2019-09-25

# APPLICATION OF MULTIVARIATE MULTILEVEL MODEL ON LONGITUDINAL DATA ANALYSIS FOR HYPERTENSION AND TYPE-II DIABETES PATIENTS JOINTLY: A CASE OF AT DEBRE TABOR REFERRAL HOSPITAL, DEBRE TABOR, ETHIOPIA.

HAILEMICHAEL, MENBERU


BAHIR DAR UNIVERSITY

DEPARTMENT OF STATISTICS

# APPLICATION OF MULTIVARIATE MULTILEVEL MODEL ON LONGITUDINAL DATA ANALYSIS FOR HYPERTENSION AND TYPE-II DIABETES PATIENTS JOINTLY: A CASE OF AT DEBRE TABOR REFERRAL HOSPITAL, DEBRE TABOR, ETHIOPIA.

BY: HAILEMICHAEL MENBERU

ADVISOR: DEMEKE LAKEW (ASSI. PROFESSOR)

A THESIS RESEARCH SUBMITTED TO THE DEPARTMENT OF STATISTICS, COLLEGE OF SCIENCE, BAHIR DAR UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF SCIENCE IN BIOSTATISTICS

JUNE, 2019

BAHIR DAR, ETHIOPIA

## Declaration

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all sources of materials used for the thesis have been duly acknowledged. The assistance received during the course of these investigations has been duly acknowledged. Therefore, we recommend it to be accepted as fulfilling the thesis requirements.

Name: Hailemichael Menberu _____ _____

. Date Signature

This thesis has been submitted for examination with my approval as a University advisor.

Demeke Lakew _____ _____

Advisor (Assi. Professor) Date Signature

## Approval Sheet

We, the undersigned, members of the board of examiners of the final open defense by Hailemichael Menberu have read and evaluated his thesis entitled **"Application of multivariate multilevel model on longitudinal data analysis for hypertension and type-II diabetes patients jointly: a case of at Debre Tabor Referral Hospital, Debre Dabor, Ethiopia"** and examined the candidate. This is therefore to certify that the thesis has been accepted in partial fulfillment of the requirement for the degree of Master of Sciences in Statistics with specialization of Biostatistics.

| | | |
|---|---|---|
| Name of chairperson | _____ | _____ |
| . | Date | Signature |
| Name of external examiner | _____ | _____ |
| . | Date | Signature |
| Name of internal examiner | _____ | _____ |
| . | Date | Signature |

**Acknowledgments**

First and for most, I would like to extend my unshared thanks to the almighty God for providing me the opportunity for what I have achieved and for his mercy.

My heartfelt gratitude goes to my advisor Demeke Lakew (Assi. Professor) for his excellent constructive scientific guidance, tireless efforts and invaluable advice to make this work reality.

I would like to thank also all Bahir Dar University staff at the department of Statistics for inspiring and encouraging us to submit our proposal on time.

My sincere appreciation also goes to my friends and colleagues who explicitly or implicitly-backed my work in what so ever means.

Finally, I would like to extend my sincere thanks go to my family, who have taught me the important things in life and supported all my endeavors.

## *Abstract*

***Background:****Hypertension and type II diabetes are major cardiovascular morbidity and mortality risk factors in the world, especially in the low and middle income country which includes Ethiopia. The objectives of this study was to assess factors that affecting hypertension and type II diabetes jointly at Debre Tabor Referral Hospital; Northeast, Ethiopia.*

***Methods:****A hospital-based retrospective cohort study was conducted on 151 regular follow-up patients aged ≥18 from January 2017 to January 2019. In the study three different model were used; Thus, multivariate analysis were used for analyze association of two response over a time; multivariate growth curve analysis were used for determine number of follow-up and its effect; Lastly multivariate multilevel analysis were used to identify the risk factors of hypertension and type II diabetes.*

***Results:****The mean age of the patients was 47.4 years and 57.6% of them were females. The association between SBP and DBP were 0.446, 0.427, 0.407, 0.394, 0.385 and 0.379 over the six follow-up respectively. Likewise the association of SBP and T2DM, DBP and T2DM were also slightly decreased over a time. Beside from the six follow-up, only the first four follow up were statistically significant and in each follow-up the multivariate response were decreased by 5.175, 0.690, 0.092, 0.012 unit respectively.The study shows that their is a significant intercept variation($u_{0i}$=13.9967, p<0.0001 and $e_{0i}$=254.11, p<0.0001) and slop variation($u_{1i}$=-0.412, p=0.0088) for number of related disease, baseline stage of SBP and baseline stage of DBP. On the fixed estimate the variable age, residence, number of related disease, baseline stage of SBP, baseline stage of DBP and baseline T2DM were significantly associated for both hypertension and type II diabetes; but gender was associated only for type II diabetes by considering P<0.05.*

***Conclusion:*** *The findings in this study conducted that hypertension and type II diabetes are becoming a serious public health concern in the country. Hence, intervention measures should be undertaken at the community level; particular emphasis should be placed on prevention by introducing lifestyle modifications and creating awareness about the problem so that early detection and intervention is possible.*



***Keywords:*** *hypertension, type II diabetes, cardiovascular disease, Ethiopia, Debre Tabor*

## Acronyms

| | |
|---|---|
| AIDS | Acquired Immune Disease Syndrome |
| BMI | Body Mass Index |
| BP | Blood Pressure |
| CVD | Cardiovascular Diseases |
| DBP | Diastolic Blood Pressure |
| DIC | Deviance Information Criteria |
| DM | Diabetic Mellitus |
| GCA | Growth curve analysis |
| HIV | Human Immune Varies |
| HT | Hypertension |
| ICC | Intra Class Correlation |
| ICMR | Indian Council of Medical Research |
| IDH | Isolated Diastolic Hypertension |
| MLM | Multilevel Model |
| MVML | Multivariate Multilevel |
| SAS | Statistical Analysis System |
| SBP | Systolic Blood Pressure |
| SEM | Standard Errors of Means |
| SPSS | Statistical Package for Social Science |
| T2DM | Type 2 diabetes mellitus |
| UACR | Urinary Albumin Excretion |

# Contents

# List of Figures

# List of Tables

# CHAPTER ONE

# 1 INTRODUCTION

## 1.1 Background the study

Hypertension and diabetes is major cardiovascular risk factors in the world, especially in the low and middle income country which includes Ethiopia. As[29] hypertension and type II diabetes have been known to be major risk factors for cardiovascular mortality and morbidity in the world. Diabetes increases the risk of cardiovascular disease 66%, while high blood pressure (BP) is associated with a 72% rise in all-cause mortality risk and a 57% rise of cardiovascular events in diabetic individuals.

According to [16] study, hypertension in type II diabetes are interrelated diseases that strongly expose patients to highly increased risk of atherosclerotic cardiovascular and kidney disease. This association has been called the deadly duet in order to emphasis the increased cardiovascular risk when the two conditions co-exist. Hypertension affects approximately 70% of patients with diabetes and is approximately twice as common in persons with diabetes as in those without.

The appropriate management of the hypertension almost 70% of patients with type II diabetes remains controversial. Hypertension in type II diabetes occur together officially considered 'co morbidities'. Since, in diabetic patients increased fluid volume and arterial stiffness so that hypertension happens. Moreover, patients impaired insulin handling can directly causes increases in blood pressure and kidney is our body's important long-term blood pressure regulator. As a result diabetes mellitus patient has twice chance to attack hypertension[4].

Hypertension is one of the most common medical problems of chronically high blood pressure with different stage. Thus, Pre-hypertension: consistent readings of 120- 139/80-89 mmHg, Stage one hypertension: consistent (two or more consecutive) readings of 140-159/90-99 mmHg, and Stage two hypertension: consistent readings of 160/100 mmHg or higher[17].

Diabetes is also another medical problem of person has a high blood sugar level, either because the body doesn't produce enough insulin, or body cells don't properly respond insulin that is produced. The most common diabetes are, Type I diabetes, previously called insulin-dependent diabetes mellitus (IDDM) or juvenile onset diabetes, may account for 5 percent to 10 percent of all diagnosed cases of diabetes. Type II diabetes was previously called non-insulin-dependent diabetes mellitus (NIDDM) or adult-onset diabetes. Type II diabetes may account for about 90 percent to 95 percent of all diagnosed cases of diabetes. Gestational diabetes: it is when pregnant women, who have never had diabetes before and have high blood glucose level during pregnancy. This gestational diabetes may precede development of type II diabetes[8].

Detecting and managing hypertension in people with type II diabetes is one of the most effective measures to prevent adverse events, and pharmacotherapy is one of the most effective ways to maintain target BP levels in primary care. Increased blood pressure (BP) is a leading risk factor for death and disability, particularly in people with diabetes[38].

Hypertension is very common in people with diabetes, in both type I and type II diabetes, and it multiplies the incidence and mortality of severe cardiovascular diseases in diabetic people. All types of cardiovascular diseases are increased in people with diabetes, most problems of notably coronary heart disease and ischemic stroke that are the leading causes of death in diabetic patients[48].

In Ethiopia, few studies have evaluated that the prevalence of hypertension in type II diabetes. For instance, a study conducted in 1982 among Ethiopian Bank employees attending a clinic in Addis Ababa found prevalence of 3.8% and 1.2% for hypertension and diabetes respectively, which is very weak prevalence[35].

Moreover, In Ethiopia, as [1] studies the cardiovascular risk factors and complications of hypertension in type II diabetes are lacking. However a retrospective study in Tikur Anbesa Hospital showed that cardiovascular diseases (CVDs) were responsible for 16% of deaths among diabetic admissions second to acute complications and infections that caused 18% of deaths[20].

## 1.2 Statement of the Problem

As shown in the background information , although hypertension is a preventable and modifiable risk factor of CVD, the prevention and control of hypertension has not yet received due attention in many developing countries . These problems in Ethiopia are also common, so that this study were important to shed some light on the problem and factors of hypertension and diabetes.

Most of the diabetes mellitus studies in Ethiopia were institution-based and urban focused with no studies of rural dwellers [1]. Hence, to study hospital-based epidemiological information from both urban and rural populations at Debre Tabor referral hospital were essential to understand the picture of hypertension and diabetes mellitus in Ethiopia.

As [26] conducted, assessing the quality care provided to patients with in South west Ethiopia has shown that, there exists a huge gap in the quality of diabetes care. Because of this gap how hypertension and type II diabetes were manage in Ethiopia is not well-known. So that these studies were gives attention based on hospital retrospective study manage the factors of both hypertension and type-II diabetes.

Accordingly, the following research question were answered.

- What are the potential risk factors that affect hypertension and type II diabetes mellitus jointly?
- How the changes in association between hypertension and type II diabetes mellitus over time?
- How much visiting time are important to identify factors of hypertension and type II diabetes mellitus?
- Are there within and between significant variations of hypertension and type II diabetes mellitus?

## 1.3 Objective of the study

### 1.3.1 General objective

The general objectives of this study was to assess factors that affecting hypertension and type II diabetes patients jointly at Debre Tabor referral hospital by using multivariate multilevel model.

### 1.3.2 Specific objectives

- To examine the effects of various risk factors on hypertension and type II diabetes mellitus patients.
- To examine the association between hypertension and type II diabetes mellitus over a time.
- To explore the number of visiting time and its effect for hypertension and type II diabetes patients.
- To analyze within and between patients variation on hypertension and type II diabetes patients.

## 1.4 Significance of the Study

Significant advances of the study were makes in the diagnosis and management of hypertension in type II diabetes as contemporary studies have increased our understanding of the risk factors and the effective treatments.

The findings of the study were give an input for policy makers to concern health specialists who work on providing care, support and treatment aspect for programs of the country. And thus, the Clinical management were benefit for the rapid testing of new treatment strategies and evolving guidelines which are often updated to maximize positive outcomes for patients. Moreover, recommendations based on the findings would serve as tools for further research into the subject matter, to make sure that acceptable and lasting

solutions are found. Coincident with the study lifestyle changes which serve as both preventative measures and early treatment, and do options for treatment of hypertension and diabetes. Finally, to offers flawless information to researcher how to use multivariate multilevel models.

## 1.5 Limitations of the study

The study were used retrospective cohort data from the hospital; that have inherent gaps such as many socio-demographic and clinical variables that affect hypertension in type II diabetes as indicated by different studies in different countries were not recorded.

## 1.6 Organization of study

The study was organized into five chapters. Following the introductory chapter one, chapter two gives a literature review on the trends and review factors using different model. Chapter three discusses the methodology and data used in the study. Chapter four deals result and discussion of the model estimation and interpretation of results. Finally, chapter five was presents conclusions and recommendation of the study.

# CHAPTER TWO

# 2 REVIEW LITERATURE

## 2.1 Operational definition

According to WHO definition, the term cardiovascular disease (CVD) is an abnormal functioning of the heart or blood vessels in the body.

Diabetes mellitus(DM) is metabolic disorder of multiple etiologies characterized by chronic hyperglycemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action, or both.

Type I diabetes: defined as body's failure to produce insulin, and presently requires the person to inject insulin; whereas, Type II diabetes is defined as a an insulin resistance of body cells fail to use insulin properly.

Hypertension (HT) is usually defined by presence of a chronic elevation of blood pressure of an individual whose systolic pressure is greater than 140 mmHg and the diastolic pressure is greater than 90 mmHg.

Systolic blood pressure(SBP) is the highest level of the blood pressure during a heartbeat and diastolic blood pressure(DBP) is the lowest level of the blood pressure during a heartbeat.

## 2.2 Trends of hypertension and type-II diabetes mellitus

According to global estimates for the prevalence of type-II diabetes for 2015 and 2040, the burden of diabetes mellitus is rapidly increasing in worldwide. In 2015 it was estimated that there were 415 million people with diabetes aged 20–79 years, from this 5.0 million deaths attributable to diabetes. According to this study and prediction three quarters (75%) of those diabetes were living in low- and middle-income countries, and in 2040 diabetes on aged 20–79 years will rise to 642 million[36].

According to international surveys study, hypertension although uncertainty increase in the world in some estimates, the rate of elevated SBP (greater than 110-115 and greater

than 140 mm Hg) increased substantially between 1990 and 2015, and dallies and deaths associated with elevated SBP also increased. Projections based on this sample suggest that in 2015, an estimated 3.5 billion adults had SBP of at least 110 to 115 mm/hg and 874 million adults had SBP of 140 mm/hg or higher[18].

According to [10] study, type-II diabetes in global population-weighted is increased by 11.2% from 1990 (39.7 $\mu$g/m$^3$) to 2015 (44.2 $\mu$g/m$^3$), increasing most rapidly from 2010 to 2015. From this study considerations, among the world's ten most populous countries, in 2010 exposures increased in Bangladesh and India and were stable but remained high in Pakistan and China. And, exposures decreased substantially in Nigeria and were low and slightly decreased in the USA, Brazil, and Russia. Population-weighted concentrations were low and stable in Japan and Indonesia.

From Cape Town study by [37], hypertension is regarded as one of Sub-Saharan Africa's (SSA) greatest health challenges after HIV/AIDS; a far cry from the early 20th century when hypertension was rare in the region. Furthermore, unlike high-income countries where mean blood pressure (BP) has decreased over the last three decades, but in Africa it remained stable or increased in most countries with hypertension emerging.

According to[34] on seminar of type-II diabetes in sub-Saharan prevalence and burden of type II diabetes are rising quickly in Sub-Saharan Africa. On that seminar, some types of diabetes arise at younger ages in African than in European populations and the rate of undiagnosed diabetes is high in most countries of sub-Saharan Africa.

According to [16] findings, the prevalence of hypertension is increasing worldwide, in Ethiopia with an estimated 972 million adults with hypertension in 2000 that is predicted to grow to 1.56 billion by 2025, while diabetes worldwide prevalence is estimated as 382 million in 2012 projected to reach 592 billion in 2030.

According to [20] finding in Ethiopia, by considering five year longitudinal study of a total of 1486 diabetes mellitus and 1907 hypertension cases recorded from December 2010 to January 2014 was conducted to analyze the trend of diabetes mellitus and hypertension at Nigist Eleni Mohammed general Hospital, Hossana. According to the study, the annual average increase of the diabetes mellitus and hypertension was 5.4% and 8% respectively.

## 2.3 Factors of hypertension and type-II diabetes mellitus separately

According to [30] a recently worldwide published study using mixed model cohort study, the major factors of patients with type II diabetes in Western communities and Asian counterparts are older age, gender male, smoking, obesity, and Cholesterol.

According to [6] Indian Council of Medical Research–India Diabetes study, by applying Multivariate regression analysis of individuals aged grater than 20 years surveyed using a stratified multistage sampling design, showed that risk factors of hypertension in urban and rural India are age, male gender; urban residence, generalized obesity, increasing diabetes; physical inactivity and alcohol consumption were significantly associated with HTN.

According to [15] in North America and Europe, by applying multivariate analysis of mixed-effect regression with random intercepts model, the major risk factors type II diabetes are gender male, body mass index, younger age, adherence to lifestyle/medication,increase diabetes, increasing treatment regimen complexity and physician-reported patient's unwillingness to intensify treatment were associated with not achieving goal.

According to [27] in Cameroon by performing Univariate and multivariate analysis for each variable of Levene's test and subsequently Student t test were used to compare the distribution of hypertension in men and women. Pearson's chi-square test for independence was used to study the relationship of hypertension and region, older age, sex male, blood glucose level, and obesity are same common risk factors.

According to [25] drug therapy management of patients in Ethiopia, by appling a systematic meta-analysis weight management is key components to reduce control blood pressure. An up-to-date and comprehensive assessment of the evidence concerning hypertension in Ethiopia is lacking. One community-based cross sectional study done in Addis Ababa showed that the age adjust prevalence of high blood pressure was 31.5 % among males and 28.9 % among females.

According to [3] by using descriptive statistics of pie chart, 28.1% of patients presented

with blood pressure in Stage two category and 31.6% were in the stage one range and only 22% presented with a blood pressure in the pre- hypertension range were affects hypertension.

As [47]study of Bivariate analysis, the association of each independent variable and type II diabetes mellitus were analyzed by using binary logistic regression model. The overall prevalence of type 2 diabetes mellitus was found Being female,no formal education,current use of alcohol, T1DM, greater than five years duration of diabetes mellitus illness, chronic complication of diabetes mellitus and other additional chronic illness were significantly associated factors depression among patients with type 2 diabetes mellitus.

According to [46] study by using univarite and multivariate analysis, there are some socio demographic, clinical, and psychosocial factors that affect type 2 diabetes mellitus, among those the clinical variables body mass index 25.0–29.9 kg/$m^2$ and, the presence of grater than three co-morbid disease, among psychosocial risk poor social support, and among socio demographic only monthly family income are statistically significant risk factors of type II diabetes after controlling of other confounding factors at Black Lion General Specialized Hospital, Addis Ababa, Ethiopia.

According to [24] study, by performing multiple logistic regression analysis with demographic and clinical characteristics conducted among independent variable age, ethnicity,anti-diabetics, medication adherence are factors for diabetes in Ethiopia.

According to [19] study, by applying multivariate and logistic regression analyses to see the association between dependent and independent variables. There were identified some risk factors of hypertension which are, age, gender male, family history, diabetes mellitus, increasing BMI, drinking coffee, and chat use in southwest Ethiopia.

According to [1] study, by applying the logistic regression analysis on cross sectional data separately for urban and rural participants revealed associated risk factors of diabetes mellitus are family history, older age, Alcohol consumption, sex male, and physical inactivity in a rural population of northwest Ethiopia.

According to above researcher [2] in another paper by applying the multivariable logistic regression analysis of population-based cross-sectional study showed that factors included

obesity, old age, alcohol consumption, and increasing waist circumference are associated with Isolated Diastolic Hypertension (IDH) in a rural population of northwest Ethiopia. According to [5] study Mizan-Aman Town, Southwest Ethiopia, on the prevalence of diabetes mellitus and its risk factors among individuals aged 15 years and above, by applying logistic regression analyses the prevention strategy to such modifiable risk factors might reduce the prevalence of diabetes mellitus and screening of DM particularly in those individuals having high WC, history of smoking habit, and hypertension needs attention. According to [53] study, at Felege Hiwot Referral Hospital, Bahir Dar, Ethiopia by applying joint mixed effect modeling of longitudinal bivariate responses with unstructured covariance structure among all covariates included in joint-mixed effect-models, sex, residence, related disease and time were statistically significant on evolution of systolic and diastolic blood pressure.

## 2.4 Factors of hypertension and type-II diabetes mellitus jointly

According to American Diabetes Association [13] study, by applying meta-analyses of stratified clinical trials, hypertension were common among patients with diabetes, with the prevalence depending on type and duration of diabetes and hypertension. Thus, age, sex, race/ethnicity, BMI, history of glycemic control, and the presence of kidney disease, are common factors of hypertension in type II diabetes in the world.

According to [16] by applying multiple drug anti-hypertensive therapy experimental study at Hai Aljamea Hospital, internal medicine study the prevalence of coexisting hypertension and diabetes appears to be increasing in industrialized nations because populations are aging and both hypertension and type2 diabetes mellitus incidence increases with age. A number of possible reasons have been adduced for this coexistence and it is postulated that both diseases share common pathogenic factors such as insulin resistance, aging, obesity, chronic sub-clinical inflammatory processes beside the use of thiazide diuretics in subjects initially with hypertension and the development of nephropathy in those initially with diabetes, especially type 1.

According to[41] reviews in Africa country, by applying fixed-effect and random-effects

method of meta-analysis, the combination of both diseases can be dangerous for human life and together they can increase the risk of heart attack or they creates blood stroke in developing country. The factors of hypertension and diabetes are including overweight (obesity), mental stress, lack of exercise and an inactive lifestyle.

According to [22] study, by applying multinomial logistic regression analysis used to identify the predictors of hypertension and DM among the participants. Predictors of hypertension were age grater than 40 years, overweight/obesity, and sex while females are less likely to develop hypertension. And also, the significant predictor of diabetes mellitus was overweight/obesity in Nigeria.

According to [28] study, hypertension in type II diabetes mellitus by using repeated measures of cognitive and physical functions were regressed, separately, on time, baseline BP, DM, and control variables. As subjects were followed from first visit to second visit examinations, it was found that people with Stage 1 and Stage 2 hypertension had a selectively faster pace of decline in reasoning performance than normotensive subjects. And also baseline DM demonstrated a faster decline on the subjects ( $\beta$= -0.97, P = 0.02) on changes of cognitive performance over 2 years.

According to [52], After multi-variable analysis using the Gompertz Cox-Regression: covariates like sex, hypertension status at baseline, protein urea at baseline, HDL-C level, LDL-C, triglyceride level were found to be independent predictors for vascular complications among hypertension in type II DM patients. The risk of developing vascular complications is decreased by 50% among male type II DM patients than female patients. The risk of vascular complications for patients who have hypertension at baseline (first or second stage) was 3.99 times higher than that of patients who have pre-stage hypertension.

According to [35] study in Ethiopia, by applying Chi-square tests used to evaluate differences in means for continuous variables, expressed as mean with standard errors of means (SEM), and categorical variables expressed as number (%). Substantially the prevalence diabetes is higher among women aged 45–54 years as compared with their similarly aged male counterparts. Likewise, the prevalence of hypertension is higher among women aged

45–54 years as compared with their similarly aged male.

According to [20] study in Ethiopia, by applying trained analysis the magnitude of hypertension and diabetes mellitus is higher in males than in females. This study also revealed that the observed trend of each disease is statistically significant with the expected trend of each disease across the year.

## 2.5 Model review based on literature

Most of the studies shown from this literature [1, 5, 19, 22, 24, 35] apply logistic regression analysis using cross-sectional data, even though this analysis show the association, but it doesn't show association over a time, and within and between variation is not clearly identified. On the other hand;[6, 27, 32] apply multivariate regression analysis using cross-sectional data, this analysis also does not show association over a time, and within and between variation.[15, 30] apply linear mixed model using cohort study separately, but dosn't apply jointly for hypertension and type-II diabetes. Lastly[25, 41] applies meta analysis, the limitations of this analysis is in their reliance on the measures at only one or two time points, which involved a much loss of data, and do not show the trend over time.

Therefor, by considering of their limitation and gaps, this study were look at the risk factors of hypertension and type II diabetes jointly, by using multivariate multilevel model association at a time and over a time.

# CHAPTER THREE

# 3 METHODOLOGY

## 3.1 Study area and study design

The study was carried out retrospectively cohort study design from Debre Tabor Referral Hospital. Debre Tabor referral hospital is a teaching and referral hospital in Amhara region, Ethiopia that serving people at and around Debre Tabor town.
According to the report of the chief executive officer of Debre Tabor referral hospital, this hospital is serving about 5 million people per year that encompasses about 300 health care professionals in addition of administrative staffs.

## 3.2 Source population and periods

The source population in the study reviews charts of all patients taking in Debre Tabor Referral hospital from $1^{st}$ January 2017 to $1^{st}$ January 2019 E.C.

## 3.3 Study population and/or study unit

Regular follow-up patients (whose age's grater than 18), that have hypertension in type II diabetes were our target population.

## 3.4 Inclusion and exclusion criteria

### 3.4.1 Inclusion criteria

Hypertension in type II diabetes patients had visited at least three times and at most six times visit in every three months irrespective of their visit during $1^{st}$ January 2017 to $1^{st}$ January 2019 E.C recorded in the hospital were include in the study.

### 3.4.2 Exclusion criteria

Records available before $1^{st}$ January 2015 and records which are vague and incomplete were not included in the study. Even though, patients that have diabetes during pregnancy and patients have type I diabetes were not included in the study; because the cases are rare, not only females.

## 3.5 Sample-size determination

The importance of sample-size calculations is a rather strange phenomenon. Firstly, sample-size calculations are based on many assumptions, which can easily be changed, and in which case the number of subjects needed, will be totally different. Secondly, sample-size calculations are (usually) based on statistical significance, which is strange, because in epidemiological and medical research the importance of significance levels is becoming more and more questionable[7]. Sample-size calculation designed for continuous outcome variable of clinical study is calculated using the following formula adopted for groups $\alpha$=0.05 and $\beta$=0.1

$$n = \frac{(\sigma_1 + \sigma_2)^2 (Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})^2}{d^2(\mu_1 - \mu_2)^2} \qquad [3.1]$$

Where $\mu_1$=148.4 and S1=30.27 are the mean and standard deviation for SBP as well as $\mu_2$=100.12 and S2=16.09 are the mean and standard deviation for DBP calculated from pilot survey. Thus, by d =0.07 margin of error the sample size in the study was n =144.Finally 5 percent of the sample will be added to determine sample size to compensate for none response rate and the total sample size become 151 (=144+7).

## 3.6 Study variables

### 3.6.1 Dependent variables

In this study the 3-variate multivariate responses (SBP, DBP and T2DM) were considered longitudinally after fasting. Here SBP and DBP were measured using mmHg whereas T2DM was measured using mmol/L

### 3.6.2 Independent variables

The explanatory variable associated with SBP, DBP and T2DM was the following.
**Socio-demographic variables:** by considering each individual as a level II unit, age, gender, residence, alcohol, smoking exposure, chat chewing, coffee drinking, race/ethnicity, anti-diabetics, adherence to lifestyle or medication, prior gestational diabetes during pregnancy, psychological stress, physical activity, and an inactive lifestyle are common factors.
**Clinical variables:** BMI, history of glycemic control, Cholesterol level, presence of related disease, prior history of HT, prior history of T2DM, and baseline HT and T2DM are common factors.

## 3.7 Methods data collection procedures

This study were used secondary data extracted from patient chart follow-up format as a data extraction tool. The chart were prepared by national health organization be uniformly used by hospitals to identify and document variables as early as possible. The data was entering by SPSS version 21.0, and analyzed by R version 3.4.4 and SAS version 9.4 statistical software.

## 3.8 Missing data treatment

One of the biggest problems in longitudinal studies is missing data. There is an enormous amount of literature dealing with this problem to the possible imputation of missing data to obtain a 'complete' dataset . However, when applying multilevel analysis to

longitudinal data, there is no need to have a 'complete' dataset, and, furthermore, it has been shown that multilevel analysis is very flexible in handling missing data. It has even been shown that applying multilevel analysis to an incomplete dataset is even better than applying imputation methods[7]

## 3.9 Ethical aspects

Letter of ethical clearance was obtained from Bahir Dar University, department of statistics and submitted to the Debre Tabor referral hospital to get permission to conduct the research. The study was developed in accordance with established legislation and complies with the norms of good clinical practice, and informed consent was being not necessary as personal identifying information was kept separate from the research data. Finally, the study protocol was approved by the ethics committee or medical directors of the Debre Tabor referral hospital.

## 3.10 Statistical analysis technique for longitudinal data

Analyzing multilevel longitudinal data is complicated, to assess the changes of outcome(s) over time to associated risk factors. For analyzing of this longitudinal data jointly, both descriptive and inferential statistics were used. Generalized estimating equations (GEE), and general linear mixed model (GLMX) are the best techniques of analyzing inferential statistics for longitudinal data. Even though, all GEE, GLMX, and MLM used for analysis of longitudinal data, multilevel analysis is (probably) the most robust and flexible of the three techniques[7]. Thus, in this study multilevel longitudinal methods were employed. But, before doing this analysis it is better to do multivariate analysis for looking the association of two response.

### 3.10.1 Multivariate Analysis

Multivariate analysis is used to analyses two or more interrelated response simultaneously. This analysis carrying out a series of univariate test to inflate type I error rate and better

power. The multivariate analysis in longitudinal data also show the association between outcomes evolve over time[21].This is also sometime known as evolutions of association (EOA). Therefore, the G covariance structures of intercept and slop for response DBP, T2DM, SBP represented by $\alpha$ ,$\beta$ and $\gamma$ respectively were:

$$G = \begin{bmatrix} \delta^2_{\alpha_1} & \delta_{\alpha_1\beta_1} & \delta_{\alpha_1\gamma_1} & \delta_{\alpha_1\alpha_2} & \delta_{\alpha_1\beta_2} & \delta_{\alpha_1\gamma_2} \\ & \delta^2_{\beta_1} & \delta_{\beta_1\gamma_1} & \delta_{\beta_1\alpha_2} & \delta_{\beta_1\beta_2} & \delta_{\beta_1\gamma_2} \\ & & \delta^2_{\gamma_1} & \delta_{\gamma_1\alpha_2} & \delta_{\gamma_1\beta_2} & \delta_{\gamma_1\gamma_2} \\ & & & \delta^2_{\alpha_2} & \delta_{\alpha_2\beta_2} & \delta_{\alpha_2\gamma_2} \\ & & & & \delta^2_{\beta_2} & \delta_{\beta_2\gamma_2} \\ & & & & & \delta^2_{\gamma_2} \end{bmatrix}$$

Now, the marginal correlation between the three responses as a function of time is given by: [3.2]

$$r_m(DBP, T2DM)/t = \frac{\delta_{\alpha_1\alpha_2} + t(\delta_{\alpha_1\beta_1} + \delta_{\beta_1\alpha_2}) + t^2(\delta_{\beta_1\beta_2})}{\sqrt{\delta^2_{\alpha_1} + 2t^2(\delta_{\alpha_1\beta_1} + \delta^2_{\beta_1}) + \delta^2_1} * \sqrt{\delta^2_{\alpha_2} + 2t^2(\delta_{\alpha_2\beta_2} + \delta^2_{\beta_2}) + \delta^2_2}}$$

$$r_m(DBP, SBP)/t = \frac{\delta_{\alpha_1\alpha_2} + t(\delta_{\alpha_1\gamma_1} + \delta_{\gamma_1\alpha_2}) + t^2(\delta_{\gamma_1\gamma_2})}{\sqrt{\delta^2_{\alpha_1} + 2t^2(\delta_{\alpha_1\gamma_1} + \delta^2_{\gamma_1}) + \delta^2_1} * \sqrt{\delta^2_{\alpha_2} + 2t^2(\delta_{\alpha_2\gamma_2} + \delta^2_{\gamma_2}) + \delta^2_3}}$$

$$r_m(T2DM, SBP)/t = \frac{\delta_{\beta_1\beta_2} + t(\delta_{\beta_1\gamma_1} + \delta_{\gamma_1\beta_2}) + t^2(\delta_{\gamma_1\gamma_2})}{\sqrt{\delta^2_{\beta_1} + 2t^2(\delta_{\beta_1\gamma_1} + \delta^2_{\gamma_1}) + \delta^2_2} * \sqrt{\delta^2_{\beta_2} + 2t^2(\delta_{\beta_2\gamma_2} + \delta^2_{\gamma_2}) + \delta^2_3}}$$

When (many) more than two longitudinally measured outcomes need to be analyzed jointly, most approaches described earlier are no longer feasible, involve numerical difficulties, or are based on extremely strong, often unrealistic, assumptions about the association structure between the various outcomes in the multivariate response vector[51]. Probably the most frequently applied multivariate technique is the multivariate analysis of variance (MANOVA), in which the average values of two or more continuous outcome variable are compared between groups. When a significant difference is found between groups, the next step is to examine which of the outcome variables differs between the

groups, or, in other words, which of the outcome variables is related to the (group) determinant. But, the specification of the random coefficient model were specific assumptions for the covariance matrix of the p repeated measurements, and the unrestricted MANOVA model as special cases[33].

Therefore, after looking the evolution of association, thus multivariate statistical methods are needed to fully answer the one goals of our research; and how much repeated measure are used in such analysis is determined by growth curve analysis.

### 3.10.2 Growth curve analysis

Growth curve analysis is offers a statistical framework for analyzing longitudinal or time course data. The primary goal of growth curve analyses in particular is to describe patterns of change over time and to determine the number of visit time (i.e., variable reduction). More specifically, the core idea of growth curve analysis in longitudinal data can estimate a best-fit line or curve to each individual's responses over time (i.e., visit time).

There are few strict requirements for the types of data that might be analyzed using growth model. First, adequate at least 100 sample size is needed to reliably estimate growth models are often preferred. Second, growth models typically require at least three repeated measures per individual. Third, for the typical method of estimation called maximum likelihood (ML), it is assumed that the repeated measures are continuous and normally distributed[12] When discrete time points are used in a longitudinal study, time can also be modeled as a categorical variable.

#### 3.10.2.1 Univariate growth curve analysis

This univariate approach is clearly the least desirable method for estimating individual growth curves. Theoretically, it does not allow for individual differences in rates of change over time and, accordingly, were result in poorly specified growth curve models and associated test statistics[12]. According [9],separate growth curves are constrained to have the same slope but allowed to have different intercepts over time. As a consequence,

this method underestimates variability and overestimates test statistics when individuals have different slopes. This limitation has been widely recognized and various approaches have been developed to provide more appropriate test statistics. The mechanisms were use multivariate approach for the improvement over the univariate growth curve methods because individual differences in slopes were allowed [45].

#### 3.10.2.2 Multivariate growth curve analysis

This multivariate approach was a marked improvement over the univariate growth curve methods because individual differences in slopes were allowed. That is, separate growth curves are estimated for each individual, and those individual growth curve parameters are used to estimate group growth curves. When discrete time points are used in a longitudinal study, time can also be modeled as a categorical variable. In this study six measurements are made of each patient, so basically a fifth-order polynomial is the highest order growth curve that can be modeled with baseline measurement usually as reference point[7].

The multivariate regression coefficients belonging to each of these dummy variables indicate the difference between a certain times.

$$d_{ti} = \begin{cases} 1\ if\ t = 1,\ otherwis\ 0 \\ 1\ if\ t = 2,\ otherwis\ 0 \\ 1\ if\ t = 3,\ otherwis\ 0 \\ 1\ if\ t = 4,\ othewis\ 0 \\ 1\ if\ t = 5,\ otherwis\ 0 \\ 1\ if\ t = 6,\ otherwis\ 0 \end{cases} \qquad [3.3]$$

$$Y_{ti} = \pi_{1i}D_{1i} + \pi_{2i}D_{2i} + \pi_{3i}D_{3i} + \pi_{4i}D_{4i} + \pi_{5i}D_{5i} + \pi_{6i}D_{6i}$$

$$\left\{\begin{array}{l}\pi_{1i}=\beta_{1i}+u_{1i}\\ \pi_{2i}=\beta_{2i}+u_{2i}\\ \pi_{3i}=\beta_{3i}+u_{3i}\\ \pi_{4i}=\beta_{4i}+u_{4i}\\ \pi_{5i}=\beta_{5i}+u_{5i}\\ \pi_{6i}=\beta_{6i}+u_{6i}\end{array}\right\} \text{where } u_{ti}=\left\{\begin{array}{llllll}0 & \sigma^2_{u1} & & & & \\ 0 & \sigma_{u1u2} & \sigma^2_{u2} & & & \\ 0 & \sigma_{u1u3} & \sigma_{u2u3} & \sigma^2_{u3} & & \\ 0 & \sigma_{u1u4} & \sigma_{u2u4} & \sigma_{u3u4} & \sigma^2_{u4} & \\ 0 & \sigma_{u1u5} & \sigma_{u2u5} & \sigma_{u3u5} & \sigma_{u4u5} & \sigma^2_{u5} \\ 0 & \sigma_{u1u6} & \sigma_{u2u6} & \sigma_{u3u6} & \sigma_{u4u6} & \sigma_{u5u6} \end{array}\right. \left.\begin{array}{l} \\ \\ \\ \\ \\ \sigma^2_{u6}\end{array}\right\}$$

Where, the dummy has a fixed effect $\beta_1$ -$\beta_6$ and a random effect $(U_1 - U_6)$. $Y^k_{ti}$ is the outcome at time t of the $i^{th}$ individual[14].

Hence; after looking dummy occasions by considering baseline visit as a reference the growth curve for this multivariate (i.e., k) response is modeled as follow.

$$\left[Y^k_{ti}=\beta_{0i}+\beta_{1i}T_{1i}+\beta_{2i}T^2_{2i}+\beta_{3i}T^3_{3i}+\beta_{4i}T^4_{4i}+\beta_{5i}T^5_{5i}\right] \qquad [3.4]$$

The average intercept (i.e.,$\beta_{0i}$) is intercept that captured the main effect of the $k^{th}$ dummy variable; similarly, (i.e., $\beta_{1i}$) is the average linear growth parameter; ($\beta_{2i}$) the average quadratic growth parameter; and so on.

### 3.10.3 Multilevel analysis

In epidemiological and medical longitudinal studies, multilevel analysis is probably most often applied construction of growth curves. Over the past 20 years multilevel modeling has become a standard approach in the analysis of clustered data [44]. Longitudinal data are one example of a hieratical structure; series of repeated measures over time at the lowest level is nested with the individual persons at the highest level. Such nested structures are typically strong hierarchies because there is much more variation between individuals in general than between occasions within individuals. These repeated measures are taken both at fixed and varying occasion. The measurements taken as a fixed occasion, all individual provided measurements at the same set of occasions, usually regularly spaced, such as in our study every four months. When occasions are varying, we have different set of measures take at different points in time for different individuals[21]. The analysis

were accommodate in two different approaches of modeling that are univariate modeling approach and the multivariate modeling approach.

#### 3.10.3.1 Univariate multilevel analysis

This analysis is used for exploring an individuals variability on longitudinal data for each response; in such random effects or multilevel modeling allows investigation of two levels variability. Therefore, within and between subjects variability were separately analyzed for SBP, DBP, and T2DM at each individual i= 1, 2. . . n[23]. These models were analyzed based on either of following mechanisms.

##### 3.10.3.1.1 Intercept only model

This is the simplest case of hierarchical model in which there are no explanatory variables at all. Then model has only an intercept term and variances at the measurement and individual level. Since the model does not contain a slope, the true individual change is a horizontal line with y-intercept $\beta_{0i}$.

The model expressed as:

$$\text{Level-1: } Y_{ti} = \pi_{0i} + e_{0i} \text{ where, } e_{ti} = N(0, \sigma_e)$$

$$\text{Level-2: } \pi_{0i} = \beta_{0i} + u_{0i} \text{ where, } u_{ti} = N(0, \sigma_u)$$

Where, the Greek letters $\pi$ and $\beta$ indicate first and second level parameter respectively. By substituting, we get

$$Y_{ti} = \beta_{00} + u_{0i} + e_{0i} \qquad [3.5]$$

$\pi_{0i}$ is the intercept for individual i for each response; $\beta_{00}$ is the mean intercept over all individuals, and $u_{0i}$ are the deviation of each individual's means; finally, $e_{0i}$ is the time- and individual-specific residual.

Now, proportion variance or the intra class correlation (ICC) refers to a set of coefficients representing the relationship between variables of the same individuals decomposes

into two independent components (i.e., level-1 and leve-2). Thus, ICC explained by the individuals (level-2) in the population is given by

$$ICC_{SBP} = \frac{{\sigma_{u_0}}^2}{{\sigma_{e_0}}^2 + {\sigma_{u_0}}^2}$$

$$ICC_{DBP} = \frac{{\sigma_{u_0}}^2}{{\sigma_{e_0}}^2 + {\sigma_{u_0}}^2}$$

$$ICC_{T2DM} = \frac{{\sigma_{u_0}}^2}{{\sigma_{e_0}}^2 + {\sigma_{u_0}}^2}$$

Where, $e_0$ and $u_0$ for each response are different and

$ICC_{measures}$=1- $ICC_{individual}$ [3.6]

#### 3.10.3.1.2 Random intercept model

A random intercepts model is a model in which intercepts are allowed to vary, and therefore, the scores on the dependent variable for each repeated measurement are predicted by the intercept that varies across patients. The prior models are sometimes called unconditional (intercept only)because there are no measured covariates used to predict the random effect.

Now, in this model often be interested in assessing how a longitudinal outcome is associated with a covariate whose value changes over time such covariant are called time-varying covariates$X_{pi}$, and whose value not changes over time called time invariant predictor $Z_{si}$.

The model is given by: Level 1: $Y_{ti} = \pi_{0i} + \pi_{1i} X_{pi} + e_{0i}$

Time invariant covariates Z inter the equation at the second level.

$$\text{Level 2: } \pi_{0i} = \beta_{00} + \beta_{01} Z_{si} + u_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} Z_{si}$$

By substituting, we get a univariate random intercept model:

$$Y_{ti} = \beta_{00} + \beta_{10} X_{pi} + \beta_{01} Z_{si} + \beta_{11} Z_{si} X_{pi} + u_{0i} + e_{0i} \quad [3.7]$$

Where, $\beta_{00}$ is the overall average intercept for each response, $\beta_{10}$ is the slop of time varying covariates, $\beta_{01}$ is the slop of time invariant covariates, $\beta_{11}$ is the mean difference

change between time varying covariates and time invariant covariates, and lastly $e_{0i}$ and $u_{0i}$ are still the within and between individual error term of the intercept.
Therefore, in this model $\beta_{00}+\beta_{10}X_{pi}+\beta_{01}Z_{si}+\beta_{11}Z_{si}X_{pi}$ is the fixed part, because the coefficients are fixed. The remaining $u_{0i}$ $+e_{0i}$ is called the random part of the model.

where, $X_{pi}$, p=1, 2,..., P, denotes the P time varying covariates that were included in the analysis, and $Z_{si}$ , s=1, 2... S denotes the s invariant covariate that were included in the analysis.

#### 3.10.3.1.3 Random coefficients model

This random coefficients model is a model in which slopes are allowed to vary in addition to intercepts for each univariate response. The relationship between an explanatory variable and the response is different across all patients with their intercept and slope by considering time varying covariate and time invariant covariate.
The model is given by: Level 1: $Y_{ti} =\pi_{0i} +\pi_{1i} X_{pi}+e_{0i}$
Time invariant covariates Z inter the equation at the second level.

$$\text{Level 2: } \pi_{0i} = \beta_{00} + \beta_{01}Z_{si} + u_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}Z_{si} + u_{1i}$$

By substituting, we get univariate random coefficient model

$$Y_{ti}=\beta_{00}+\beta_{10}X_{pi}+\beta_{01}Z_{si}+\beta_{11}Z_{si}X_{pi}+u_{1i}X_{pi} + u_{0i} \ +e_{0i} \qquad [3.8]$$

Where, the only additional term in this model is $u_{1i}$ the random slop of time varying covariates for each response,
In above random coefficient model explained variance ($R^2$) is test improvement of how much variation is explained by each level, by computing difference of the deviance baseline model (b) and the current (m) model.

$$R^2_{individual}=\frac{\delta^2_{u0/b} \ -\delta^2_{u0/m}}{\delta^2_{u0/b}} \text{ and } R^2_{measurmet}=\frac{\delta^2_{e0/b} \ -\delta^2_{e0/m}}{\delta^2_{e0/b}} \qquad [3.9]$$

#### 3.10.3.2 Multivariate multilevel analysis

Multivariate multilevel analysis is a methodology for the analysis of two or more response variables of data manifesting complex patterns of variability, with a focus on nested sources of variability. Multivariate multilevel analysis is a powerful analysis tool, like to analysis of variance using several response variables may lead to more powerful than universities. There are also numbers of advantages to this model expression, a key one of which is the inclusion of more complex error structures at both level 1 and level 2 than is possible within the univariate model. The reason is that the covariance structure not only holds within each construct separately, but it also holds across construct. In MVML analysis there are two somewhat-odd things about this expression relative to the usual univariate model. First, there is no overall intercept term for this reduced-form model. Instead, the intercept for the first outcome is captured in the main effect of the first dummy variable (i.e., $\beta_{1i}^{1} d_{ti}^{1}$) as shown below; similarly, the intercept for the second and third outcome is captured in the main effect of the second and third dummy variable (i.e., $\beta_{1i}^{2} d_{ti}^{2}$ and $\beta_{1i}^{3} d_{ti}^{3}$ respectively). Second, the linear slope for each outcome is captured in the interaction between each dummy variable and (i.e., time varying and time invert) predictors[11]. Thus we have k dummy variable $d_{ti}^{k}$, define k= 1, 2, and 3 for SBP, DBP and T2DM respectively.

$$d_{ti}^{k} = \begin{cases} 1 \; if \; k = SBP, \; otherwis \; 0 \\ 1 \; if \; k = DBP, \; otherwis \; 0 \\ 1 \; if \; k = T2DM, \; otherwis \; 0 \end{cases} \qquad [3.10]$$

Now, like to above UML analysis these models were also analyzed based on either of following mechanisms by considering this dummy variable.

##### 3.10.3.2.1 Intercept only model

Intercept only model is useful as a null model that serve as a benchmark with which other models are compared, if there is variation by computing the interclass correlation (ICC). Since this model has no explanatory variables, the model is known as unconditional model or null model. To fit simultaneous variables of interest model without any explanatory

variables, the k response dummy variables in a model must exclude the usual intercept only model.

The model expressed as: Level-1: $Y_{ti} = \pi_{1i} d^1_{ti} + \pi_{2i} d^2_{ti} + \pi_{3i} d^3_{ti}$,

$$\text{Level-2: } \pi_{1i} = \beta^1_{1i} + U^1_{1i}$$

$$\pi_{2i} = \beta^2_{1i} + U^2_{1i}$$

$$\pi_{3i} = \beta^3_{1i} + U^3_{1i}$$

By substituting this model is rewritten as

$$Y_{ti} = \beta^1_{1i} d^1_{ti} + \beta^2_{1i} d^2_{ti} + \beta^3_{1i} d^3_{ti} + U^1_{1i} d^1_{ti} + U^2_{1i} d^2_{ti} + U^3_{1i} d^3_{ti} \qquad [3.11]$$

$$= \sum_{k=1}^{k=3} \beta^k_{1i} d^k_{ti} + \sum_{k=1}^{k=3} U^k_{1i} d^k_{ti}$$

Where, $U^k_{1i}$ has an individual random effect that captures level-2 variation for the $k^{th}$ response, and in the case there is no lowest level error term, because the lowest level exists solely to define the multivariate response structure.

In above [3.6], the fixed part has only the intercept, which is the overall mean, and the random part has within-between intercept variance. In the equivalent multivariate model the fixed part has k regression coefficients for the dummy variables, with k overall means for the $k^{th}$ response variables. The random part has covariance matrix G and $R$, which contains variance and covariance of regression slops of individual and intercept measures. In this matrix the upper and lower diagonal matrix are equal and that reflect the between patient level covariance for each two response; the main diagonal is the variance for each response.

Like to the above [3.7]; in multivariate analysis the inter-class correlation were show how much variation are explained by level-2 were analyzed.

$$\text{ICC}_{mv} = \frac{\Omega^2_{u0}}{\Omega^2_{u0} + \Omega^2_{e0}}, \qquad [3.12]$$

Where, $\text{ICC}_{measures} = 1 - \text{ICC}_{individua}$

#### 3.10.3.2.2 Random intercept model

Like to above [3.8] model; in multivariate random intercept model intercept are allowed to vary. Now looking the association of outcome by adding time-varying covariates $X_{pi}$ and by adding time invariant predictor Z, the models were used to assess the contribution, significance and variance changes of multivariate outcome.

The model expressed as: Level-1: $Y_{ti}^{k} = \pi_{1i} d_{ti}^{1} + \pi_{2i} d_{ti}^{2} + \pi_{3i} d_{ti}^{3}$,

$$\text{Level-2: } \pi_{1i} = \beta_{1i}^{1} + \beta_{2i}^{1} X_{pi} + \beta_{3i}^{1} Z_{si} + U_{1i}^{1}$$

$$\pi_{2i} = \beta_{1i}^{2} + \beta_{2i}^{2} X_{pi} + \beta_{3i}^{2} Z_{si} + U_{1i}^{2}$$

$$\pi_{3i} = \beta_{1i}^{3} + \beta_{2i}^{3} X_{pi} + \beta_{3i}^{3} Z_{si} + U_{1i}^{3}$$

By substituting the model is

$$Y_{ki} = \sum_{k=1}^{k=3} \beta_{1i}^{k} d_{ti}^{k} + \sum_{k=1}^{k=3} \beta_{2i}^{k} d_{ti}^{k} X_{pi} + \sum_{k=1}^{k=3} \beta_{3i}^{k} d_{ti}^{k} Z_{si} + \sum_{k=1}^{k=3} U_{1i}^{k} d_{ti}^{k} \qquad [3.13]$$

Where,$(u_{1i}^{k})'$ is MVN(0,G)

#### 3.10.3.2.3 Random coefficients model

The last step in the multilevel analysis is the addition of a random slope. This random coefficients model is a model in which intercept and slopes are allowed to vary; the relationship between explanatory variable and the response is different across all patients with their intercept and slope. If we fit a model based on the same predictors on the response variable for all patients jointly, we may obtain different intercept and slopes for each patient in multilevel analysis.

This model also considers time varying covariate $X_{pi}$ and time invariant covariate $Z_{si}$ like to the above random intercept model.

The model expressed as: Level-1: $Y_{ti}^{k} = \pi_{1i} d_{ti}^{1} + \pi_{2i} d_{ti}^{2} + \pi_{3i} d_{ti}^{3}$,

$$\text{Level-2: } \pi_{1i} = \beta_{1i}^{1} + \beta_{2i}^{1} X_{pi} + \beta_{3i}^{1} Z_{si} + U_{2i}^{1} X_{pi} + U_{1i}^{1}$$

$$\pi_{2i} = \beta_{1i}^{2} + \beta_{2i}^{2} X_{pi} + \beta_{3i}^{2} Z_{si} + U_{2i}^{2} X_{pi} + U_{1i}^{2}$$

$$\pi_{3i} = \beta_{1i}^{3} + \beta_{2i}^{3} X_{pi} + \beta_{3i}^{3} Z_{si} + U_{2i}^{3} X_{pi} + U_{1i}^{3}$$

By substituting the model is

$$Y_{ki}=\sum_{k=1}^{k=3}\beta_{1i}^{k}d_{ti}^{k}+\sum_{k=1}^{k=3}\beta_{2i}^{k}d_{ti}^{k}X_{pi}+\sum_{k=1}^{k=3}\beta_{3i}^{k}d_{ti}^{k}Z_{si}+\sum_{k=1}^{k=3}U_{2i}^{k}d_{ti}^{k}X_{pi}+\sum_{k=1}^{k=3}U_{1i}^{k}d_{ti}^{k} \qquad [3.14]$$

Where,$(u_{1i}^{k},u_{2i}^{k})'$ is MVN(0,G)

After all like to above univariate multilevel analysis; for multivariate multilevel analysis explained variance ($R^2$) is tested by computing difference of the deviance baseline model (b) and the current (m) model.

$$R_{individual}^{2}=\frac{\delta_{u0/b}^{2}-\delta_{u0/m}^{2}}{\delta_{u0/b}^{2}} \text{ and } R_{measurmet}^{2}=\frac{\delta_{e0/b}^{2}-\delta_{e0/m}^{2}}{\delta_{e0/b}^{2}} \qquad [3.15]$$

### 3.10.4 General approach of multivariate multilevel analysis

It has been mentioned before that there is also an 'alternative', more general way to perform a multivariate multilevel analysis. It must be realized that to perform this more general multivariate multilevel analysis, the outcome variables must be scaled in the same way. One of the possibilities is to calculate standardized values (i.e. z-scores) of the (three) continuous outcome variables under consideration. This means that from each total systolic blood pressure observation, the average total systolic blood pressure value has to be subtracted, and this value must be divided by the standard deviation of the total systolic blood pressure values in the population under study. The same has to be done for the diastolic blood pressure and type II diabetes mellitus observations, with the average and standard deviation of the diastolic blood pressure and type-II diabetes values. It should be noted that the data structure that is needed to perform the more general approach for multivariate multilevel analyses.

### 3.10.5 Selection of covariance structure

The most common covariance structure in repeated measure are: First, Simple structure species that the observations are independent, even on the same patient, and have homogeneous variance. Second, Compound symmetric structure species that observations on the same patient have homogeneous covariance and homogeneous covariance. Third,

unstructured structure species no patterns in the covariance matrix, and is more appropriate for balanced data nature. fourth, Autoregressive (order 1) covariance structure species homogeneous variance; and more appropriate for unbalanced data and equally spaced measurement times such that $t_{n+1}$-$t_n$ is a constant for all n [31]. In addition to this, AR (1) model use only two parameters yield AIC and SBC statistics are considerable superior than totally unstructured model even though -2RLL is larger or worse[42].

### 3.10.6 Variable selection for multilevel analysis

In order to select variables to be included in multi variable-analysis, forward variable selection was used. The first step in this selection is to fit a univariable multilevel model for each covariate at the 25% level. Next multi-variable model is fitted that contains all covariates that are significant in univariable analysis for each response separately; Lastly based on this selected variable multi-variable model is fitted for multivariate response jointly (SBP, DBP and T2DM).

### 3.10.7 Model selection and comparison

In order to select the best and final model which is appropriately fits with the given longitudinal data, it is necessary to compare the different models by using different techniques and methods. Hence, Akaki information criteria (AIC) and Bayesian information criteria (BIC) that calculated from deviance based on number of estimated parameter p is also most convenient at 5% level of significance, but for multilevel model deviance information criteria (DIC) is appropriate.

$$DIC = -2log\frac{\lambda_0}{\lambda_1} \qquad [3.16]$$

Where, -2log is the twice negative log-likelihood value for the model $\lambda_0$ and $\lambda_1$ for the null and alternative hypothesis respectively. After all, Deviance compares chi squared distribution with degrees of freedom equal to the difference (p) in the number of parameters fitted under the two models smaller values is better[21]

### 3.10.8 Parameter estimation

#### 3.10.8.1 Maximum likelihood estimation

Parameter estimates of multilevel model were derived for both fixed components and random components. To estimate this component, there is different of parameter estimation technique. Among that, maximum likelihood estimation is most commonly used estimation method in multilevel model. Because an advantage of ML estimation method is that generally robust, and produce estimate that are asymptotically efficient (the lower-variance of two estimators) and consistent (estimate $\beta\hat{}$ converge to the true parameter value $\beta$). Maximum likelihood estimate (MLE) for parameters $\beta$ be estimated and defined that the likelihood of the model parameters, given the vector of $n$ observations, defined as:

$$L = l\left(\beta, Y_{ti}^{k}\right) = \prod \left\{ 2\pi^{\frac{-n_k}{2}} \det(V)^{\frac{-1}{2}} \exp(\frac{-1}{2}(Y_{ti}^{k} - Y_{ti}^{k}\beta)^{t} V^{-1}(Y_{ti}^{k} - Y_{ti}^{k}\beta) \right\}$$

Then, the MLE of $\widehat{\beta}$ on combining all information from all $n$ subjects equals

$$\widehat{\beta}=(\sum x_i V^{-1} x_i)^{-1}(\sum x_i V^{-1} Y_{ti}^{k})^{-1} \qquad [3.17]$$

Where det refers to the determinant, and the elements $V = var(Y_{ti}^{k})$ matrix are functions of the covariance parameters in $\beta$.

Whereas in a multilevel model, we shall generally have several residuals at different levels and for the variance components model we obtain:

$$\hat{u}_i = \frac{n_k {\delta_u}^2}{(n_k {\delta_u}^2 + {\delta_{e0}}^2)} \tilde{y}_k \;\; and \; e_{0i} = \tilde{y}_j - \hat{u}_i \left( i.e., \; \tilde{y}_k = \frac{\sum \tilde{y}_{ti}}{n_k} \right) \qquad [3.18]$$

Where, $n_k = n$ is the number of level 1 unit in the $i^{th}$ level 2 units of $k^{it}$response

### 3.10.9 Goodness of fit test

Once a model has been developed through different techniques in estimating the model parameters, there was several mechanisms involved in assessing the appropriateness, adequacy and usefulness of the model.

#### 3.10.9.1 Test for individual predictors

The t-test statistic is commonly used to test significance of individual parameter regression coefficients for each independent variable. Let denote an arbitrary parameter. The hypothesis is given as follows: $H_0$: $\beta_i = 0\, Versus\ H_\alpha$: $\beta_i \neq 0$ The t-test statistic is given by: $Z_i = \frac{\hat{\beta}_i}{S.E(\hat{\beta}_i)}$ [3.19] Where, $\beta_i$, i =1,2,3,....., p is coefficient of the variable. Under the null hypothesis, $Z_i$ has approximately a t-distribution with the number of degree of freedom $(d.f.) = n-r-1$ where $n$ is the total number of sample and $r$ is total number of explanatory variables. If $n - r$ is large enough say larger than 40, the t-distribution be replaced by a standard normal distribution.

#### 3.10.9.2 Test for overall predictors

In applications of a hierarchical linear model, the deviance-based test, or likelihood ratio test, is a general principle for testing fixed multi-parameter and for testing about the random part of the model. When parameters of a statistical model are estimated by maximum likelihood (ML) method, the estimation also provides the likelihood,

$$deviance = -\log\left(\frac{\lambda_0}{\lambda_1}\right) \qquad [3.20]$$

Where, $\lambda_0\ and\ \lambda_1$ are the likelihoods for the null and alternative hypotheses and this is referred totables of the chi squared distribution with degrees of freedom equal to the difference (p) inthe number of parameters fitted under the two models.

### 3.10.10 Model diagnosis

For multilevel analysis making inference about the model depends on whether the data met the required assumptions or not. In this hierarchical regression models some residuals, graph and other techniques were used to assess peculiarities or the distinctive features of the model with regards to the data.

Residuals: Like in other regression-type models, residuals ($e_{0i}$ and $u_{0i}$ ) play an important exploratory role for model checking in multilevel models. If their is outliers researchers recommended that remove outliers before conducting analysis and use robust regression technique for detecting influential observation[43].

Heteroscedasticity: The comprehensive nature of most algorithms for estimating the HLM makes it relatively straightforward to include some possibilities for modeling non-constant variances of the random effects; these were indicated by complex variation.

Normality: The Normal score plots are fairly straight, suggesting that the normal distribution assumption is reasonable for residuals that checked by plots of residuals.

Linearity: This assumes that multivariate expected value of dependent variable is a straight-line function of each independent variable, and slope of that line does not depend on other variables by looking histogram or a Q-Q-Plot.

# CHAPTER FOUR

# 4 RESULTS

## 4.1 Data exploratory

There were a total of 906 visits from 151 subjects; the number of visits per subject varied from 3 to 6 with equally three month interval for all patients. The sample sizes at the six consecutive time points were 151, 124, 120, 76, 58 and 40. There is a sharply increasing degree of missing data over time due to dropouts, missed clinic visits and transferring to other hospital.

Moreover, different socio-demographic and clinical related covariates were encompassed in table 1 bellow. Among a total of 151 respondents, 87 (57.6%) were female and three fourth, 115(76.2%), of these respondents were have no prior history of hypertensive. At baseline, about one fourth, 40 (26.5%) of the respondents were categorized in pre stage of type II diabetes while, 62 (41.1%) were in second stage. The result also displayed, the baseline mean and standard deviation of respondents age were 47.4(SD=13.4 in years); refer (Table 1) for the rest details.

The mean and standard deviation of responses for each follow-up are also presented in Table 2 bellow. There was a general decrement of mean value up to third visiting time and then oscillate after third visiting time. However, when we look at the standard deviations there was slight variation among the follow-up and smaller variation taken on the forth measurement.

As can be seen from this Table, the baseline mean of patients were 95.2 (SD=16.6 mmHg), 171.1 (SD=30.7 mmHg) and 347.5(SD=71.3 mmHg) with respect to DBP, SBP and T2DM. Likewise; the overall mean for respondents were 91.2 (SD=16.7 mmHg), 163.9 (SD=30.6 mmHg) and 332.5(SD=72.4 mmHg) with respect to DBP, SBP and T2DM. In general, when we look at the trend of the mean for the first three consecutive visits the mean is slightly decreased and after the fourth visit time the mean is osculated.

Table 1: Frequency distribution of categorical variable and mean of continuous variable

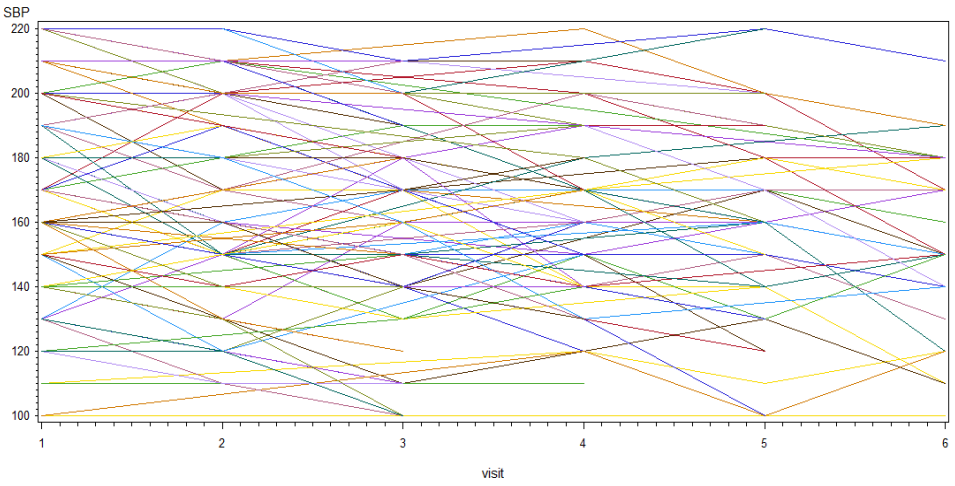
| variables | Characteristics | Frequency (n) | Percent |
|---|---|---|---|
| Gender | Female | 87 | 57.6 |
| | Male | 64 | 42.4 |
| Residence | Urban | 88 | 48.3 |
| | Rural | 63 | 41.7 |
| Number of related disease | No disease | 55 | 36.4 |
| | one disease | 61 | 40.4 |
| | two disease | 35 | 23.2 |
| Previous history of hypertension | Yes (present) | 36 | 23.8 |
| | No (absent) | 115 | 76.2 |
| baseline of Stage DBP | Pre-stage | 40 | 26.5 |
| | first stage | 49 | 32.5 |
| | second stage | 62 | 41.1 |
| baseline of Stage SBP | Pre-stage | 38 | 25.2 |
| | first stage | 72 | 47.7 |
| | second stage | 41 | 27.2 |
| Continuous predictor | Mean | Std dev | |
| Age | 47.4 | 13.4 | |

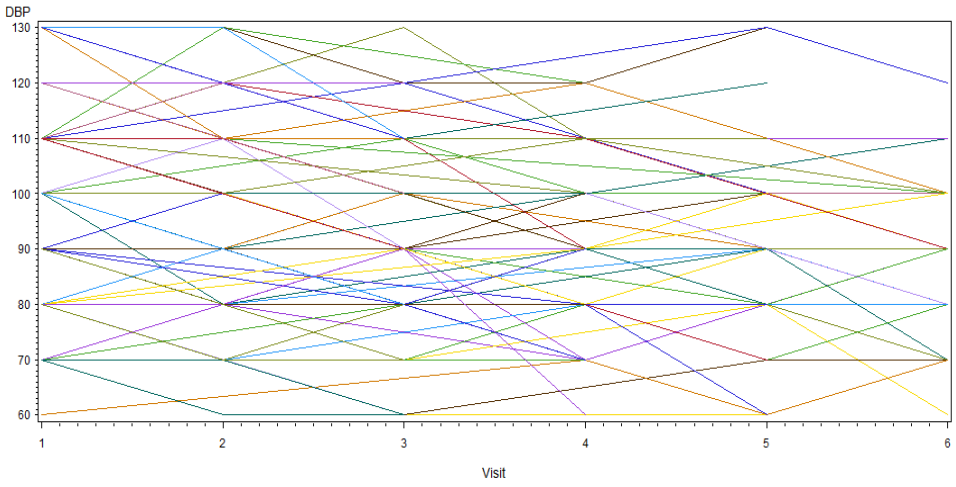
Table 2: Summery statistics of response variable at each follow-up

| Time | DBP | | SBP | | DM | |
|---|---|---|---|---|---|---|
| | Mean | std dev | Mean | std dev | Mean | std dev |
| Time=1 | 95.2 | 16.6 | 171.1 | 30.7 | 347.5 | 71.3 |
| Time=2 | 92.6 | 17.7 | 167.3 | 31.0 | 338.9 | 73.6 |
| Time=3 | 87.5 | 16.3 | 158.3 | 30.5 | 323.7 | 71.2 |
| Time=4 | 89.1 | 14.0 | 161.4 | 26.2 | 324.7 | 65.8 |
| Time=5 | 90.5 | 17.9 | 159.8 | 32.0 | 316.7 | 78.9 |
| Time=6 | 88.1 | 14.7 | 152.6 | 28.3 | 319.5 | 69.9 |
| Overall | 91.2 | 16.7 | 163.9 | 30.6 | 332.5 | 72.4 |

### 4.1.1 Individual profile plot of growth curve analysis

To underpin the model building and visualize the pattern of DBP, SBP and T2DM measurements of the patient's overtime, the overall individual plots were considered. Figures 1 indicated that the variability (within and between patients) is slightly decreasing trend on each respondents throughout the follow up. For each responses most (but not all) observations were slightly turn down throughout the follow up. Likewise, there is variation with in subject throughout the time by decreasing each response from visit to visit.



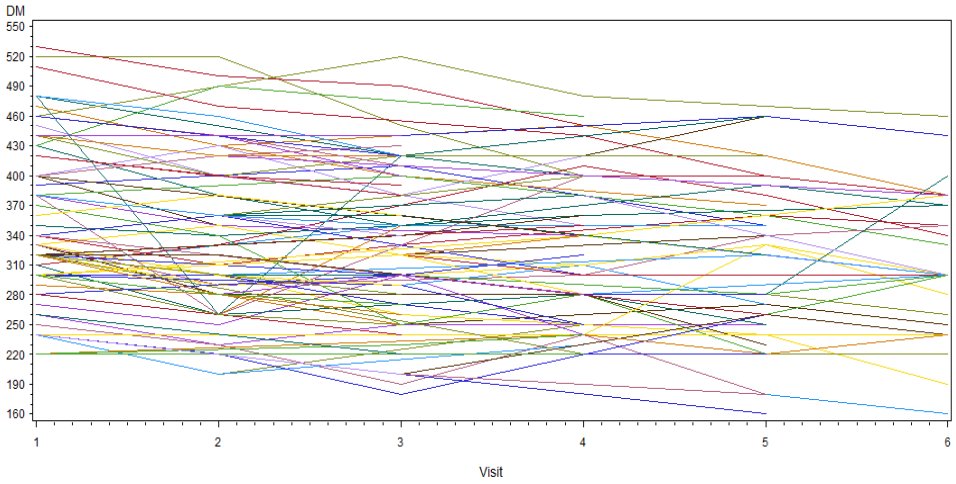



Figure 1: Individual profile plots of growth curve for each response

### 4.1.2 Mean plot by some selected predictor

The mean plot and confidence intervals of respondents measured on groups of subjects over six clinical visits were shown in the bellow graph. Thus, the average mean plot of respondent were decreases with increasing time for residence. Specifically, rural patients was less mean plot than urban patients.

Similarly, in the appendix C, the mean plot of respondents for each response were decreases with increasing time in gender. In addition to this, the mean plots respondent were decreases with increasing time for number of related disease; specifically, patients that have no disease and have one disease were lower mean plot as compared with two or more disease(i.e., patients that have two or more related disease mean plot were slightly increased with increasing time for all responses).

Lastly, the mean plot for baseline stage of hypertension (SBP and DBP) were slightly decreasing over the six follow-up. In this result the mean plot were lower for pre-stage and first stage as compared with second stage.

General, we observe that after the fourth visiting time the confidence interval is large (wide) that indicates poorly estimate the response (i.e., only the first four visit were best explain the model).

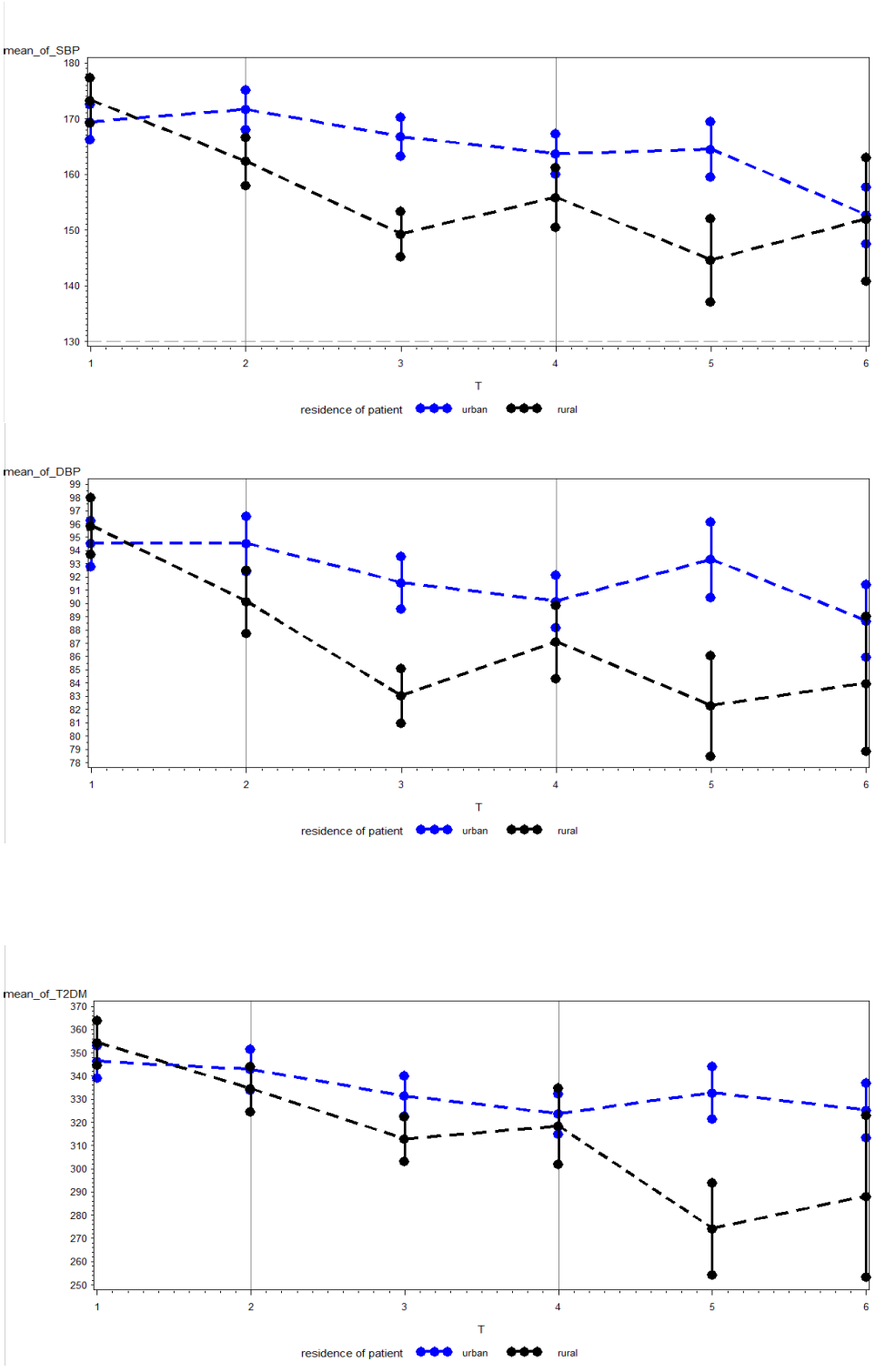



Figure 2: Mean plot of residence with their confidence interval for each response

## 4.2 Multivariate analysis

Before doing multivariate growth curve analysis and multivariate multilevel analysis to check the association of those three responses over time is interesting; Thus, autoregressive (order 1) covariance structure is better to fit the model; because the data are measured in equal spaced every three month and due to missing the data are unbalanced. Now, the estimated variance-covariance matrix for the multivariate analysis though the follow-up were

$$G = \begin{bmatrix} 1122.52 & 509.51 & 231.27 & 104.97 & 47.6470 & 21.6270 \\ & 1122.52 & 509.51 & 231.27 & 104.97 & 47.6470 \\ & & 1122.52 & 509.51 & 231.27 & 104.97 \\ & & & 1122.52 & 509.51 & 231.27 \\ & & & & 1122.52 & 509.51 \\ & & & & & 1122.52 \end{bmatrix}$$

$$R = \begin{bmatrix} \delta_1^2 & \delta_{12} & \delta_{13} \\ & \delta_2^2 & \delta_{23} \\ & & \delta_3^2 \end{bmatrix} = \begin{bmatrix} 16.5115 & 62.4122 & 33.1621 \\ & 215.20 & 278.12 \\ & & 46.5933 \end{bmatrix}$$

Now, the marginal correlation between DBP and T2DM as a function of time is given by:

$$\frac{(104.97 + t(509.51 + 231.27) + t^2(1122.52))}{\sqrt{1122.52 + 2t^2(509.51 + 1122.52) + 16.5115} * \sqrt{1122.52 + 2t^2(509.51 + 1122.52) + 215.20}}$$

The marginal correlation between DBP and SBP as a function of time is given by:

$$\frac{(104.97 + t(231.27 + 509.51) + t^2(1122.52))}{\sqrt{1122.52 + 2t^2(231.27 + 1122.52) + 16.5115} * \sqrt{1122.52 + 2t^2(231.27 + 1122.52) + 46.5933}}$$

The marginal correlation between T2DM and SBP as a function of time is given by:

$$\frac{(104.97 + t(509.51 + 231.27) + t^2(104.6470))}{\sqrt{1122.52 + 2t^2(509.51 + 1122.52) + 215.20} * \sqrt{1122.52 + 2t^2(509.51 + 1122.52) + 46.5933}}$$

The result revealed that the correlation between the diastolic blood pressure and type II diabetes mellitus at baseline was 0.437 and the remaining marginal correlation for visit 2 to visit 6 were 0.425, 0.406, 0.393, 0.379 and 0.379 respectively.
Likewise for diastolic and systolic blood pressure at baseline the correlation was 0.446 and the remaining marginal correlation for visit 2 to visit 6 were 0.427, 0.407, 0.394, 0.385 and 0.379 respectively.
Lastly for type II diabetes mellitus and systolic blood pressure at baseline the correlation was 0.436 and the remaining marginal correlation for visit 2 to visit 6 were 0.423, 0.405, 0.392, 0.385 and 0.379 respectively over the 2 year period follow up time.

## 4.3 Growth curve analysis

Based on deviance information criteria's, random intercept and slope model was more parsimonious for separate response than intercept only model and random intercept model based on autoregressive-1 covariance structure.

Table 3: Deviance information criteria for univariate growth curve analysis

| Separate model for GCA | Deviance information criteria | | |
|---|---|---|---|
| | DBP | SBP | T2DM |
| Intercept only GCA | 19839.0 | 22881.9 | 27143.7 |
| Random intercept GCA | 20162.1 | 23229.1 | 27470.6 |
| Random intercept and slope GCA | **18055.8** | **21004.9** | **25497.4** |

Therefor, multivariate growth curve is the only mechanism since in univariate growth curve doesn't allow random intercept and slope. Thus; the multivariate growth curve is as follow:

Table 4: Growth curve analysis for multivariate response by creating dummy

| Effect | Estimate | SE | Lower | Upper | P-value |
|---|---|---|---|---|---|
| First-order polynomial | -5.1750 | 1.3004 | -7.7240 | -2.6260 | <0.0001 |
| second-order polynomial | -0.6896 | 0.2371 | -1.1544 | -0.2247 | 0.0036 |
| third-order polynomial | -0.09237 | 0.04317 | -0.1770 | -0.00775 | 0.0324 |
| fourth-order polynomial | -0.01244 | 0.007602 | -0.02734 | -0.002464 | 0.0491 |
| order polynomial | cov-parameter | Estimate | SE | P-value | |
| | Variance($u_{0i}$) | 69.2085 | 9.6915 | <0.0001 | |
| | Variance($u_{1i}$) | 69.6328 | 9.7721 | <0.0001 | |
| | Variance($u_{2i}$) | 70.3776 | 9.9542 | 0.0003 | |
| | Variance($u_{3i}$) | 71.2416 | 10.1535 | 0.0034 | |
| | Variance($e_{0i}$) | 11232 | 158.54 | <0.0001 | |

From table 4 outputs shows that, the effect of each visiting time to develop hypertension in type II diabetes by creating five dummy variables to considering time as discrete variable. In contrast, the four approaches describe patterns of change quite differently because the profile analysis allows for individual differences in individual growth curve linear, quadratic, cubic and quartic slopes (i.e., $p<0.05$).

The baseline time shows that hypertension in type-II diabetes were decreasing over time. The regression coefficient belonging to the first dummy variable represents the difference of hypertension in type-II diabetes between the first and the second measurement. This difference is -5.1750, with a standard error of 1.3004. This means that as time increase from (first visit to second visit) hypertension in type-II diabetes were decreases by 5.1750 unite. The corresponding p-value is again highly significance ($p <0.001$). The regression coefficient for the second dummy variable represents the difference of hypertension in type-II diabetes between the second and the third measurement. This difference is -0.6896 (i.e., as time increase from second visit to third visit, hypertension in type-II diabetes were decrease by 0.6896 unite) with the corresponding p-value ($p= 0.0036$). The regression coefficient belonging to the third dummy variable represents the difference

between the third measurement and the fourth measurement. In this measurement, developing hypertension in type II diabetes were also decreased by 0.09237 unite; which is significant (p-value 0.0324); and the like. Furthermore, the fifth and sixth visit dummy measurement has no significant difference. Therefor we conclude that from the above analyses developing hypertension in type-II diabetes over time is best described by a first, second, third and fourth order polynomial function with random intercept and slope GCA(i.e., fifth and sixth visiting time is not necessary).

As demonstrated from the table, there was statistically significant intercept and slope variability on respondents at Debre Tabor referral hospital (i.e., p<.0001). Since only the first four measurement is best fit on the fixed effect, the random effect of those visit time have significance slope difference.

## 4.4 Multilevel analysis

Before doing multilevel analysis intra-class correlation was give strong evidence that variability is occurring between the patients or not by supporting of HLM. From the null model without adding any explanatory analysis intra-class correlation for multivariate response were calculated as

$$ICC_{individual} = \frac{{\sigma_{u_0}}^2}{{\sigma_{e_0}}^2 + {\sigma_{u_0}}^2} = \frac{1335.11}{1335.11 + 779.18} = 0.6315$$

Therefore, 63.15% of the variation for hypertension in type II diabetes mellitus were exists between patient's variation and the remaining 36.85% of variation is explained by within variation. Since that, the highest variation is explained in the higher level and this variation is statistically significant (p<0.0001), multilevel model is appropriate. Multilevel model is scope with missing data on the response variables[21]

### 4.4.1 Results of univariate multilevel analysis

Using deviance information criteria's the best univariate multilevel model were selected in table bellow. From this table random intercept and slope model with time invariant covariate was more parsimonious (the small values of deviance information criteria

have the best fit the model) than all other model based on autoregressive-1 covariance structure.

Table 5: Deviance information criteria for univariate multilevel analysis

| Random effect | DBP | SBP | T2DM |
|---|---|---|---|
| Intercept only model | 4773.6 | 5455.3 | 6475.3 |
| Random intercept time varying covariate model | 3957.2 | 4531.5 | 5345.7 |
| Random intercept time invariant covariate model | 3953.7 | 4526.7 | 5336.6 |
| Random coefficient time varying covariate model | 3887.0 | 4440.5 | 5265.6 |
| Random coefficient time invariant covariate model | 3883.0 | 4435.3 | 5257.0 |
| Random coefficient interaction model | 3926.5 | 4484.2 | 5282.2 |

Therefore, the univariate multilevel analysis were fitted in table bellow by using this selected model. In case all predictors were significant in multi-variable analysis except previous (prior) history of hypertension; so that this model were analyzed by the remaining variable.

Table 6: Univariate multilevel analysis output

| **SBP** | | | | | | |
|---|---|---|---|---|---|---|
| Predictors | Category | Estimate | SE | Lower | Upper | Pr > \|T\| |
| Intercept | Continuous | 100.95 | 7.2721 | 82.2613 | 119.65 | <.0001 |
| Age | Continuous | 0.4592 | 0.08709 | 0.2881 | 0.6303 | <.0001 |
| Residence | Rural | -8.8828 | 2.7975 | -16.0741 | -1.6916 | 0.0247 |
| (reference= Urban) | | | | | | |
| related | No disease | -11.7174 | 3.0059 | -18.4148 | -5.0199 | 0.0030 |
| disease | One disease | -11.0499 | 2.7704 | -17.2227 | -4.8771 | 0.0026 |
| (reference ≥ Two disease) | | | | | | |
| Baseline | Pre-stage | -12.3113 | 3.9316 | -21.0714 | -3.5512 | 0.0107 |
| stage of SBP | First stage | -7.7819 | 2.8370 | -14.1032 | -1.4606 | 0.0207 |
| (reference= Second stage) | | | | | | |
| Baseline | Pre-stage | -11.8265 | 3.8872 | -20.4877 | -3.1653 | 0.0124 |
| stage of DBP | First stage | -9.1436 | 2.8308 | -15.4511 | -2.8361 | 0.0090 |
| (reference= Second stage) | | | | | | |
| BaselineT2DM | Continuous | 0.1781 | 0.01730 | 0.1441 | 0.2121 | <.0001 |
| **DBP** | | | | | | |
| Intercept | Continuous | 57.0932 | 4.2893 | 46.0671 | 68.1193 | <.0001 |
| Age | Continuous | 0.3662 | 0.05306 | 0.2619 | 0.4704 | <.0001 |
| Residence | Rural | -4.5831 | 1.6616 | -8.8543 | -0.3118 | 0.0399 |
| (reference= Urban) | | | | | | |
| related | No disease | -5.9622 | 1.7944 | -9.9604 | -1.9640 | 0.0077 |
| disease | One disease | -5.6449 | 1.5618 | -9.1248 | -2.1649 | 0.0047 |
| (reference ≥ Two disease) | | | | | | |
| Baseline | Pre-stage | -9.3625 | 2.2719 | -14.4246 | -4.3005 | 0.0021 |
| stage of SBP | First stage | 6.7165 | 1.6842 | -10.4692 | -2.9638 | 0.0026 |
| (reference= Second stage) | | | | | | |
| BaselineT2DM | Continuous | 0.08058 | 0.01055 | 0.05986 | 0.1013 | <.0001 |

| T2DM | | | | | | |
|---|---|---|---|---|---|---|
| Age Continuous | | 0.8282 | 0.1757 | 0.4829 | 1.1734 | <.0001 |
| Gender (reference= male) | Female | -6.5768 | 3.1175 | -12.7013 | -0.4524 | 0.0354 |
| Residence (reference= Urban) | Rural | -16.7501 | 6.1541 | -32.5697 | -0.9305 | 0.0417 |
| related disease (reference ≥ Two disease) | No disease | -23.8505 | 6.5422 | -38.4274 | -9.2736 | 0.0045 |
| | One disease | -20.2931 | 6.0001 | -33.6621 | -6.9240 | 0.0070 |
| BaselineT2DM Continuous | | 0.8356 | 0.03486 | 0.7671 | 0.9041 | <.0001 |

As we can observe from the fitted univariate multilevel analysis in table 6 above final model was modeled with sets of covariates that include fixed effect parameters age, gender, residence, number of related disease, baseline stage of SBP, baseline stage of DBP and baseline stage of T2DM. Among those covariates there was significance difference (p = 0.05) by intercept, age, residence, number of related disease, baseline stage of DBP, and baseline stage of type II DM for diastolic blood pressure; and those predictors were also significant (p = 0.05) for systolic blood pressure in addition to baseline stage of SBP; lastly age, gender, number of related disease, and baseline stage of T2DM were significant (p = 0.05) for type II diabetes mellitus.

According to result of univariate analysis, residence (rural), gender (female), number of related disease (no disease and one disease), baseline stage of DBP (pre-stage and first stage), and baseline stage of SBP (pre-stage and first stage), were negatively associated with each responses that mean the repeatedly follow up made a particular decrease on each outcomes with (P<0.0001) in this study. But, intercept, age and baseline stage of T2DM were positively associated with each response.

Therefore, from this univariate analysis we observe that SBP, DBP, and T2DM were increase as age and baseline T2DM increase; However, SBP, DBP, and T2DM were

decrease as residence (rural), gender (female), number of related disease (no disease and one disease), baseline stage of DBP (pre-stage and first stage) and baseline stage of SBP (pre-stage and first stage) as compare to their reference.

Table 7: Random coefficient for univariate multilevel analysis

| **Covariance Parameter Estimates for SBP** | | | | |
|---|---|---|---|---|
| Cov Parm | Subject | Estimate | Std. Error | Pr > \|T\| |
| Variance($u_{0i}$) | ID | 5.4192 | 2.3320 | 0.0101 |
| AR(1) | ID | 0.0417 | 1.2400 | 0.9732 |
| Residual($e_{0i}$) | | 197.67 | 13.4497 | <.0001 |
| **Covariance Parameter Estimates for DBP** | | | | |
| Cov Parm | Subject | Estimate | Std. Error | Pr > \|T\| |
| Variance($u_{0i}$) | ID | 2.5978 | 0.8945 | 0.0018 |
| AR(1) | ID | 0.1461 | 0.5982 | 0.8070 |
| Residual($e_{0i}$) | | 62.8670 | 4.2982 | <.0001 |
| **Covariance Parameter Estimates for T2DM** | | | | |
| Cov Parm | Subject | Estimate | Std. Error | Pr > \|T\| |
| Variance($u_{0i}$) | ID | 25.9973 | 9.3038 | 0.0026 |
| AR(1) | ID | 0.04900 | 0.5507 | 0.9291 |
| Residual($e_{0i}$) | | 775.84 | 52.5250 | <.0001 |

Note:- $e_{0i}$ and $u_{0i}$ are random intercept for lowest and highest level, and AR(1) is random slope for those predictor in univariate multilevel analysis.

From table, random slops were fitted for residence, number of related disease, baseline stage of systolic blood pressure, and baseline stage of diastolic blood pressure for subject specific effect. Thus; even if this random coefficient model is better, but the overall random slop were insignificant. Therefore calculating the random variation(G matrix) for each predictor were not that much important.

Now, $R^2_{\cdot}$ is calculated based on the fitted and previous model for showing how much

variation is explained by predictors as follow.

$$R^2_{individual\ for\ SBP}=\frac{\delta^2_{u0/b}\ -\delta^2_{u0/m}}{\delta^2_{u0/b}}=\frac{21.7443\ -5.4192}{21.7443}=0.7508$$

$$R^2_{individual\ for\ DBP}=\frac{\delta^2_{u0/b}\ -\delta^2_{u0/m}}{\delta^2_{u0/b}}=\frac{10.5927\ -2.5978}{10.5927}=0.7548$$

$$R^2_{individual\ for\ T2DM}=\frac{\delta^2_{u0/b}\ -\delta^2_{u0/m}}{\delta^2_{u0/b}}=\frac{104.47\ -25.9973}{104.47}=0.7512$$

Thus, 75.08%, 0.75.48% and 75.12% of the proportion of variance is explained by the second level (individual) for SBP, DBP, and T2DM with predictors.

### 4.4.2 Results of multivariate multilevel analysis

Like to above univariate multilevel analysis, deviance information criteria's is used to select the best multivariate multilevel model.

Table 8: Deviance information criteria for multivariate multilevel analysis

| Random effect | Deviance |
|---|---|
| Intercept only model | 14003.2 |
| Random intercept time varying covariate model | 12123.6 |
| Random intercept time invariant covariate model | 12112.6 |
| Random intercept and slope model with time varying covariate | 12118.4 |
| Random intercept and slope model with time invariant covariate | 12108.0 |
| Random intercept and slope model with interaction | 12160.9 |

Thus, from table 8 above model random intercept and slope model with time invariant covariate was more parsimonious than all other model.

Now, Likelihood ratio test were compares the sum differences between univariate deviances, and multivariate deviance of the models to a chi-square distribution; since the univariate models were considered to be nested within the multivariate model. Thus, the -2log likelihood of the univariate random intercept and random slope model, and multivariate random intercept and random slope model were ($X^2$= 1467.3, p-value <0.0001), which follows a Chi-square distribution with twelve degrees of freedom. Again, there are 12 degrees of freedom, because the random slope(i.e., residence, number of related disease, baseline stage of SBP and DBP) random intercept were estimated. This significant likelihood ratio test indicates that MVML random intercept and random slope model is a better fit to the data than the separate univariate random intercept and slope models. Therefore, multivariate random coefficient model were fitted for both fixed part of predictors, and the random part (within and between) subject correlations by using the selected autoregressive variance-covariance structure.

Table 9: Multivariate multilevel analysis output

| **Solution for fixed effects** | | | | | | | |
|---|---|---|---|---|---|---|---|
| Predictors | Category | response | Estimate | SE | Lower | Upper | Pr > \|T\| |
| Intercept | Continuous | SBP | 99.6681 | 10.3319 | 79.3913 | 119.94 | <.0001 |
| Intercept | Continuous | DBP | 57.4218 | 10.2752 | 37.2561 | 77.5874 | <.0001 |
| Intercept | Continuous | T2DM | 13.8753 | 10.2808 | -6.3012 | 34.0518 | 0.1775 |
| Age | Continuous | SBP | 0.3739 | 0.1450 | 0.08933 | 0.6586 | 0.0101 |
| Age | Continuous | DBP | 0.3367 | 0.1448 | 0.05242 | 0.6209 | 0.0203 |
| Age | Continuous | T2DM | 0.5707 | 0.1449 | 0.2864 | 0.8550 | <.0001 |
| Gender | Female | SBP | 1.4942 | 2.4419 | 6.2865 | 3.2981 | 0.5407 |
| . | Female | DBP | 0.9541 | 2.4451 | 5.7527 | 3.8446 | 0.6965 |
| . | Female | T2DM | -7.5109 | 2.4451 | -2.7124 | 12.3095 | 0.0022 |
| (reference= male) | | | | | | | |
| Residence | rural | SBP | -7.2478 | 2.5395 | -12.2318 | -2.2638 | 0.0044 |
| . | rural | DBP | -3.6497 | 1.5375 | -8.6298 | -1.3303 | 0.0407 |
| . | rural | T2DM | -11.3837 | 2.5389 | -16.3664 | -6.4010 | <.0001 |
| (reference= urban) | | | | | | | |
| Number | No disease | SBP | -7.7468 | 3.0708 | -13.7734 | -1.7203 | 0.0118 |
| Of | No disease | DBP | -4.3127 | 3.0696 | -10.3368 | 1.7115 | 0.1604 |
| related | No disease | T2DM | -10.5187 | 3.0703 | -16.5444 | -4.4930 | 0.0006 |
| disease | One disease | SBP | -7.7148 | 2.3012 | -12.2309 | -3.1986 | 0.0008 |
| . | One disease | DBP | -3.9097 | 2.2910 | -8.4060 | 0.5866 | 0.0882 |
| . | One disease | T2DM | -13.3050 | 2.2980 | -17.8150 | -8.7950 | <.0001 |
| (reference $\geq$ Two disease) | | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Baseline | Pre-stage | SBP | -14.3142 | 4.7323 | -23.6015 | -5.0269 | 0.0026 |
| stage | Pre-stage | DBP | -2.8178 | 0.7282 | -4.2451 | -1.3905 | 0.0014 |
| of SBP | pre-stage | T2DM | -8.5467 | 2.6798 | -13.8060 | -3.2874 | 0.0015 |
| . | First stage | SBP | -0.4640 | 4.7298 | -9.7465 | 8.8184 | 0.9219 |
| . | First stage | DBP | -2.2906 | 2.6768 | -7.5439 | 2.9627 | 0.3924 |
| . | First stage | T2DM | -6.5288 | 2.6766 | -11.7817 | -1.2759 | 0.0149 |
| (reference= Second stage) | | | | | | | |
| Baseline | Pre-stage | SBP | -12.4275 | 4.7077 | -21.6665 | -3.1884 | 0.0266 |
| stage | Pre-stage | DBP | -10.0463 | 4.7051 | -19.2803 | -0.8123 | 0.0330 |
| of DBP | pre-stage | T2DM | -5.5895 | 4.7054 | -14.8241 | 3.6451 | 0.2352 |
| . | First stage | SBP | -8.4662 | 2.8123 | -13.9855 | -2.9469 | 0.0027 |
| . | First stage | DBP | -6.4454 | 2.8104 | -11.9610 | -0.9299 | 0.0220 |
| . | First stage | T2DM | 1.0138 | 2.8111 | -4.5032 | 6.5307 | 0.7185 |
| (reference= Second stage) | | | | | | | |
| Baseline | Continuous | SBP | 0.1976 | 0.02821 | 0.1422 | 0.2530 | <.0001 |
| DM | Continuous | DBP | 0.08509 | 0.02801 | 0.03012 | 0.1401 | 0.0024 |
| . | Continuous | T2DM | 0.8708 | 0.02802 | 0.8158 | 0.9258 | <.0001 |

From this multivariate multilevel analysis, most parameters are statistically significant; among those residence, gender, number of related disease, baseline stage of SBP, and baseline stage of DBP were negatively associated with those three outcomes. On the other hand, age, and baseline T2DM were positively associated with those three outcomes.

The overall mean of SBP and DBP were increased by 99.6681(S.E. =10.3319) and 57.4218 (S.E. = 10.2752) mmHg respectively keeping all predictors constant.

From the output age was positively associated with all three responses. Thus, as age of patients increased by one year, the average rate change of systolic blood pressure were increased by 0.3739 mmHg with corresponding ($p<0.0101$) value; and the diastolic blood pressure of patients were increased by 0.3367 mmHg with its ($p=0.0203$) value; lastly the type II diabetes mellitus of patients were also increased by 0.5707 mmol/L ($p<0.0001$)

by keeping all other predictors constant.

In addition, gender were also significant mean difference (p=0.0022) for type II diabetes mellitus; for the case, the average type II diabetes mellitus were 7.5109 mmol/L times lower for females patients as compare with male by keeping all other predictors constant. Moreover, residence was significantly associated with three response. Thus, the average systolic blood pressure, diastolic blood pressure and type II diabetes mellitus were 7.2478 mmHg (P=0.0044) times, 3.6497 mmHg (P=0.0407) times, and 11.3837 mmol/L (p<.0001) times lower respectively for rural patients as compared to urban patients by keeping all other variables as constant.

Similarly, number of related disease was significantly associated with systolic blood pressure and type II diabetes mellitus with small p-value; for instance, the average systolic blood pressure and type II diabetes mellitus of patients were -7.7468 mmHg (P = 0.0118) and 10.5187 mmol/L (P =0.0006) times lower average for patients has no disease as compared with patients that has two or more disease by keeping all other variables as constant; In similar way, the average systolic blood pressure and type II diabetes mellitus of patients were 7.7148 mmHg (P = 0.0008) and 13.3050 mmol/L (p< 0.0001) times lower average for patients has one related disease as compared with patients that has two or more disease by keeping all other variables as constant.

The baseline stages of SBP were also significantly associated with those three response jointly; for instance patients baseline stages of SBP were at pre-stage are 14.3142 mmHg (p=0.0026), 2.8178 mmHg (p=0.0014) and 8.5467 mmol/L(p=0.0015) times lower evolution of systolic blood pressure, diastolic blood pressure and type II diabetes mellitus respectively as compare with patients who are in second (last) stage by keeping all other variables as constant; and patients baseline stages of SBP were at first stage were 6.5288 mmol/L(p=0.0149) times lower average type II diabetes mellitus as compare with patients who are in second stage by keeping all other variables as constant.

On other hand, baseline stages of DBP were also significantly associated for both systolic blood pressure and diastolic blood pressure; Thus, the average systolic and diastolic blood pressure of patients were 12.4275 mmHg (p = 0.0266) and 10.0463 mmHg (p = 0.0330)

times lower for pre-stage patients as compare with patients who are in second stage by keeping all other variables as constant. Likewise, the average systolic and diastolic blood pressure of patients were 8.4662 (p=0.0027) mmHg and 6.4454 (p=0.0220) mmHg times lower for first stage as compare with patients who are in second stage by keeping all other variables as constant

Lastly, baseline T2DM were positively associated with all three responses. As baseline T2DM increase by one mmol/L, the rate change of systolic blood pressure of patients were increase by 0.1976 mmHg (p<0.0001), the diastolic blood pressure of patients were increased by 0.08509 mmHg (p=0.0024), and the type II diabetes mellitus of patients were also increased by 0.8708 mmol/L (p<0.0001) by keeping all other variables as constant.

Table 10: Random coefficient for multivariate multilevel analysis

| **Covariance Parameter Estimates** | | | | | | |
|---|---|---|---|---|---|---|
| Cov Parm | Subject | Estimate | Std. Error | Lower | Upper | Pr > \|T\| |
| Variance($u_{0i}$) | ID | 13.9967 | 1.9197 | 10.8825 | 18.6768 | <.0001 |
| AR(1) | ID | -0.4118 | 0.1572 | -0.7200 | -0.1036 | 0.0088 |
| Residual($e_{0i}$) | | 254.11 | 3.9298 | 246.58 | 261.99 | <.0001 |

Note:-like to table 7, $e_{0i}$ and $u_{0i}$ are random intercept for lowest and highest level, and AR(1) is random slope for predictors.

Alike fixed parameter estimation and testing, variability analysis of random effects are also another important aspects. High variability is the indicator of less accuracy or high error on prediction of the association of outcome evolutions with respective risk factors. Then as shown in Table 10, the between random intercept variance is estimated to be 13.9967(S.E. = 1.9197) and the within random intercept variance is estimated to be 254.11(S.E. = 3.9298). Thus, the variability of between random intercepts is lower than that of variability of within intercepts. This multivariate random coefficient model shows small slope variation 0.4118 (p=0.0088) and the slope variation of this model is due to the predictor number of related disease, baseline stage of SBP, and baseline stage of DBP see on the appendix D using covariance matrix.

Now; alike to above univariate multilevel analysis, for multivariate multilevel analysis explained variance ($R^2$) is tested by computing difference of the deviance baseline model (b) and the current (m) model.

$$R^2_{individual}=\frac{\delta^2_{u0/b} \ -\delta^2_{u0/m}}{\delta^2_{u0/b}} = \frac{24.9727 \ - 13.9967}{24.9727} = 0.4395$$

For multivariate response the second level predictors are explain 43.95% of variation; and the remaining 56.05% are explained by lowest level variation.

## 4.5 Goodness of fit test

Goodness of fit of the fitted multivariate random coefficient models was assessed using deviance-based chi-square test. Accordingly, the deviance-based chi-square test provided chi-square value of 64.26 ($p<0.001$) which would imply good fit for the hypertension in type II diabetes mellitus.

## 4.6 Model diagnosis

Based on figure 3 the fitted multivariate random intercepts and slope model possible presence of outliers and influential values were checked. Thus, from the multivariate multilevel model of residual Vs predicted graph shows there are some outliers(but their leverage is between -3 and 3). The graph indicate that the variability of error in multivariate response was almost constant, and it has no any pattern. That means the error does not far deviate from each other and distance of individual residual were equally far from the horizontal line without any pattern. Therefore, linearity and heteroscedastie of the errors were satisfied. From the residual Vs quintile plots of multivariate random intercepts and random slopes shows, even if it seems slight deviation of normality at the bottom and top the normality assumption is fulfilled; and also from the graph of percentile Vs residual, the multivariate normality plot is satisfied.

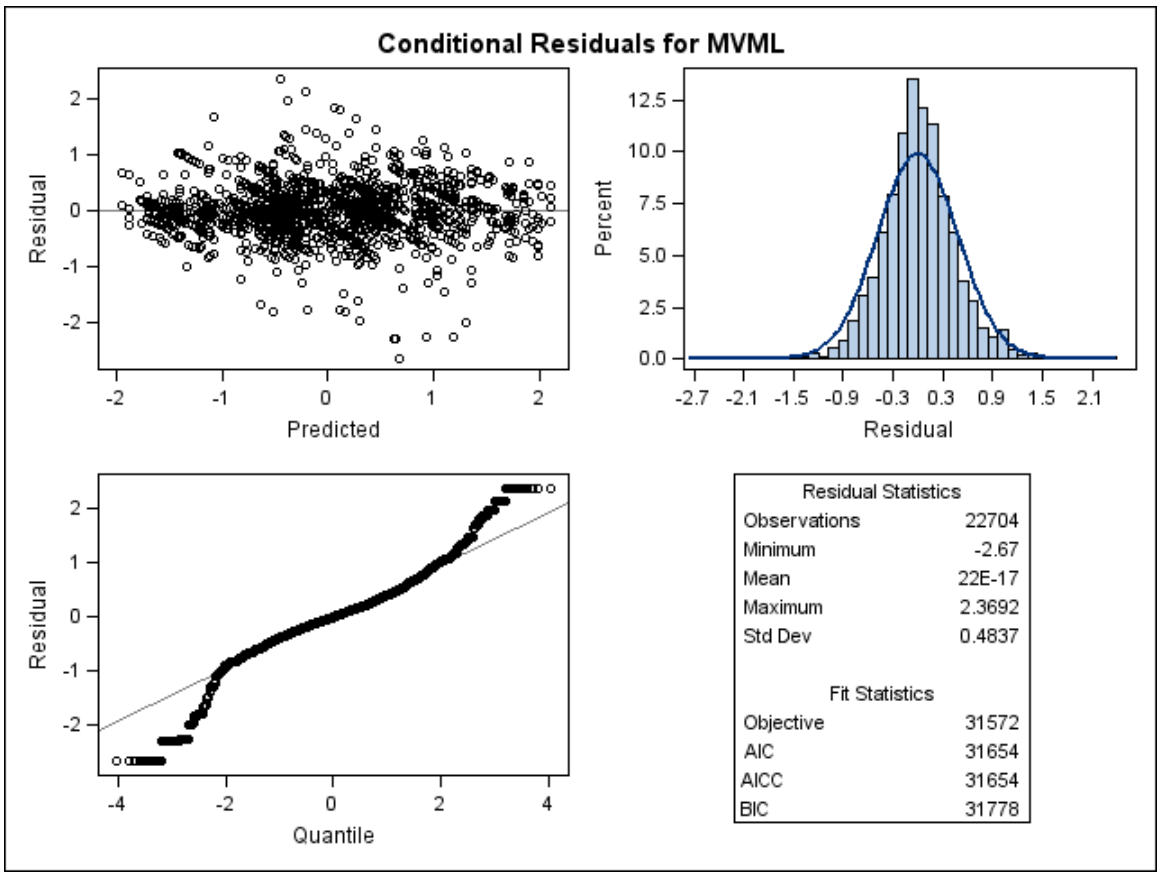



Figure 3: Residual plot for the fitted multivariate multilevel model

## 4.7 Discussion

Based on different well organized literatures, some discussions were organized as follows. In the study multivariate model were best to fit for longitudinal data than univariate model. This is supported by different researcher[39, 40, 50] the joint multivariate models were simpler as compared to the separate univariate models as their effective number of parameters was smaller.

Based on multivariate analysis of longitudinal data the association between systolic and diastolic blood pressure were decrease over a time. This finding is consistent with[53]. In this finding, the association between type II diabetes with systolic and diastolic blood pressure were also decrease over a time; but I don't get any supportive or contradict idea. Beside from growth curve analysis the finding of the study provides direct evidence that an increasing visiting time, systolic BP, diastolic BP, and type II diabetes mellitus were reasonably decreasing through out the follow up time. The finding is consistent with the latest literature [28, 53, 54]

According to the results, age was an important demographic predictor for hypertension (SBP and DBP) and type II diabetes mellitus; that as age increase in year, the rate of hypertension (SBP and DBP) and type II diabetes mellitus were increase. This result in-lined with the previous study that,the older age has significance factor for hypertension and type II diabetes mellitus[13, 16, 22]; and in Ethiopia study also consistent with [1, 24] older age has significance effect for diabetes mellitus, and as [2, 25] consist older age has also significance effect for hypertension.

According to the study, gender was an important predictor that revealed males are more affected in type-II diabetes mellitus than females.This estimated result also consistent with similar previous studies conducted by different scholars [1, 15, 20].

The study also showed that residence was an important demographic predictor that revealed rural patients were lower average hypertension(SBP and DBP) and type II diabetes mellitus as compared with urban patients. This finding is consistent or similar with the previous studies conducted by different scholars [1, 6, 52].

Related disease was also an important clinical variable that patients have no disease and

one related disease were less likely average systolic blood pressure and type II diabetes mellitus than those patients who have two or more disease. This result were consistent with the study [46, 47] related disease were increase type II diabetes mellitus; and consistent with the study [53] patients that have related disease were increase their systolic blood pressure.

Results from this study demonstrate that patients whose baseline stage of hypertension are at pre-stages were less likely average hypertension(SBP and DBP) and type II diabetes mellitus as compared with patient who were in second stage. But this estimated result was contradict with previous studies conducted by [28]. The reason may be personal lifestyle or medication after they diseased.

Lastly the finding provides an evidence that hypertension increase as a patients baseline type II diabetes increase. This also consistent with previous finding [15, 19, 27]

# CHAPTER FIVE

# 5 CONCLUSION AND RECOMMENDATION

## 5.1 Conclusion

The main aim of this thesis was to identify factors that affect SBP, DBP and T2DM in multivariate form. Thus, data analysis shows that for all categorical predictors after the fourth visit (follow-up) the confidence interval were wider and weaker to estimate the response. Inferentially, multivariate analysis, growth curve analysis and multilevel analysis were considered for fitting hypertension and type II diabetes mellitus measured longitudinally. Thus, multivariate growth curve analysis and multivariate multilevel analysis with random intercept and slope were better to fit the analysis over univariate growth curve analysis and univariate multilevel analysis by using the preferred autoregressive covariance structure.

Based on this multivariate growth curve analysis out of six repeated measure only the first four visit(i.e., linear, quadratic, cubic, and quartic ) are important to determine an appropriate significant model.

From the multivariate multilevel analysis of random coefficient model, the fixed part shows that the evolution of systolic blood pressure and diastolic blood pressure were significantly differ with intercept, age, residence, number of related disease, baseline stage of SBP, baseline stage of DBP,and baseline T2DM; Like wise, type II diabetes mellitus was significantly differ by age, gender, residence, number of related disease, baseline stage of SBP and baseline T2DM. Likewise, the covariance parameter (random part) of this model concludes that more than half percent of variation are explained by within patients(i.e., level-1)

Finally we conclude that the multivariate model suggests, there was a moderate positive association between the response DBP with T2DM, DBP with SBP, and T2DN with SBP overtime.

## 5.2 Recommendation

Based on the findings of the study, we forward the following possible recommendations: First and most interesting thing to identify the factors of hypertension and type II diabetes were accesses and quality of data; Therefor, the national health minister should prepare the chart patients card by including the following variable alcohol, smoking exposure, chat chewing, coffee drinking, race/ethnicity, anti-diabetics, adherence to lifestyle or medication, prior history of gestational diabetes, psychological stress, physical activity, lifestyle, BMI, history of hypoglycemic control and Cholesterol level in addition to age, gender, residence, co-related disease prior history of HT and prior history of DM; and the health professionals(Doctor's) must be recored each variables properly.

Thus, the policy makers should make this chronic illness a part of the public health agenda, and they should plan timely interventions. Intervention measures at the community level should be undertaken using health education and other measures by providing an emphasis on the prevention, early detection, and treatment of hypertension and type II diabetes mellitus. Furthermore, researchers and health care providers should work to uncover the burden of hypertension and type II diabetes mellitus overall.

Further studies should be conducted by taking three or four level multivariate model using longitudinal data into account to assess the variation of hypertension and type II diabetes mellitus across regional level.

# References


[1] Solomon Mekonnen Abebe, Yemane Berhane, Alemayehu Worku, and Abebayehu Assefa. Diabetes mellitus in north west ethiopia: a community based study. *BMC public health*, 14(1):97, 2014.

[2] Solomon Mekonnen Abebe, Yemane Berhane, Alemayehu Worku, and Assefa Getachew. Prevalence and associated factors of hypertension: a crossectional community based study in northwest ethiopia. *PloS one*, 10(4):e0125210, 2015.

[3] Chalachew Misganaw Alemayehu and Sofia Kebede Birhanesilasie. Assessment of stoke patients: occurrence of unusually high number of haemorrhagic stroke cases in tikur anbessa specialized hospital, addis ababa, ethiopia. *Clin Med Res*, 2(5):94–100, 2013.

[4] KM Ashraful. An observation on a small group of hypertensive and diabetic patients in bangladesh. *MOJ Biol Med*, 2(1):181–184, 2017.

[5] Shiferaw Birhanu Aynalem and Ayalew Jejaw Zeleke. Prevalence of diabetes mellitus and its risk factors among individuals aged 15 years and above in mizan-aman town, southwest ethiopia, 2016: a cross sectional study. *International journal of endocrinology*, 2018, 2018.

[6] A Bhansali, VK Dhandania, M Deepa, RM Anjana, SR Joshi, PP Joshi, SV Madhu, PV Rao, R Subashini, V Sudha, et al. Prevalence of and risk factors for hypertension in urban and rural india: the icmr–indiab study. *Journal of human hypertension*, 29(3):204, 2015.

[7] Liza Bowen. Applied multilevel analysis: A practical guide. jos wr twisk., 2007.

[8] Pedro Bullon, Hubert N Newman, and Maurizio Battino. Obesity, diabetes mellitus, atherosclerosis and chronic periodontitis: a shared pathology via oxidative stress and mitochondrial dysfunction? *Periodontology 2000*, 64(1):139–153, 2014.

[9] Margaret R Burchinal, Lauren Nelson, and Michele Poe. Iv. growth curve analysis: An introduction to various methods for analyzing longitudinal data. *Monographs of the Society for Research in Child Development*, 71(3):65–87, 2006.

[10] Aaron J Cohen, Michael Brauer, Richard Burnett, H Ross Anderson, Joseph Frostad, Kara Estep, Kalpana Balakrishnan, Bert Brunekreef, Lalit Dandona, Rakhi Dandona, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. *The Lancet*, 389(10082):1907–1918, 2017.

[11] Patrick J Curran, James S McGinley, Daniel Serrano, and Chelsea Burfeind. A multivariate growth curve model for three-level data. *APA handbook of research methods in psychology*, 3:335–358, 2012.

[12] Patrick J Curran, Khawla Obeidat, and Diane Losardo. Twelve frequently asked questions about growth curve modeling. *Journal of cognition and development*, 11(2):121–136, 2010.

[13] Ian H De Boer, Sripal Bangalore, Athanase Benetos, Andrew M Davis, Erin D Michos, Paul Muntner, Peter Rossing, Sophia Zoungas, and George Bakris. Diabetes and hypertension: a position statement by the american diabetes association. *Diabetes Care*, 40(9):1273–1284, 2017.

[14] Bieke De Fraine, Georges Van Landeghem, Jan Van Damme, and Patrick Onghena. An analysis of wellbeing in secondary school with multilevel growth curve models and multilevel multivariate models. *Quality and Quantity*, 39(3):297–316, 2005.

[15] Pedro de Pablos-Velasco, Klaus G Parhofer, Clare Bradley, Eveline Eschwege, Linda Gönder-Frederick, Pierre Maheux, Ian Wood, and Dominique Simon. Current level

of glycaemic control and its associated factors in patients with type 2 diabetes across europe: data from the panorama study. *Clinical endocrinology*, 80(1):47–56, 2014.

[16] NK Elnaggar. Diabetes and hypertension. *J Hypertens*, 3(192):2167–1095, 2014.

[17] Centers for Disease Control, Prevention (CDC, et al. Racial/ethnic disparities in the awareness, treatment, and control of hypertension-united states, 2003-2010. *MMWR. Morbidity and mortality weekly report*, 62(18):351, 2013.

[18] Mohammad H Forouzanfar, Patrick Liu, Gregory A Roth, Marie Ng, Stan Biryukov, Laurie Marczak, Lily Alexander, Kara Estep, Kalkidan Hassen Abate, Tomi F Akinyemiju, et al. Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm hg, 1990-2015. *Jama*, 317(2):165–182, 2017.

[19] Esayas Kebede Gudina, Yadani Michael, and Sahilu Assegid. Prevalence of hypertension and its risk factors in southwest ethiopia: a hospital-based cross-sectional survey. *Integrated blood pressure control*, 6:111, 2013.

[20] Degefa Helamo, Romedan Delil, and Temesgen Dileba. Trends of diabetes mellitus and hypertension at nigist eleni mohammed general hospital, hossana, ethiopia (december 2010-january 2014): a five year longitudinal study. *Safety in Health*, 3(1):1, 2017.

[21] Joop J Hox, Mirjam Moerbeek, and Rens Van de Schoot. *Multilevel analysis: Techniques and applications*. Routledge, 2017.

[22] Alphonsus Rukevwe Isara and Patrick Otamere Okundia. The burden of hypertension and diabetes mellitus in rural communities in southern nigeria. *The Pan African medical journal*, 20, 2015.

[23] Maribeth Johnson. Individual growth analysis using proc mixed. *SAS User Group International*, 27, 2002.

[24] Tefera Kassahun, Tesfahun Eshetie, and Hailay Gesesew. Factors associated with glycemic control among adult patients with type 2 diabetes mellitus: a cross-sectional survey in ethiopia. *BMC research notes*, 9(1):78, 2016.

[25] Kelemu Tilahun Kibret and Yonatan Moges Mesfin. Prevalence of hypertension in ethiopia: a systematic meta-analysis. *Public Health Reviews*, 36(1):14, 2015.

[26] Yibeltal Kiflie, Challi Jira, and Dereje Nigussie. The quality of care provided to patients with chronic non-communicable diseases: A retrospective multi-setup study in jimma zone, southwest ethiopia. *Ethiopian journal of health sciences*, 21(2):119–130, 2011.

[27] Samuel Kingue, Constant Ndong Ngoe, Alain Patrick Menanga, Ahmadou M Jingi, Jean Jacques N Noubiap, Betrand Fesuh, Christophe Nouedoui, Gervais Andze, and Walinjom FT Muna. Prevalence and risk factors of hypertension in urban areas of cameroon: A nationwide population-based cross-sectional study. *The Journal of Clinical Hypertension*, 17(10):819–824, 2015.

[28] Hsu-Ko Kuo, Richard N Jones, William P Milberg, Sharon Tennstedt, Laura Talbot, John N Morris, and Lewis A Lipsitz. Effect of blood pressure and diabetes mellitus on cognitive and physical functions in older adults: a longitudinal analysis of the advanced cognitive training for independent and vital elderly cohort. *Journal of the American Geriatrics Society*, 53(7):1154–1161, 2005.

[29] Guido Lastra, Sofia Syed, L Romayne Kurukulasuriya, Camila Manrique, and James R Sowers. Type 2 diabetes mellitus and hypertension: an update. *Endocrinology and Metabolism Clinics*, 43(1):103–122, 2014.

[30] Ryan Lee, Tien Y Wong, and Charumathi Sabanayagam. Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss. *Eye and vision*, 2(1):17, 2015.

[31] Ramon C Littell, Jane Pendergast, and Ranjini Natarajan. Modelling covariance structure in the analysis of repeated measures data. *Statistics in medicine*, 19(13):1793–1819, 2000.

[32] Chun Luo, Yuting Ke, Yanyan Yuan, Ming Zhao, Fuyan Wang, Yisheng Zhang, and Shizhong Bu. A novel herbal treatment reduces depressive-like behaviors and increases brain-derived neurotrophic factor levels in the brain of type 2 diabetic rats. *Neuropsychiatric disease and treatment*, 12:3051, 2016.

[33] Cora JM Maas and Tom AB Snijders. The multilevel approach to repeated measures for complete and incomplete data. *Quality and Quantity*, 37(1):71–89, 2003.

[34] Jean Claude N Mbanya, Ayesha A Motala, Eugene Sobngwi, Felix K Assah, and Sostanie T Enoru. Diabetes in sub-saharan africa. *The lancet*, 375(9733):2254–2266, 2010.

[35] Lemba D Nshisso, Angela Reese, Bizu Gelaye, Sebelewengel Lemma, Yemane Berhane, and Michelle A Williams. Prevalence of hypertension and diabetes among ethiopian adults. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 6(1):36–41, 2012.

[36] K Ogurtsova, JD da Rocha Fernandes, Y Huang, U Linnenkamp, L Guariguata, NH Cho, D Cavan, JE Shaw, and LE Makaroff. Idf diabetes atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes research and clinical practice*, 128:40–50, 2017.

[37] Nasheeta Peer, Krisela Steyn, Carl Lombard, Nomonde Gwebushe, and Naomi Levitt. A high burden of hypertension in the urban black population of cape town: The cardiovascular risk in black south africans (cribsa) study. *PloS one*, 8(11):e78567, 2013.

[38] Cynthia Robitaille, Sulan Dai, Chris Waters, Lidia Loukine, Christina Bancej, Susan Quach, Joellyn Ellison, Norman Campbell, Karen Tu, Kim Reimer, et al. Diag-

nosed hypertension in canada: incidence, prevalence and associated mortality. *cmaj*, 184(1):E49–E56, 2012.

[39] Awol Seid, Muluye Getie, Belay Birlie, and Yehenew Getachew. Joint modeling of longitudinal cd4 cell counts and time-to-default from haart treatment: a comparison of separate and joint models. *Electronic Journal of Applied Statistical Analysis*, 7(2):292–314, 2014.

[40] Awoke Seyoum and Zewotir Temesgen. Joint longitudinal data analysis in detecting determinants of cd4 cell count change and adherence to highly active antiretroviral therapy at felege hiwot teaching and specialized hospital, north-west ethiopia (amhara region). *AIDS research and therapy*, 14(1):14, 2017.

[41] Andrea Siebenhofer, Klaus Jeitler, Karl Horvath, Andrea Berghold, Nicole Posch, Jutta Meschik, and Thomas Semlitsch. Long-term effects of weight-reducing drugs in people with hypertension. *Cochrane Database of Systematic Reviews*, (3), 2016.

[42] Judith D Singer. Fitting individual growth models using sas proc mixed. *Modeling intraindividual variability with repeated measures data: Methods and applications*, pages 135–170, 2002.

[43] T. A. Snijders and J. Berkhof. Diagnostic checks for multilevel models.in handbook of multilevel analysis:. pages 141–175, 2008.

[44] Fiona Steele. Multilevel models for longitudinal data. *Journal of the Royal Statistical Society: series A (statistics in society)*, 171(1):5–19, 2008.

[45] Barbara G Tabachnick, Linda S Fidell, and Jodie B Ullman. *Using multivariate statistics*, volume 5. Pearson Boston, MA, 2007.

[46] Henok Mulugeta Teshome, Getenet Dessie Ayalew, Fasil Wagnew Shiferaw, Cheru Tesema Leshargie, and Dube Jara Boneya. The prevalence of depression among diabetic patients in ethiopia: a systematic review and meta-analysis, 2018. *Depression research and treatment*, 2018, 2018.

[47] T Tiki. Prevalence and associated factors of depression among type 2 diabetes mellitus patients on follow up at ambo general hospital, oromia regional state, ethiopia, institutional based cross sectional study. *Journal of Depression and Anxiety*, 6(01):2167–1044, 2017.

[48] Jaakko Tuomilehto, Daiva Rastenyte, Qing Qiao, and Dimitrije Jakovljevic. Epidemiology of macrovascular disease and hypertension in diabetes mellitus. *International textbook of diabetes mellitus*, 2003.

[49] Steven van de Vijver, Hilda Akinyi, Samuel Oti, Ademola Olajide, Charles Agyemang, Isabella Aboderin, and Catherine Kyobutungi. Status report on hypertension in africa-consultative review for the 6th session of the african union conference of ministers of health on ncd's. *Pan African Medical Journal*, 16(1), 2014.

[50] Geert Verbeke and M Davidian. Joint models for longitudinal data: Introduction and overview. 2009.

[51] Geert Verbeke, Steffen Fieuws, Geert Molenberghs, and Marie Davidian. The analysis of multivariate longitudinal data: A review. *Statistical methods in medical research*, 23(1):42–59, 2014.

[52] Haileab Fekadu Wolde, Asrat Atsedeweyen, Addisu Jember, Tadesse Awoke, Malede Mequanent, Adino Tesfahun Tsegaye, and Shitaye Alemu. Predictors of vascular complications among type 2 diabetes mellitus patients at university of gondar referral hospital: a retrospective follow-up study. *BMC endocrine disorders*, 18(1):52, 2018.

[53] Demeke Lakew Workie, Dereje Tesfaye Zike, and Haile Mekonnen Fenta. Bivariate longitudinal data analysis: a case of hypertensive patients at felege hiwot referral hospital, bahir dar, ethiopia. *BMC research notes*, 10(1):722, 2017.

[54] Sung Sug Yoon, Qiuping Gu, Tatiana Nwankwo, Jacqueline D Wright, Yuling Hong, and Vicki Burt. Trends in blood pressure among adults with hypertension: United states, 2003 to 2012. *Hypertension*, 65(1):54–61, 2015.

## APPENDIX

**Appendix A: Explanatory variable**

| Variables from literatures | | | Variables from hospital | | |
|---|---|---|---|---|---|
| No | Variables | Value/Code | No | Variables | Value/Code |
| 1. | psychological stress | Stressed=0<br>Non-stressed=1 | 1. | Age | Continuous(year) |
| 2. | Lifestyle | Inactive =0<br>Active =1 | 2. | Gender | Female=0<br>Male=1 |
| 3. | anti-diabetics | Use=0<br>Not use=1 | 3. | Residence | Rural=0<br>Urban=1 |
| 4. | Adherence medication | Poor =0<br>Moderate =1<br>Good =2 | 4. | Presence of related disease | No disease =0<br>One disease=1<br>≥two disease=2 |
| 5. | physical activity min/day | Poor(<30)=0<br>Moderat(30-45)=1<br>Good (>45)=2 | 5. | stage of SBP | Pre-stage≤130 =0<br>stagone 140-150 =1<br>Stage two≥160=2 |
| 6. | history of glycemic in mm/hg | Low(<55)=0<br>med(56-69)=1<br>high>70 =2 | 6. | stage of DBP | Pre-stage≤ 89=0<br>Stage one 90-99=1<br>Stage two≥ 100=2 |
| 7. | prior gestational DM | Present =0<br>Absent =1 | 7. | Baseline DM | Continuous (mmol/L) |
| 8. | Cholesterol level | low =0<br>high =1 | 8. | prior HT history | Present=0<br>Absent=1 |
| 9. | Race/ Ethnicity | Black =0<br>White =1 | 9. | prior DM history | Present=0<br>Absent=1 |
| 10. | Chat chewing | Not chewing=0<br>chewing =1 | 10. | Time | Discrete |
| - | | | | | |

| | | |
|---|---|---|
| 11. | BMI | Continuous |
| 12. | Smoking exposure | Non-smoker=0<br>Smoker=1 |
| 13. | drinking coffee | Use=0<br>Not use=1 |

**Appendix D:**

$U_{1i}^{k} = MVN\,(0, \mathrm{G})$ Where, $\mathrm{G} = \Sigma_{k*k}$ variance-covariance matrix;(i.e., $u_{0i}$ to $u_{11i}$ are variance of intercept, number of related disease(no, one, and two or more), stage of SBP(pre, first,and second), and stage of DBP(pre, first,and second) respectively).

$$G = \begin{bmatrix} 13.99 & -5.76 & 2.37 & -0.98 & 0.40 & -0.17 & 0.07 & -0.03 & 0.01 & -0.01 \\ & 13.99 & -5.76 & 2.37 & -0.98 & 0.40 & -0.17 & 0.07 & -0.03 & 0.01 \\ & & 13.99 & -5.76 & 2.37 & -0.98 & 0.40 & -0.17 & 0.07 & -0.03 \\ & & & 13.99 & -5.76 & 2.37 & -0.98 & 0.40 & -0.17 & 0.07 \\ & & & & 13.99 & -5.76 & 2.37 & -0.98 & 0.40 & -0.17 \\ & & & & & 13.99 & -5.76 & 2.37 & -0.98 & 0.40 \\ & & & & & & 13.99 & -5.76 & 2.37 & -0.98 \\ & & & & & & & 13.99 & -5.76 & 2.37 \\ & & & & & & & & 13.99 & -5.76 \\ & & & & & & & & & 13.99 \end{bmatrix}$$

## Appendix B: Normality of the data

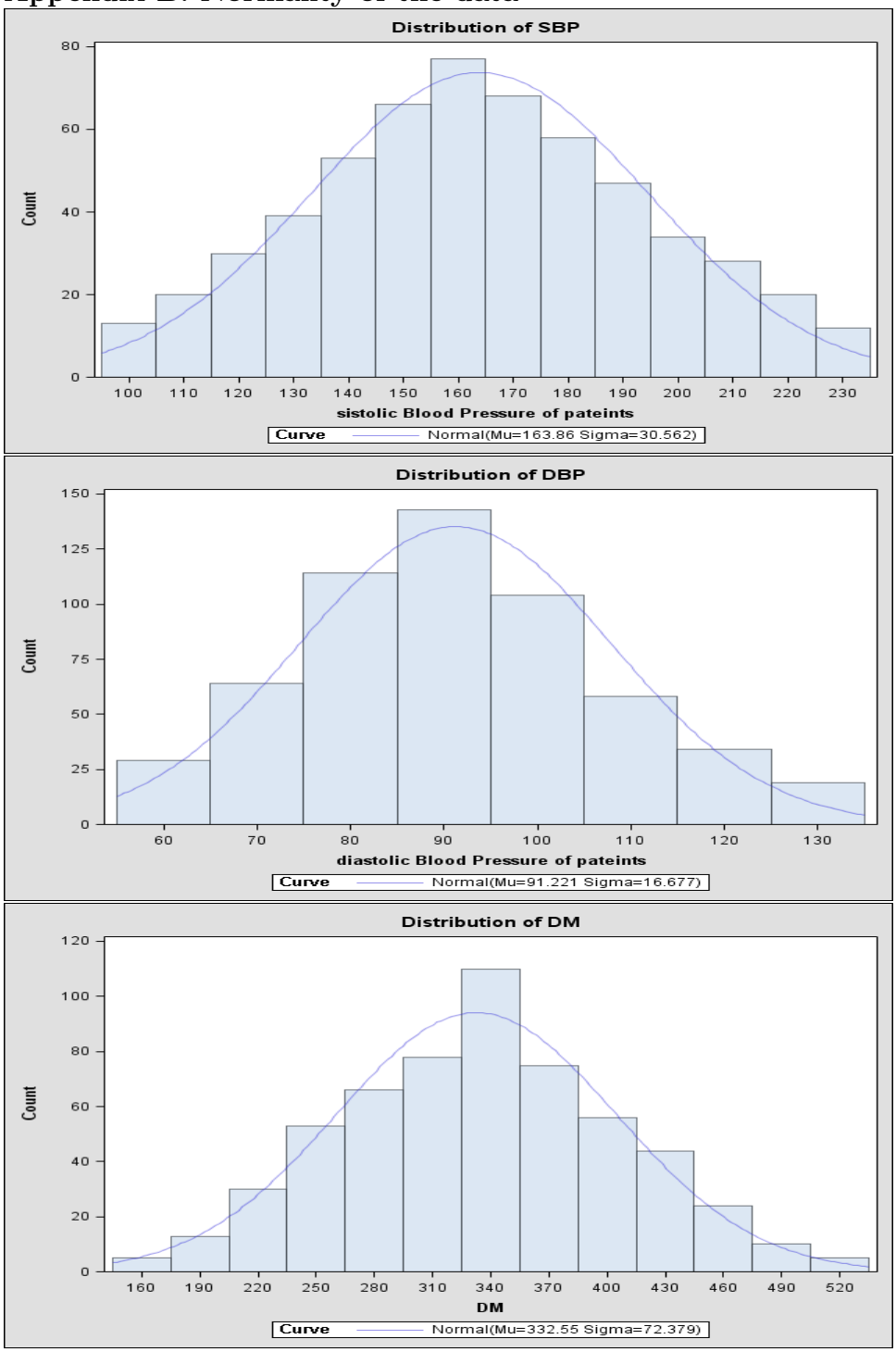



Figure 4: Sample data are comes from normal population

## Appendix C: mean plot of each response with their confidence interval

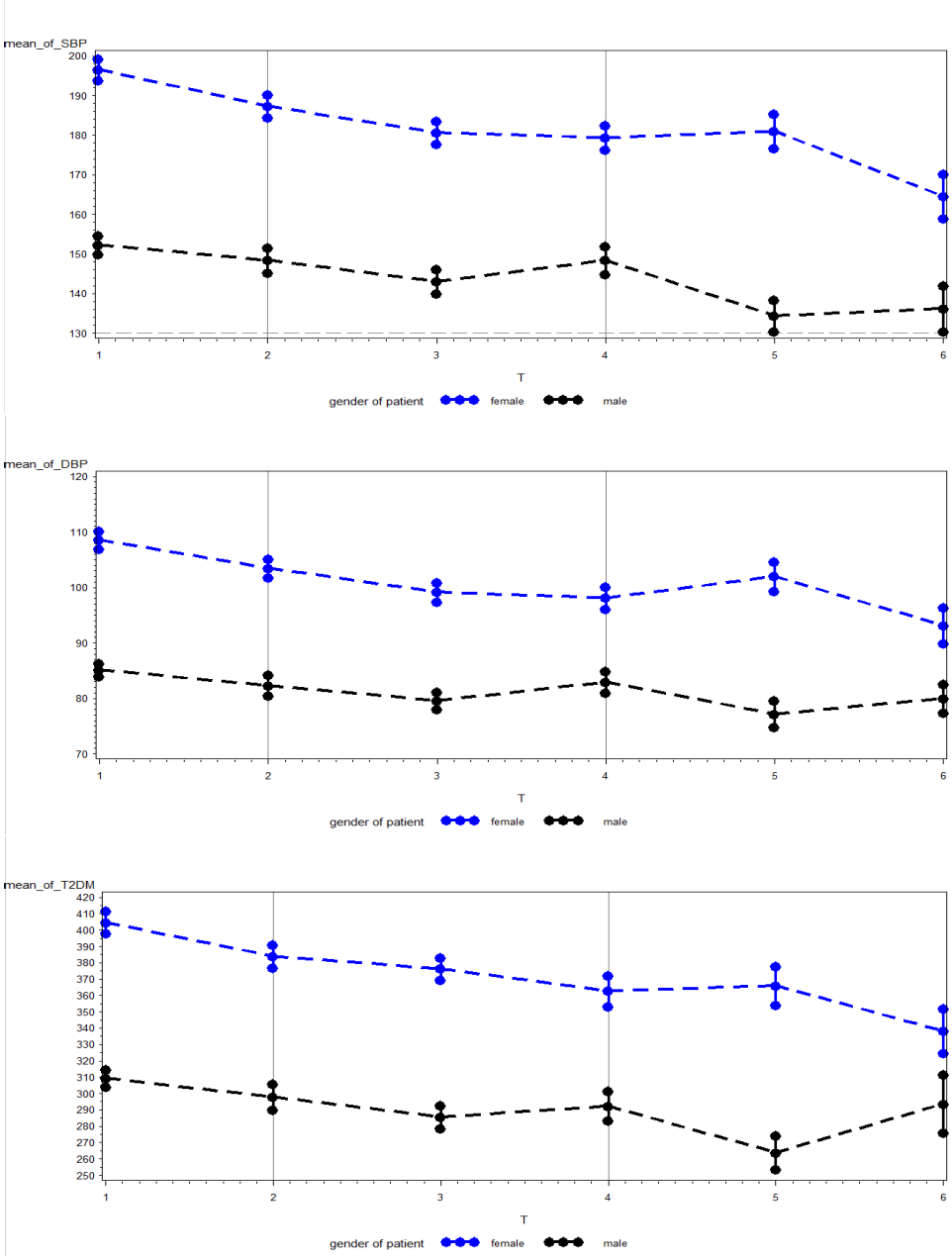



Figure 5: Mean plot of gender with their confidence interval for each response

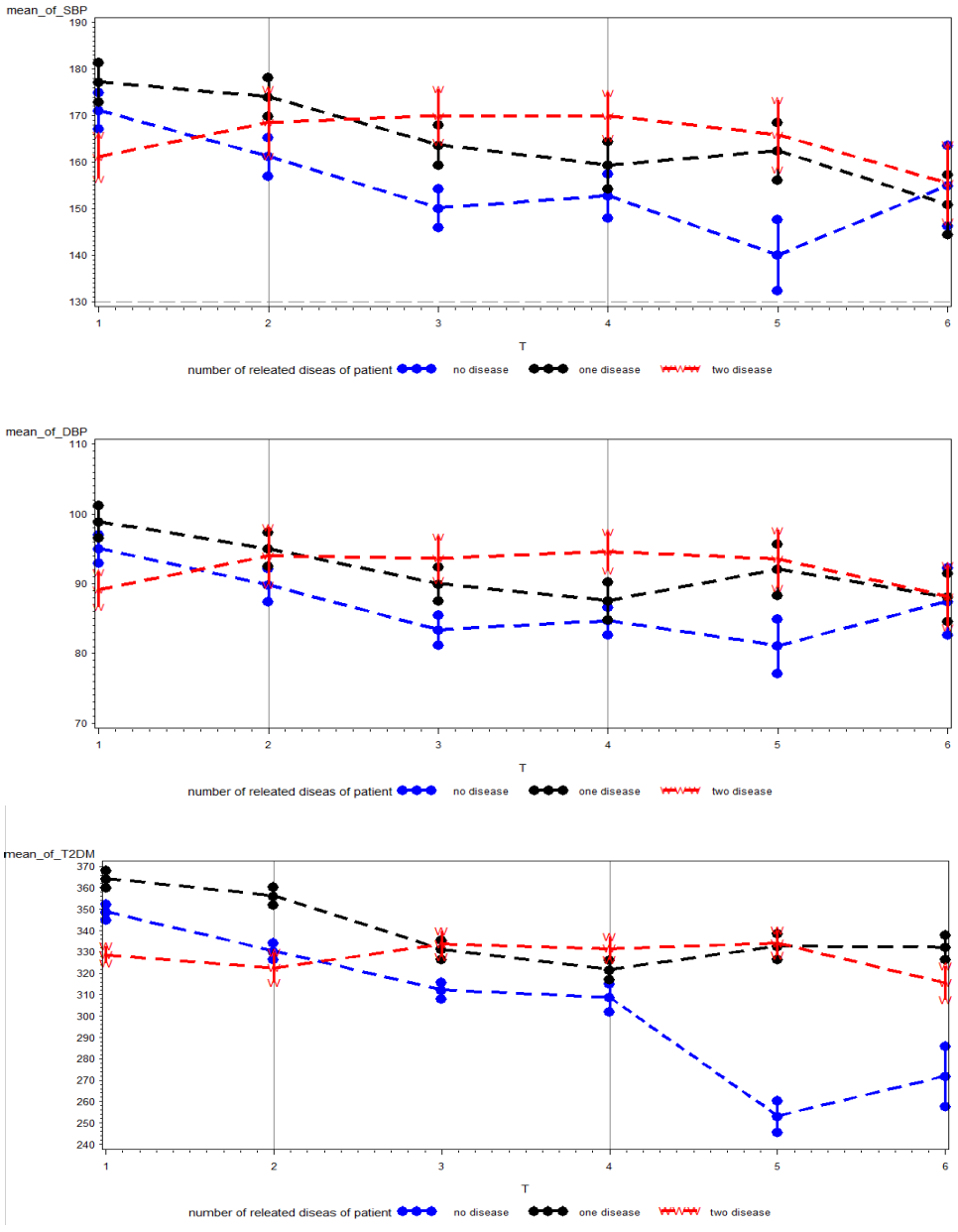



Figure 6: Mean plot of co-related disease with their confidence interval for each response

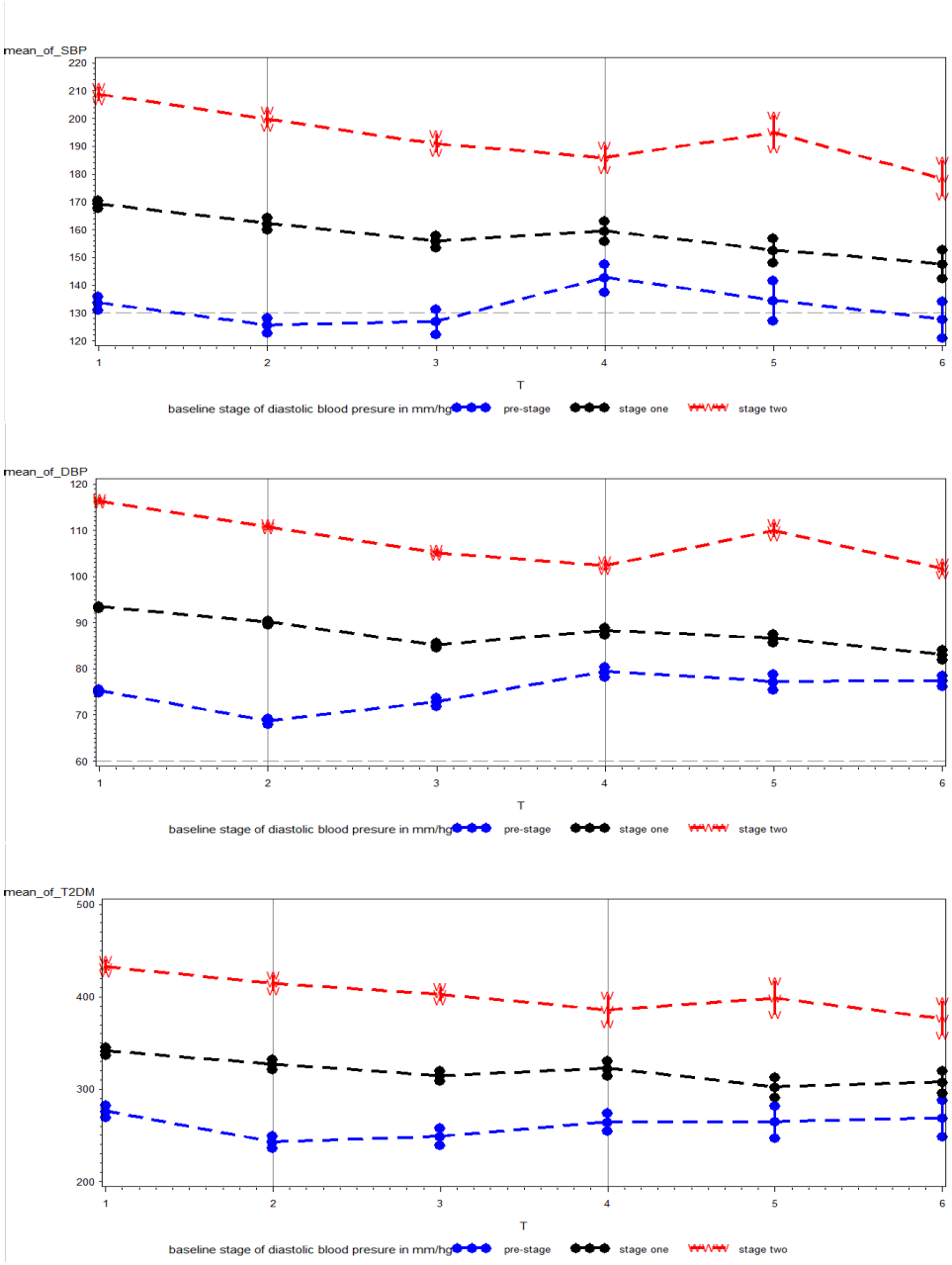



Figure 7: Mean plot of baseline stage of DBP with their confidence interval for each response