interval (1.47 – 2.26), which concluded that the Male with a history of Barber Shave had approximately double risk of getting Hepatitis B. We assume that the base category was no History of Shaving. The value of Wald statistic was 22.47 which is statistically significant. For that reason, we reject the null hypothesis and interprets that Hepatitis B and males who get their shave from Barber has significant association at 1% level of significance.

The binary logistic regression coefficient of males who get their shave from Barber and at Home was 0.56 and Odds Ratio (O.R.) = 1.754 with a confidence interval (1.61 – 2.35), which found that the Male with a history of Barber Shave had approximately double risk of getting Hepatitis B. We assume that the base category was no History of Shaving. The value of Wald statistic was 22.47 which was significant. Consequently, we reject the null hypothesis and find that Hepatitis B and males who get their shave from Barber and at home have significant relationship at 1% level of significance.

Table 4.57: Hepatitis B vs. History of Jaundice

Hepatitis B (Males)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
History of Jaundice	0.984	0.252	15.288	.000	2.68*	1.63	4.38

Table 4.57 revealed the independent association between Hepatitis B and those Males which ever had jaundice in their past. The binary logistic regression coefficient for Jaundice was 0.984 and Odds Ratio (O.R.) = 2.68 with a confidence interval (1.63 – 4.38), which concluded that the Male with a history of jaundice has 2.68 times more risk of obtaining Hepatitis B as compared to those who had no jaundice ever. The value of Wald statistic was 15.288 which was statistically significant. Therefore, we reject the null hypothesis and analyze that Hepatitis B and History of Jaundice has significant association at 1% level of significance.

Table 4.58: Hepatitis B vs. Marital Status

Hepatitis B (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Ever Married	1.26	0.25	46.43	.000	3.53*	3.02	3.96

Table 4.58 illustrated the autonomous association of Hepatitis B and the Marital Status. The binary logistic regression coefficient of Males which are ever married was 1.26 and Odds Ratio (O.R.) = 3.53 with a confidence interval (3.02 – 3.96), which

showed that the married male had more than three times risk of getting Hepatitis B as compared to those who had no history of Marriage. The value of Wald statistic was 46.43 which was statistically significant. Therefore, we reject the null hypothesis that Hepatitis B and Ever Married Males have no significant effect at 1% level of significance.

Table 4.59: Hepatitis B vs. Share Cigarettes/ Hookah

Hepatitis B (Males)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Share cigarettes/ hookah	0.71	0.21	17.98	0.000	2.03*	1.67	2.46

Table 4.59 illuminated the independent relationship of Hepatitis B and the Males who Share cigarettes/hookah. The binary logistic regression coefficient was 0.71 and Odds Ratio (O.R.) = 2.03 with a confidence interval (1.67 – 2.46). The result suggested that the risk of Hepatitis B became double approximately in those males who Share cigarettes/ hookah. We assume that the base category was no sharing. The value of Wald statistic was 17.98 which was statistically significant. And so, we reject the null hypothesis and conclude that Hepatitis B and sharing cigarettes/ hookah have significant relationship at 1% level of significance.

Table 4.60: Hepatitis B vs. Family Suffering from Hepatitis

Hepatitis B (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Family Suffering from Hepatitis	0.837	0.253	21.96	0.000	2.31*	1.41	3.79

Table 4.60 investigated the relationship of Hepatitis B and the family suffering from Hepatitis. The binary logistic regression coefficient was 0.837 and Odds Ratio (O.R.) = 2.31 with a confidence interval (1.41 – 3.79), which concluded that the Male with a family history of Hepatitis had 2.31 times higher chances of getting Hepatitis B. We assume that the base category was no family history of Hepatitis. The value of Wald statistic was 21.96 which was statistically significant. Therefore, we reject the null hypothesis and find that Hepatitis B and family history of Hepatitis have significant association at 1% level of significance.

Table 4.61: Hepatitis B vs. Use of Injection

Hepatitis B (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
Use of Injections	0.357	0.124	8.35	0.004	1.43*	1.12	1.82

Table 4.61 showed the independent association between Hepatitis B and those Males who ever used Injections in their past. The binary logistic regression coefficient is 0.357 and Odds Ratio (O.R.) = 1.43 with a confidence interval (1.12 – 1.82), which concluded that the Male with a history of Injections had 1.43 times more risk of obtaining Hepatitis B as compared to those who had no history of Injections. The value of Wald statistic was 8.35 which was statistically significant. Therefore, we reject the null hypotheses and suggest that Hepatitis B and use of Injections have significant association at 1% level of significance.

Table 4.62: Hepatitis B vs. Syringe Type

Hepatitis B (Male)	B	S.E.	Wald Statistic	Sig.	Odds Ratio	95% C.I. for O.R.	
						Lower	Upper
New/Disposable	0.097	.13	0.56	0.452	1.102	0.856	1.420
Re-Use	0.37	0.124	8.821	0.003	1.45*	1.135	1.850

Table 4.62 described the autonomous relationship of Hepatitis B (Males) and Type of Syringe. The binary logistic regression coefficient of males who used New/Disposable syringe was 0.097 and Odds Ratio (O.R.) = 1.102 with a confidence interval (0.856 – 1.420). The value of Wald statistic was 0.56 which was statistically insignificant. Consequently, we accept the null hypothesis that Hepatitis B and the males who use New/Disposable syringe have no significant association at 1% level of significance.

The logistic coefficient of males who used re-use syringes was 0.37 and Odds Ratio (O.R.) = 1.45 with a confidence interval (1.135 – 1.850), which concluded that the Male with a history of re-use syringe had 1.45 times greater risk of having Hepatitis B. We assume that the base category was no history of Injections. The value of Wald statistic was 8.821 which was statistically significant. For that reason, we reject the null hypothesis and interpret that Hepatitis B and males who use re-use syringes have significant association at 1% level of significance.

Table 4.63: Hazard Factors of Hepatitis B (For Males)

Risk Factors		Hepatitis B		Total	O.R.	95% C.I.
		Negative	Positive			
Age Group	Under 5	1018	20	1038	Reference	
	5-19	4934	108	5042	1.092	.675 – 1.769
	20-29	2365	50	2415	1.055	.625 – 1.781
	30-39	1368	62	1430	2.262*	1.357–3.768
	40-49	1028	66	1094	3.204*	1.928–5.323
	50-59	580	40	620	3.441*	1.993–5.944
	Above 60	744	45	789	3.018*	1.767–5.154
				12428		
Shaving (Age Above 19 Years)	No	2012	63	2075	Reference	
	Home	982	39	1021	1.27	1.01 - 2.01
	Barber	2110	106	2216	1.604*	1.47 - 2.26
	Both	1001	55	1056	1.754*	1.61 – 2.35
				6368		
Had Jaundice	No	10236	302	10538	Reference	
	Yes	228	18	246	2.68*	1.63 - 4.38
	Don't Know	1573	71	1644	1.5	1.18 - 1.99
				12428		
Marital Status (Age Above 19 Years)	Unmarried	2004	32	2036	Reference	
	Married	4101	231	4332	3.53*	2.28 - 3.43
				6368		
Share Cigarettes/ Hookah (Age Above 19 Years)	No	10909	396	11305	Reference	
	Yes	651	48	699	2.03*	1.3 - 2.72
				12004		
Family Suffering From Hepatitis	No	7631	227	7858	Reference	
	Yes	262	18	280	2.31*	1.41 - 3.79
	Don't Know	4144	146	4290	1.19	0.96 - 1.46
				12428		

Use Of Injections	No	3424	86	3510	Reference	
	Yes	8467	304	8771	1.43*	1.12 - 1.82
	Don't Know	146	1	147	0.27	0.037-1.97
				12428		
Type Of Syringe Used	Don't Know	4455	124	4579	Reference	
	New/ Disposable	4010	123	4133	1.102	0.86- 1.42
	Re-Use	3572	144	3716	1.45*	1.14- 1.85
				12428		

O.R. = Odds Ratio, C.I. = Confidence Interval, * shows significant at 1% level

4.6.4 Multivariate Analysis of Hepatitis B (for Males)

Table 4.64 demonstrated the multivariate model for Hepatitis B (Males). Forward selection criteria is adapted to select the model. Hepatitis B is taken as a binary response variable, whereas all other factors (i.e. Risk Factors) are taken as independent variables.

Table 4.64: Multivariate Model of Hepatitis B (for Males)

Risk Factors	B	S.E.	Wald	p-value
Age	.047	0.015	9.818*	0.001
Home Shaving	0.12	0.083	2.072	0.315
Barber Shaving	0.38	0.097	15.35*	0.0007
Jaundice's History	0.815	0.256	10.128*	0.002
Ever Married	0.801	0.158	25.541*	0.000
Family History of Hepatitis	0.697	0.257	7.346*	0.007
Constant	-2.587	0.432	35.017*	0.000

Note: * shows significant at 0.01 level.

The equation for overall fitted Logit Model for Hepatitis B (for Males) is

$$\ln\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = -2.857 + 0.047\,(Age) + 0.38\,(Barber\ Shaving) + 0.815\,(Jaundice) + 0.801\,(Ever\ Married) + 0.697\,(Family\ History\ of\ Hepatitis) \quad (4.4)$$

4.6.4.1 Estimated Probabilities

The equation (4.4) can be used to estimate the different probabilities of Hepatitis B in Males in the absence or presence of significant hazard factors. Few examples are stated below,

- If Age=50, Barber Shaving=1, Jaundice=1, Ever Married=1 and Family history of Hepatitis=1, then the estimated probability is

$$\hat{\pi} = \frac{1}{1 + e^{(-2.816)}} = 0.9435$$

It means that the probability of getting Hepatitis B in male is 0.9435 in the existence of above declared hazard factors.

- If Age=50, Barber Shaving=1, Jaundice=0, Ever Married=1 and Family history of Hepatitis=0, then the estimated probability is

$$\hat{\pi} = \frac{1}{1 + e^{(-0.674)}} = 0.6624$$

It means that the probability of getting Hepatitis B in male is 0.6624. It reveals that the less exposure to hazard factors lead to lower probability of getting Hepatitis B in Males.

Table 4.65: Hosmer - Lemeshow Statistic

$\hat{C}$	P-value
6.916	0.438

The H.L. (Hosmer Lemeshow) statistic has a value i.e. $\hat{C}$ = 6.916 & P-value evaluated from chi square distribution is 0.438, that leads to accept H_0 and specifies that the model has been fitted well.

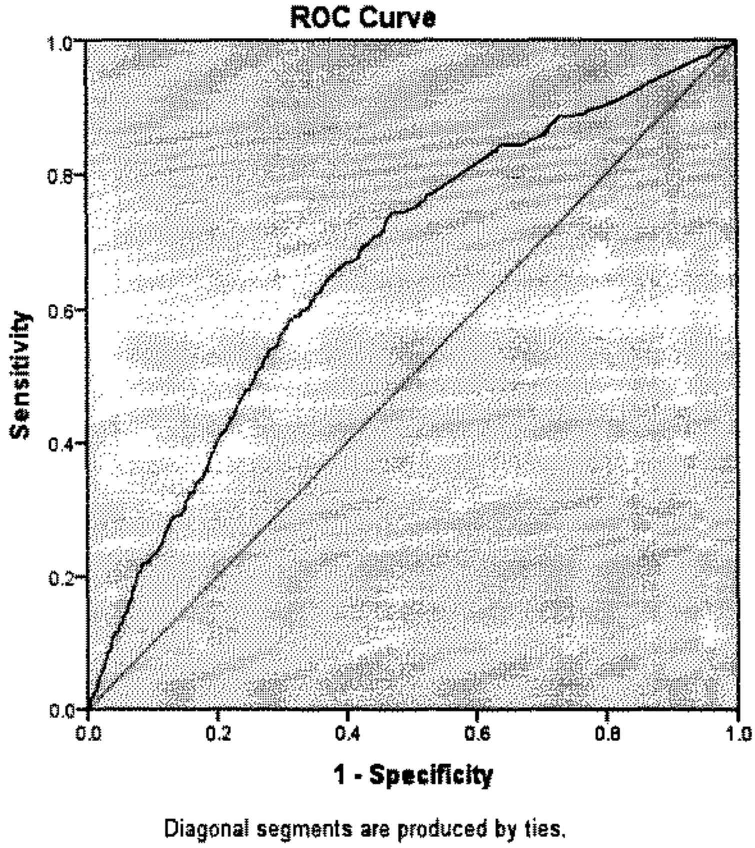


Figure 4.4: R.O.C. Curve of Hepatitis B (for males)

Table 4.66: Area under the Curve (A.U.C.)

Test Result Variable(s): Predicted probability				
A.U.C.	S.E.	Asymptotic Sig.	Asymptotic 95% C.I.	
			Lower Bound	Upper Bound
0.666	0.014	.000	0.638	0.693

Table 4.66 indicates the area under the R.O.C. curve. Since A.U.C. is 0.666, which is significantly different from 0.5, depicted an acceptable discrimination. The significance of A.U.C. (p-value=0.000) provides a strong evidence of the accuracy of fitted model. This means that the data has a specified logistic regression distribution.

4.7 Conclusion

The pervasiveness rate of Hepatitis in Punjab is relatively high as compared to the other provinces in Pakistan. Therefore, this research study is the first large sample provincial based study with the perseverance to reveal the significant hazard factors which are related to the pervasiveness of Hepatitis B & C among the people of Punjab. But if we consider prior research studies on Hepatitis, many had limitations & comprised of small samples which did not represent the overall population. In this study, the factors were divided as Demographic, Socio-economic & Clinical risk factors. The data were analyzed according to Descriptive as well as Inferential Statistics.

The *demographic factors* illuminated that Hepatitis B & C were relatively higher in males as compared to females, ever-married persons were more likely to have Hepatitis than never-married persons. Hepatitis B & C were increased with the increase in age. According to the *socio-economic factors*, the rates of Hepatitis B & C were lower in the group of graduate people, but achieved the minimum rates among the group of people living in well-furnished houses.

In Inferential section, bivariate & multivariate analysis were applied separately to determine the principal hazard factors for both males and females. In bivariate analysis, association between Hepatitis and each hazard factor was tested independently. In bivariate analysis, the common significant risk factors of Hepatitis B & C (for females) were age group (above 60 years), ear and nose piercing, jaundice's history, ever-married and family history of hepatitis. Whereas, the common significant risk factors of Hepatitis B & C (for males) were age group (above 50 years), barber shaving, re-use syringes, jaundice's history, share cigarettes/hookah, ever-married, family history of hepatitis.

The Multivariate model of Hepatitis C (for females) demonstrated that age, re-use syringes, ear and nose piercing, jaundice's history, ever-married and family history of hepatitis were the significant hazard factors associated with the Hepatitis C (for females). However, the dominant hazard factors related to Hepatitis C (for males) were age, barber shaving, re-use syringes, jaundice's history, share cigarettes/hookah, ever-married, family history of hepatitis and tattooing/acupuncture.

The Multivariate analysis of Hepatitis B (for females) revealed that age, ear and nose piercing, jaundice's history, ever-married and family history of hepatitis were the momentous hazard factors concomitant with the Hepatitis B (for females). In spite of this, the substantial hazard factors interrelated with Hepatitis B (for males) were age, barber shaving, jaundice's history, ever-married and family history of hepatitis. Mutual hazard factors of hepatitis B & C for both genders were age, jaundice's history, ever-married and family history of hepatitis. Barber shaving in males played a crucial role in the pervasiveness of both Hepatitis B & C, whereas, ear and nose piercing was the major cause of extensiveness of Hepatitis in females.

4.8 Recommendations

Hepatitis B & C are prevailing day by day in Pakistan. Lack of awareness about the Hepatitis played a key role in the extensiveness of liver cancers. The need of an hour is to educate people about the factors that are responsible for the commonness of Hepatitis. Frequent use of injections and cure from intravenous drips should be eluded.

Re-use of filthy syringes must be avoided, while disposable and good quality syringes are recommended. Also the reuse of blades, razors and needles should be discouraged. It is mandatory to disinfect the surgical apparatuses.

Barber shaves, ear & nose piercing with proper hygienic equipment is recommended. Both Barber and a customer are at high risk due to non-sterile instruments. Appropriate blood screening and national as well as global precautionary measures against the pervasiveness of Hepatitis should be preferred. Proper vaccination against hepatitis is also recommended.

Chapter 5
References

Abbasi, A., Saleem, A., Butt, S.A., Habib, N., Hussain, W. and Arooj, S. (2013). "A Statistical Study Regarding Knowledge and Risk Factors of Hepatitis B and C in Muzaffarabad Azad Jammu & Kashmir", *IOSR Journal of Nursing & Health Science*, **1**(3), 55-62.

Abbas, Z., Jeswani, N.L., Kakepoto, G.N., Islam, M., Mehdi, K. and Jafri, W. (2008). "Prevalence and mode of spread of hepatitis B and C in rural Sindh, Pakistan", *Tropical Gastroenterology*, **29**(4), 210–216.

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. John Wiley and Sons, Inc., Hoboken, New Jersey.

Akhtar, S., Younus, M., Adil, S., Jafri, S.H. and Hassan, F. (2004). "Hepatitis C virus infection in asymptomatic male volunteer blood donors in Karachi, Pakistan", *Journal of Viral Hepatitis,* **11**(6), 527-535.

Ali, S.A., Donahue, R.M.J., Qureshi, H. and Vermund, S.H. (2009). "Hepatitis B and Hepatitis C in Pakistan: Prevalence and risk factors", *International Journal of Infectious Diseases,* **13**(1), 9-19.

Bari, A., Akhtar, S., Rahbar, M.H. and Luby, S.P. (2001). "Risk factors for Hepatitis C virus in male adults in Rawalpindi-Islamabad, Pakistan", *Tropical Medicine and International Health*, **6**(9), 732-738.

Christensen, R. (1997). *Log-Linear Models and Logistic Regression.* Springer-Verlag New York, Inc.

David, W. and Hosmer, S. L. (1989). *Applied Logistic Regression Analysis*. John Wiley & Sons, New York.

Ghias, M., Pervaiz, M.K. and Aslam, A. (2010). "Risk Factors for Hepatitis C virus among Urban/Rural Settings of Patients Visiting Tertiary Care Hospitals at Lahore, Pakistan", *Journal of Statistics*, **17**, 33-46.

Ghias, M. and Pervaiz, M.K. (2009). "Identification of epidemiological risk factors for Hepatitis C in Punjab, Pakistan", *Journal of Ayub Medical College, Abbottabad*, **21**(2), 156-161.

Hafeez-ud-din, Siddique, T.H., Lahrasab, W. and Sharif M.A. (2012). "Prevalence of Hepatitis B and C in healthy adult males of paramilitary personnel in Punjab", *Journal of Ayub Medical College, Abbottabad*, **24**(3), 138-140.

Hanley, J.A. and McNeil, B.J. (1982). "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve", *Radiology*, **143**(1), 29-36.

Machado, D.F., Martins, T., Trevisol, D.J., Vieira e Silva, R.A., Narciso-Schiavon J.L., Schuelter-Trevisol, F. and Schiavon, L.L. (2013). "Prevalence and Factors Associated With Hepatitis B Virus Infection Among Senior Citizens in a Southern Brazilian City", *Kowsar Hepatitis Monthly*, **13**(5), 78-87.

Masud, I., Khan, H. and Khatak, A.M. (2004). "Relative frequency of Hepatitis B and C virus in patients with hepatic cirrhosis at DHQ teaching hospital D.I. Khan", *Journal of Ayub Medical College, Abbottabad*, **16**(1), 32-34.

Qureshi, H., Arif, A., Riaz, K., Alam, S.E., Ahmed, W. and Mujeeb, S.A. (2009). "Determination of risk factors for Hepatitis B & C in male patients suffering from chronic hepatitis", *BMC Research Notes*, **2**: 212.

Qureshi, H., Bile, K.M., Jooma, R., Alam, S.E. and Afridi H.U.R. (2010). "Prevalence of Hepatitis B and C viral infections in Pakistan: findings of a national survey appealing for effective prevention and control measures", *Eastern Mediterranean Health Journal*, **16**, 15-23.

Shah, F., Salih, M., Malik, I.A. and Hussain, I. (2002). "Increasing prevalence of chronic hepatitis and associated risk factors", *Pakistan Journal of Medical Research*, **41**(2), 55-70.

Shazi, L. and Abbas, Z. (2006). "Comparison of risk factors for hepatitis B and C in patients visiting a gastroenterology clinic", *Journal of the college of Physicians & Surgeons- Pakistan,* **16** (2), 104-107.

Syspa, V., Hadjipaschali, E. and Hatzakis, A. (2001). "Prevalence, risk factors and evaluation of a screening strategy for chronic hepatitis C & B virus infections in healthy company employees", *European Journal of Epidemiology,* **17**, 721-728.

World Health Organization. Report of Hepatitis C in Pakistan. August 08, 2012. Islamabad.

Zweig, M.H. and Campbell, G. (1993). "Receiver Operating Characteristic (ROC) plots: A Fundamental Evaluation Tool in Clinical Medicine", *Clinical Chemistry*, **39**(4), 561-577.