Longitudinal Analysis of Cardiovascular Disease Risk Factors based on the Framingham Heart Study

Group 1:
Emily Li, 2427297
Bryan Ng, 2427348
Junyi Fang, 2427308

Introduction

The Framingham Heart Study is an epidemiological study in cardiovascular research among a population of free living subjects in the community of Framingham, Massachusetts. The study aimed to identify the common factors or characteristics that contribute to cardiovascular disease (CVD). It is a large-scale cohort study to systematically collect data on health, lifestyle, and clinical measures from individuals in a community setting.

The data used here is a subset of this study, including 4,434 participants and 11,627 total observations from three examination periods spanning approximately 1956 to 1968.

These data include clinical, laboratory, questionnaire-based, and adjudicated cardiovascular event outcomes, such as myocardial infarction, stroke, and heart failure.

Clustering Structure and Variables

Clustering Structure

In this longitudinal study, the primary clustering structure is based on individuals. Each subject, identified by a unique identifier(RANDID), has 1 to 3 measurements taken at different time points.

Variable	Description	Units	Range
BMI	Body Mass Index	kg/m²	14.43–56.8
DIABP	Diastolic Blood Pressure	mmHg	30–150
SYSBP	Systolic Blood Pressure	mmHg	83.5–295
TOTCHOL	Total Cholesterol	mg/dL	107–696
GLUCOSE	Serum Glucose	mg/dL	39–478
HEARTRTE	Heart Rate	beats/min	37–220
LDLC / HDLC	Low / High Density Lipoprotein Cholesterol	mg/dL (available for period 3 only)	20-565 / 10-189
CURSMOKE	Current Smoker	0/1	Binary
BPMEDS	Use of Anti-hypertensive medication at exam	0/1	Binary
PREVSTRK	Prevalent Stroke	0/1	Binary
PREVMI	Prevalent Myocardial Infarction	0/1	Binary
PREVAP	Prevalent Angina Pectoris	0/1	Binary
PREVCHD	Prevalent Coronary Heart Disease	0/1	Binary
...	...	...	...

Specific Aim 1

Overall

We aim to investigate how demographic factors (Age, Sex, BMI) impact physiological measures such as Glucose, Blood Pressure, and Cholesterol. Additionally, we aim to examine how these physiological markers differ among individuals with or without a history of cardiovascular conditions, and to explore how Anti-hypertensive medication use and smoking behavior are associated with these risk factors and disease prevalence.

Question 1

Do variables such as Glucose (GLUCOSE), LDL Cholesterol (LDLC), HDL Cholesterol (HDLC), Systolic Blood Pressure (SYSBP), Diastolic Blood Pressure (DIABP), and Heart Rate (HEARTRTE) differ by demographic factors including Sex (SEX), Age (AGE), and Body Mass Index (BMI) while adjusting for Education (EDUC) and Time (TIME)?

Associated Variables

Clustering Variable: Participant ID (RANDID)
Exposure Variables: Sex (SEX), Age (AGE), Body Mass Index (BMI)
Covariates: Education (EDUC), Time (TIME)
Outcome Variables: Glucose (GLUCOSE), LDL Cholesterol (LDLC), HDL Cholesterol (HDLC), Systolic Blood Pressure (SYSBP), Diastolic Blood Pressure (DIABP), Heart Rate (HEARTRTE)

Specific Aim 2

Question 2

Are elevated levels of physiological markers such as Glucose (GLUCOSE), LDL Cholesterol (LDLC), HDL Cholesterol (HDLC), Systolic Blood Pressure (SYSBP), Diastolic Blood Pressure (DIABP), and Heart Rate (HEARTRTE) associated with a higher prevalence of cardiovascular conditions such as Prevalent Stroke (PREVSTRK), Prevalent Myocardial Infarction (PREVMI), Prevalent Coronary Heart Disease (PREVCHD), and Prevalent Angina Pectoris (PREVAP) at the time of examination?

Associated Variables

Clustering Variable: Participant ID (RANDID)
Exposure Variables: Glucose (GLUCOSE), LDL Cholesterol (LDLC), HDL Cholesterol (HDLC), Systolic Blood Pressure (SYSBP), Diastolic Blood Pressure (DIABP), Heart Rate (HEARTRTE)
Covariates: Sex (SEX), Age (AGE), Time(TIME), Body Mass Index (BMI)
Outcome Variables: Prevalent Stroke (PREVSTRK), Prevalent Myocardial Infarction (PREVMI), Prevalent Coronary Heart Disease (PREVCHD), Prevalent Angina Pectoris (PREVAP)

Specific Aims 3 and 4

Question 3

What is the effect of Anti-hypertensive Medication (BPMEDS) on Systolic Blood Pressure (SYSBP) and Diastolic Blood Pressure (DIABP)?

Associated Variables

Clustering Variable: Participant ID (RANDID)
Exposure Variable: Anti-hypertensive Medication Use (BPMEDS)
Covariates: Sex (SEX), Age (AGE), Time(TIME), Body Mass Index (BMI)
Outcome Variables: Systolic Blood Pressure (SYSBP), Diastolic Blood Pressure (DIABP)

Question 4

What is the effect of Smoking Status (CURSMOKE) and Smoking Intensity (CIGPDAY) on Glucose (GLUCOSE), LDL Cholesterol (LDLC), HDL Cholesterol (HDLC), Systolic Blood Pressure (SYSBP), Diastolic Blood Pressure (DIABP), and Heart Rate (HEARTRTE)?

Associated Variables

Clustering Variable: Participant ID (RANDID)
Exposure Variables: Smoking Status (CURSMOKE), Cigarettes per Day (CIGPDAY)
Covariates: Sex (SEX), Age (AGE), Time(TIME), Body Mass Index (BMI)
Outcome Variables: Glucose (GLUCOSE), LDL Cholesterol (LDLC), HDL Cholesterol (HDLC), Systolic Blood Pressure (SYSBP), Diastolic Blood Pressure (DIABP), Heart Rate (HEARTRTE)