의학박사 학위논문

Cardiovascular risk assessment and machine learning prediction model of metabolic comorbidity

대사 질환 동시 이환의 심혈관계 질환 위험 평가 및 기계학습 예측 모형 개발

2022년 8월

서울대학교 대학원

의과학과 의과학전공

안 서 경

Cardiovascular risk assessment and machine learning prediction model of metabolic comorbidity

지도 교수 박 수 경

이 논문을 의학박사 학위논문으로 제출함
2022년 4월

서울대학교 대학원
의과학과 의과학전공
안 서 경

안서경의 의학박사 학위논문을 인준함
2022년 7월

위 원 장 박 상 민 (인)

부위원장 박 수 경 (인)

위 원 이 해 영 (인)

위 원 고 광 필 (인)

위 원 박 보 영 (인)

Abstract

Cardiovascular risk assessment and machine learning prediction model of metabolic comorbidity

Seokyung An
Biomedical Science
The Graduate School
Seoul National University

Introduction: The growing aging population and westernized lifestyle have increased the prevalence of disease comorbidity, which is defined as having more than two metabolic diseases including hypertension (HTN), diabetes mellitus (DM), dyslipidemia (LIP), obesity, and metabolic syndrome (MetS). The combination of these diseases is related to an increased risk of cardiovascular disease (CVD) outcomes. The Global Burden of Disease 2016 Study reported that CVD are by far the leading cause of death globally and one of the major health challenges of the 21st century. In Korea, CVD is the second largest cause of death following cancer.

As those diseases share risk factors, the World Health Organization (WHO) designated healthy lifestyle, including alcohol reduction, weight loss, smoking cessation, physical activity, and healthy diet, as modifiable factors of CVDs. Thus, it is necessary to estimate the amount of comorbidity prevalence, identify the combined association of metabolic comorbidity and other risk factors (family history of CVD and lifestyle factors) with CVD outcomes, and develop predictive model for comorbidity for detecting the high-risk of metabolic comorbidity and preventing the future risk of CVD through intervention strategies.

Methods: This study mainly used population-based cohort study from the Korea Genome and Epidemiology Study (KoGES) including Health Examinee-Gem study (HEXA), cardiovascular disease association study (CAVAS), and Ansan and Ansung Study from 2001-2014, in addition to United States (US) National Health and Nutrition Examination Survey 2003-2014 (NHANES), Korea NHANES (KNHAENS) 2007-2014, and Asia Cohort Consortium (ACC) study.

For the statistical analyses, direct standardization methods using the WHO world standard population was performed to estimate the age-standardized prevalence of metabolic diseases. The baseline characteristics were compared using Chi-squared test for categorical variables and Student's t-test for continuous variables. Cox proportional hazards regression analysis was performed to estimate hazard ratios (HRs) with 95% confidence intervals (CIs) of CVD outcomes. To calculate the odds ratios (ORs) of metabolic diseases, logistic regression models were used. For prediction model, cox proportional hazard regression, and random survival forest (RSF) models were developed in the training set (70% of the total population) and performance evaluations of each model were performed in the test set (30% of the total population) with concordance statistics (c−index). For self−assessed biological age (BA) prediction model, elastic net regression analysis with 10−fold cross validation was performed.

Results: According to the comparison of the prevalence of metabolic disease and comorbidity in Korea and the US, Korea had a lower prevalence of metabolic comorbidity than the US. In both Korean and the US population, the most common combination was HTN and obesity. Among the Korean population, individuals living in rural areas had the higher comorbidity prevalence than those who lived in urban areas.

In the association between metabolic comorbidity, family history of CVD, and the risk of CVD study, we found that individuals with DM, HTN, LIP, and a positive family history of CVD had a 2.88−fold increased risk of CVD, a 3.30−fold increased risk of MI, and a 2.52−fold increased risk of stroke compared to the individuals with a negative family history of CVD and none of metabolic diseases.

In the impact of lifestyle factors with cardiometabolic disease (CMDs) such as HTN, DM, coronary heart disease (CHD), and stroke on CVD death study, the healthy lifestyle status was defined as 'never smoker', 'never drinker', and 'body mass index (BMI) 18.5–27.4kg/m^2' in Asian population. Among the lifestyle factors, non−smoking had the strongest association with decreasing risk of all cause and CVD death among the healthy lifestyle factors. A significant association of healthy lifestyle score with lower CVD death was observed among individuals with HTN, DM, and CHD (HR 0.76, 95% CI: 0.63−0.93). For individuals with cardiometabolic comorbidity, having three of healthy lifestyle factors was significantly associated with decrease in CVD (HR 0.51, 95% CI: 0.42−0.61) and premature CVD death (HR 0.38, 95% CI: 0.27−0.54).

Based on the repeated measurements for assessing change in lifestyle factors study, unhealthy lifestyle modification including increased dose of cigarette smoking (HR 1.49, 95% CI: 1.09−2.03) and increased their intensity of consumption from light/moderate to heavy had a significantly increased risk for MetS (HR 1.42, 95% CI: 1.10−1.84). For obesity, individuals who newly became obesity had a significant increase in risk for MetS (HR 1.88, 95% CI: 1.44−2.45).

For improving the individualized health status, we developed machine learning–based disease prediction model and self–assessed BA as a predictor for metabolic comorbidity. We found that compared to the individuals in same BA as chronological age (CA) group, those in younger BA than CA group were associated with a decreased risk of DM (HR = 0.63, 95% CI: 0.55–0.72), HTN (HR = 0.74, 95% CI: 0.68–0.81), and combination of HTN and DM (HR = 0.65, 95% CI: 0.47–0.91). For machine learning–based disease prediction model study, predictive models achieved a high discriminatory ability for comorbidity of HTN and DM.

Conclusions: This study highlights the necessity of accounting to metabolic comorbidity to reduce the future risk of CVD outcomes in Korean population. Although individuals already have had cardiometabolic comorbidity, healthy lifestyles (smoking cessation, abstaining from alcohol, and maintaining BMI) are effective to reduce the further risk of CVD death. Moreover, lifestyle changes help to decrease the risk of a cluster of metabolic conditions. At last, machine learning–based self–assessed BA and disease prediction model may be an effective indicator for identifying the high–risk group and decreasing burden of metabolic comorbidities in Korea through prevention.

Keywords: metabolic comorbidity, lifestyle prevention, cardiovascular disease, biological age, prediction model

Student number: 2016-21993

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

HTN	Hypertension
DM	Diabetes mellitus
LIP	Dyslipidemia
CVDs	Cardiovascular diseases
US	United states
MetS	Metabolic syndrome
MI	Myocardial infarction
WHO	World health organization
CMDs	Cardiometabolic diseases
BMI	Body mass index
CA	Chronological age
SES	Socioeconomic status
BA	Biological age
RSF	Random survival forest
NHANES	National Health and Nutrition Examination Survey
KNHANES	Korea National Health and Nutrition Examination Survey
KoGES	Korean Genome and Epidemiology Study
HEXA	Health−examinees study
CAVAS	Cardiovascular disease association study

HDL	High−density lipoprotein
TG	Triglyceride
IRS	Institutional Review Boards
ACC/AHA	American College of Cardiology and American Heart Association
IDF	International Diabetes Federation
NCEP−ATP−III	National Cholesterol Education Program’s Adult Treatment Panel III
ASR	Age−standardized rate
HRs	Hazard ratios
CIs	Confidence intervals
WHR	Waist−to−hip ratio
ACC	Asian Cohort Consortium
CHD	Coronary heart disease
HLS	Healthy lifestyle score
SD	Standard deviation
Age−Diff	Age−difference
VIF	Variance inflation factor
RF	Random forest
C−index	Concordance index
CoxPH	Cox proportional hazard model

I. Introduction

1.1. Background

As the national life span increases, the prevalence rate of metabolic disease and comorbidity is accelerating globally [1-4]. Metabolic comorbidity, which is defined as having more than two metabolic diseases including hypertension (HTN), diabetes mellitus (DM), dyslipidemia (LIP) and obesity is the major risk factors of cardiovascular diseases (CVDs)[5]. In the United States (US), nearly a half of adults have heart diseases, and around one in every four people die from it.[6] In Korea, CVD is the second largest cause of death following cancer.[7]

Lifestyle factors such as cigarette smoking, alcohol consumption, obesity, dietary intake, and exercise are widely known as a risk factor for high blood pressure, fasting glucose levels, cholesterol levels, and metabolic syndrome (MetS), all of which impact on CVDs.[8-12] In Korea, the growing westernized lifestyle and aging population were associated with increasing the prevalence of metabolic disease.[13-17]

Prevalence of metabolic disease and comorbidity

Estimating the prevalence of metabolic disease and comorbidities in the nationally representative population can be an key indicator for forecasting future CVD risk. Moreover, considering such dissimilarities in Korea and the US, it is important the understand the differences in prevalence of metabolic comorbidity between the two countries in order to better interpret data from the US and apply the implications to the Korean population.

Several studies have investigated the prevalence of metabolic disease in each country,[18, 19] but no previous study has compared the prevalence of metabolic comorbidities between two countries based on the nationally representative population dataset. Moreover, there have been few population-based studies in Korea that estimate the comorbidity prevalence between urban and rural areas. Therefore, estimation of metabolic disease and comorbidities prevalence is necessary to measure the burden of disease and suggest the future preventive strategies to mitigate the risk of CVD outcomes in Korea.

Metabolic comorbidity, family history of CVD, and the risk of CVD

Metabolic comorbidities are the major risk factors of CVD including myocardial infarction (MI) and stroke, which is the leading cause of death [20-22]. The risk of MI in diabetic patients with high blood pressure has been reported to be more than 2-fold higher than that in patients without these conditions [23]. In another study, patients with a prevalence of DM and LIP had a 1.3-fold increase in CVD risk [24]. Another remarkable risk factor of CVD is a family history of CVD [25]. Moreover, having a positive family history of CVD is associated with a higher prevalence of metabolic disease [26]. With the aging population, the prevalence of metabolic comorbidity is constantly increasing, and a continued increase in CVD is inevitable [22, 27, 28]. The relationship between metabolic comorbidities and CVD may differ depending on the family history of CVD. Previous studies have found a relationship between metabolic diseases and an increased risk of CVD events and mortality [23, 24, 29, 30]. However, evidence regarding the risk of CVD incident among patients with metabolic comorbidities among people with family history of CVD is lacking.

Cardiometabolic comorbidity, lifestyle factors, and the risk of CVD death

The World Health Organization (WHO) has announced healthy lifestyle guidelines such as smoking cessation, the reduction of alcohol intake and body weight, sufficient regular exercise, and a healthy diet [31], as modifiable factors for risk reduction, prevention, and treatment of cardiometabolic diseases (CMDs) [32]. Moreover, a previous cohort study suggested that having multiple healthy lifestyle factors may be significant to decrease the risk of CVD [33] and CVD−related death than adherence to only one healthy lifestyle factor [34]. For alcohol consumption, observational studies reported the J−shaped curve in alcohol consumption, however, prospective, and clinical trials showing that light to moderate alcohol consumption was beneficial were lacking [35]. Moreover, the guidelines from primary prevention of stroke are not advised to begin drinking due to the alcohol dependency [35]. It is well known that there is different association of body mass index (BMI) with healthy outcomes between Asian and European populations. According to the WHO [36], the 3 categories including 18.5−22.9 (normal weight), 23−27.4 (overweight), and 27.5+ (obesity) were suggested for Asian population. However, the study stated that there the available data were not sufficient to conclude Asian−specific cut points. The optimal BMI range associated with a reduced risk of death in Asian population remains controversial. Based on the more than 1 million Asian population study *[37]*, the lowest risk of death was seen among Asian population with a BMI in the range of 22.6−27.4.

Based on this association, many studies have investigated the impact of the number of healthy lifestyle factors on death. However, these studies were also limited in that they tended to examine the association between lifestyle factors and death due to CVDs in a healthy population without a significant past medical history[34, 38–41]. Only a few studies have assessed the association between healthy lifestyles and life expectancy in the Western populations with chronic diseases[42, 43]. However, the impact of healthy lifestyle factors and CMDs comorbidity on CVD death in Asian population remains unclear. Moreover, no previous study has investigated whether multiple healthy lifestyle factors significantly lower the risk of CVD-specific death in patients with varying combination of CMDs.

Therefore, the study to find the impact of healthy lifestyle factors and varying combination of CMDs on CVD-specific death in Asian population is needed.

Change in lifestyle factors and metabolic syndrome

MetS is a combination of metabolic disorders including high blood pressure, fasting glucose level, cholesterol level, and obesity [44]. Nearly a quarter of World population have MetS and the prevalence of MetS is continuously increasing [45, 46]. As MetS is associated with increasing risk of DM, CVD, and mortality, it has become a major concern in public health [47].

Lifestyle modification including non-cigarette smoking, limited alcohol consumption, regular physical activity, and weight loss is primary approach of prevention and management of MetS. The previous studies suggested that baseline healthy lifestyle factors are related to decrease the risk of MetS. However, whether changes in lifestyle factors is associated with MetS risk is unclear. Thus, it is necessary to identify the effects of changes in lifestyle factors that may influence the risk of MetS for decreasing the burden of disease. In this study, we investigated the association of changes/trajectories in lifestyle factors (dose of cigarette smoking, dose of alcohol consumption, physical activity, and BMI) with HTN, DM, and MetS in Korean population.

Biological age

Although chronological age (CA) is an key risk factor for metabolic diseases, the impact of CA on diseases may differ according to the body composition, socioeconomic status (SES) and lifestyle behaviors [48-50]. There are changes in body composition occur with aging, and these changes can have an effect on disease [50]. Differences in aging have also been reported based on the socioeconomic status [48]. Based on this difference, individuals with the same CA may have different biological ages (BA). Thus, BA, which is calculated based on aging-related factors, has been considered as a more precise predictor for indicating disease risk compared to the CA [51-54].

Previous studies of BA have been conducted based on clinical information such as laboratory blood tests, frailty-related physical factors, physiological factors, metabolomics, and deoxyribonucleic acid-methylation [51, 55-57]. The BA models based on these markers were useful to consider the biological mechanism of aging, however, were inflexible in lifestyle recommendations and interventions to manage health status. Moreover, only two studies assessed the BA as a predictor for the risk of metabolic disease [58, 59] and no study for comorbidity. Thus, the necessity for an individualized self-assessed BA model for metabolic comorbidity is emphasized to improve health management.

Disease prediction model

Nearly half of Korean population aged over 40 years had HTN [44]. And the prevalence of HTN in DM increased in Korean adults aged over 30 years [60]. Comorbidity of HTN and DM is a risk factor that increased the risk of CVDs and death [61, 62], which contributed to immense health and economic burdens in Korea [63–65]. Therefore, it is essential to provide practical model to help early-detection of these conditions in order to decrease the risk of further multimorbidity and premature death. Previous studies have used several machine learning algorithms for analyzing time-series data to predict DM, CVD, and mortality risk, respectively [66–69]. However, the evidence on machine learning approaches for predicting the metabolic comorbidity is limited. Developing a machine learning-based prediction models for HTN and DM comorbidity using the common risk factors is necessary to detect high-risk groups. Therefore, we aimed to identify the risk factors for HTN and DM, develop a predictive model predicting HTN and DM simultaneously, and evaluate the predictive performance of the models.

1.2. Objectives

The principle aim of this study was to find the combined association of metabolic comorbidity and other risk factors (family history of CVD and lifestyle factors) with CVD outcome and develop prediction models for comorbidity based on the machine learning approaches in Korean population.

In detail, study objectives related to 1) estimate and compare the prevalence of metabolic disease and comorbidity in Korea and the US; 2) assess the risk of CVD in relation to metabolic comorbidity and family history of CVD; 3) evaluate the impact of lifestyle factors and cardiometabolic comorbidity on CVD-specific death; 4) find the association of change in lifestyle factors with metabolic syndrome; 5) develop a machine learning-based biological age and prediction models for metabolic comorbidity (Figure 1). To achieve the goals, the following eight hypothesis were tested in this study:

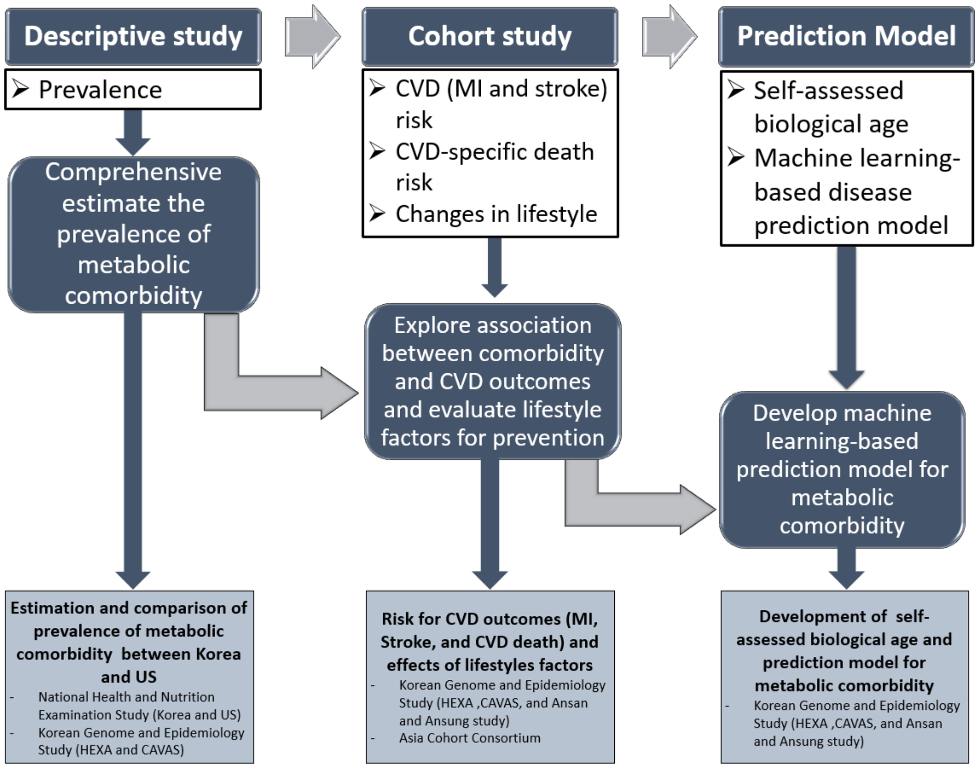


Figure 1. Overview of study objectives

1.3. Hypothesis

<u>Metabolic comorbidity, history of CVD, and the risk of CVD</u>

Hypothesis 1: Metabolic comorbidity is associated with CVD (MI and stroke) risk.

Hypothesis 2: Individuals with metabolic comorbidity and a positive family history of CVD have increased risk of CVD (MI and stroke).

<u>Cardiometabolic comorbidity, lifestyle factors, and the risk of CVD death</u>

Hypothesis 3: Each combination of lifestyle factors (cigarette smoking, alcohol drinking, and obesity) has different impact on CVD death risk in Asian population.

Hypothesis 4: The impact of healthy lifestyle factors on CVD death is different according to the combination of CMDs.

<u>Change in lifestyle factors and metabolic syndrome</u>

Hypothesis 5: Healthy lifestyle changes (reduced dose of cigarette smoking, alcohol consumption, regular physical activity, and weight loss) can benefit from reduced risk of MetS.

<u>Prediction models for metabolic comorbidity</u>

Hypothesis 6: Self-assessed BA based is associated with metabolic comorbidity.

Hypothesis 7: Disease predictive models using statistical and machine learning approaches can predict DM and HTN comorbidity (Table 1 & Figure 2).

Table 1. Overview of study hypothesis

Study	Data	Hypothesis
1. Prevalence	USNHANES, KNHANES, and KoGES	
2. Family CVD	KoGES	1. Metabolic comorbidity → CVD ↑ 2. Family CVD + comorbidity → CVD ↑
3. CVD death	ACC	3. HLS ↑ → CVD death ↓ 4. CMDs + HLS ↑ → CVD death ↓
4. Change in lifestyle	KoGES	5. Healthy lifestyle change → MetS.↓
5. Biological age	KoGES	6. BA < CA → comorbidity ↓
6. Prediction model	KoGES	7. Prediction of disease comorbidity

Abbreviation: NHANES: National Health and Nutrition Examination Survey; KoGES: Korean Genome and Epidemiology Study; CVD: Cardiovascular disease; ACC: Asian Cohort Consortium; CMDs: Cardiometabolic diseases; HLS: healthy lifestyle score; MetS: Metabolic syndrome; BA: biological age; CA: chronological age;

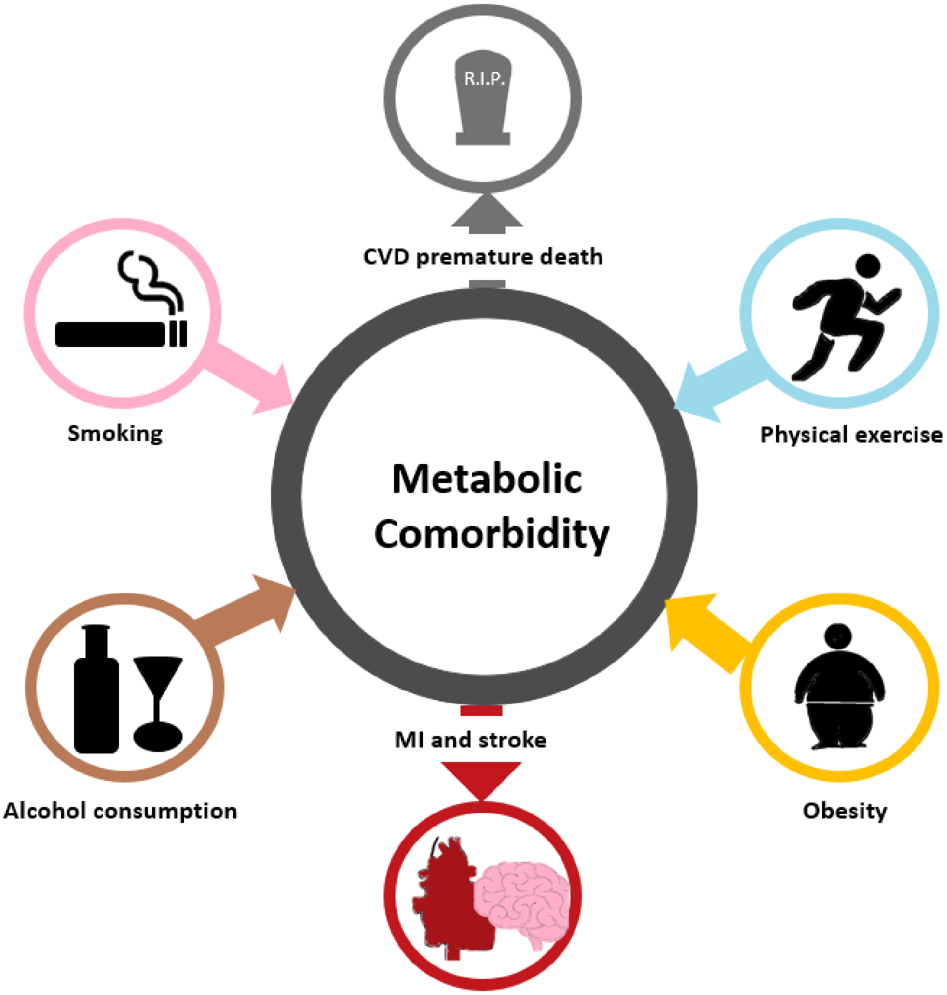


Figure 2. Overview of metabolic comorbidity mechanisms

II. Materials and methods

2.1. Data source

2.1.1. National Health and Nutrition Examination Survey

US National Health and Nutrition Examination Survey (US NHANES)

The NHANES is a series of multistage probability surveys designed to be representative of the non-institutionalized population in the US [70]. Since 1999, the NHANES has been collecting data in 2-year phases. In this study, participants recruited between 2003 to 2014 were used.

Korea National Health and Nutrition Examination Survey (KNHANES)

The KNHANES is a nationally representative cross-sectional survey that collects data on demographic status, lifestyle habits, anthropometric measurements, and clinical profiles [71]. The data is collected annually through a health questionnaire and examination done by certified physicians and medical technicians. Individuals recruited between 2007 to 2014 were included in this study.

2.1.2. Korean Genome and Epidemiology Study (KoGES)

This study was based on population-based cohorts from the KoGES, including the health examinees study (HEXA) from 2004 to 2017, the cardiovascular disease association study (CAVAS) from 2005 to 2014, and the Ansan and Ansung study from 2001 to 2014. This cohort consisted of participants recruited from the National Health Examinee Registry, and including data on demographics, health examinations, laboratory blood tests, and disease diagnoses obtained by trained interviewers. The detailed information of the KoGES is described previous studies [72, 73].

2.1.3. Asia Cohort Consortium (ACC)

The Asian cohort consortium (ACC) is a cooperative study design involving several cohort studies from multiple countries in Asia. The dataset has the advantage of being able to use the pooling of raw data to prove hypotheses of small effect size [37]. It contains information on demographic variables, lifestyle behaviors, and disease history. The database was integrated based on the structured questionnaires and managed by the ACC coordinating center. More details of the ACC and its study framework are described elsewhere[37, 74-76].

2.2. Study population

2.2.1. Prevalence study

The eligible criteria were those with who were (i) 40 to 69 years old; (ii) having the information on body measurements, blood pressure measurements, blood tests (including fasting glucose level, HbA1C, total cholesterol, high-density lipoprotein (HDL) cholesterol, triglyceride (TG)), and use of antihypertensive and antidiabetic drugs to define the metabolic prevalence status. Based on the inclusion criteria, this study is eligible to 15,872 individuals from US NHANES, 26,492 individuals from KNHANES, 139,345 participants from HEXA, and 24,994 participants from CAVAS. The entire study protocol was approved by the Institutional Review Boards (IRB) of Seoul National University Hospital (Approval No. 0608-018-179 and 1912-063-1088). Informed consent was confirmed by the IRB.

2.2.2. Family history of CVD and the risk of CVD study

Among 211,721 adults aged 40-89 years who had undergone health examinations from KoGES integrated data, we initially included participants who had received at least two health examinations. We excluded the participants lost to follow-up and had lack of information on family history of CVD, metabolic comorbidity at baseline, and age of onset of MI or stroke. Individuals with a history of MI or stroke at the baseline were further excluded. Finally, a total of 72,111 participants were included in the study (Figure 3). The study protocol was approved by the IRB of the Seoul National University (No. 1912−063−1088).

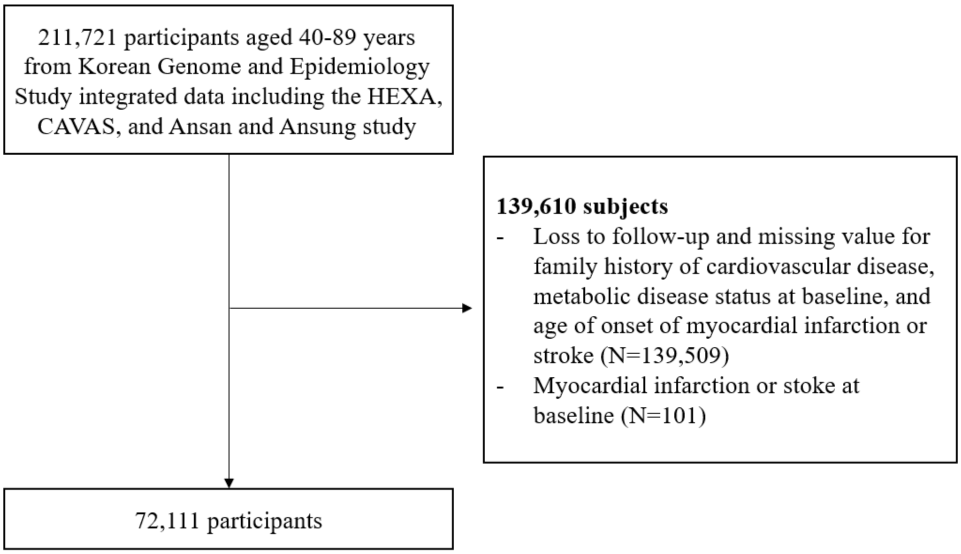


Figure 3. Flow chart of the study population selection from the Korean Genome and Epidemiology Study

2.2.3. Lifestyle factor and the risk of CVD death study

Of the 619,518 eligible study participants from ACC, 135,247 participants with missing value on age, sex, follow−up time, and lifestyle factors including cigarette smoking, alcohol drinking, and body mass index at baseline were excluded. These were classified as "healthy lifestyle factors" for the study. 80,419 participants without past medical history of CMDs including hypertension, DM, coronary heart disease (CHD), and stroke were also excluded from the study. The final study population of 403,852 participants aged over 18 years were included in this study (Figure 4).

This study was approved by the IRB of the ACC coordinating center (National Cancer Center, Tokyo, Japan) (approval no. 2014-041) and Seoul National University Hospital (approval no. H-0110-084-002 and H-0901-040-269). The requirement for informed consent from the participants was waived by ACC coordinating center according to confidentiality guidelines.

Figure 4. Flow chart of the study population selection from the Asian Cohort Consortium

2.2.4. Change in lifestyles study

The Ansan and Ansung study is a prospective cohort study, which conducted 6th biannual repeated survey since baseline recruitment enrolled between 2001 and 2003 in Korea (Figure 5). Study design for assessing the association of lifestyle trajectories over time with MetS was shown in Appendix 1.

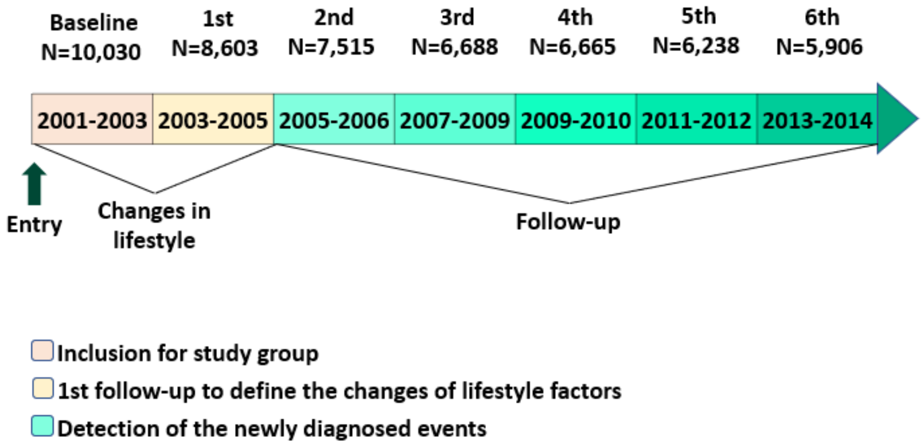


Figure 5. Flow chart of basline entry and follow-up for the Ansan and Ansung cohort study.

Among the 8,603 participants at the 1st follow-up, we excluded 286 participants those who with no information on lifestyle factors (smoking, alcohol drinking, physical activity, BMI, and waist circumstance) and 742 participants who were lost to follow-up from 2005 to 2016. Among a total of 7,575 participants, we included 4,638 participants without HTN at the baseline and 2nd follow-up period for the analysis of HTN risk, 6,709 participants for the analysis of DM risk, and 3,292 participants for the analysis of MetS risk (Figure 6). For lifestyle trajectories, a total of 3,888 participants for HTN, 5,930 for DM, and 2,683 for MetS were included (Appendix 2). The study protocol was reviewed and approved by the IRB of the Seoul National University (No. 1912-063-1088).

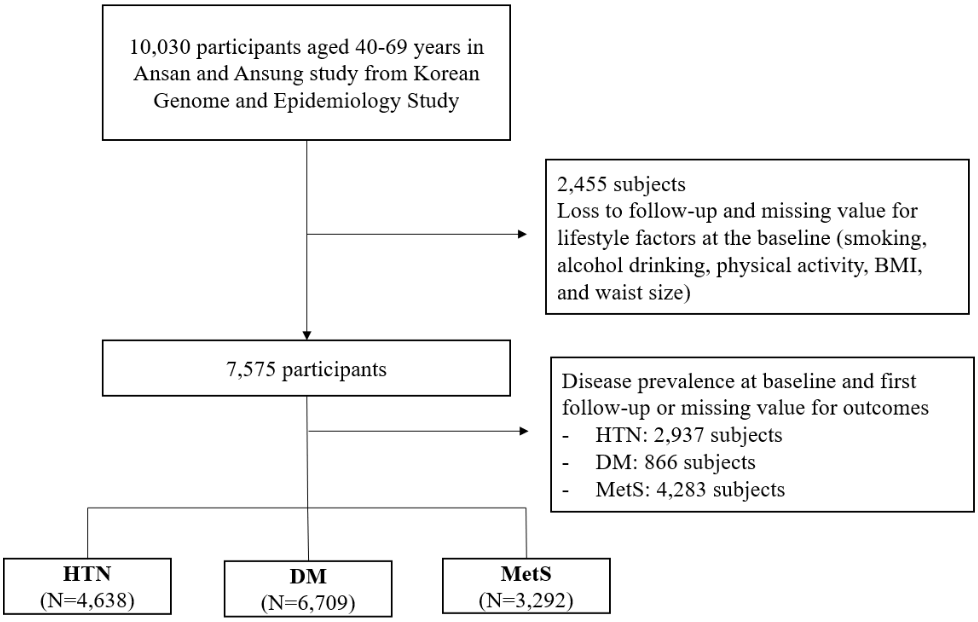


Figure 6. Flow chart of the study population selection from the Ansan and Ansung follow-up study

2.2.5. Biological age study

Of the 211,721 participants in the KoGES integrated data, a total of 101,980 healthy individuals aged 40–79 years with a Charlson's comorbidity index [77] of 0, and with body measurements, SES, history of disease, and lifestyle behaviors were finally included to calculate the BA. To estimate the risk of HTN, DM, and comorbidity of DM and HTN, a total of 43,143 individuals who had at least 2 years of follow-up years were included (Figure 7) [78]. The study protocol was reviewed and approved by the IRB of the Seoul National University (No. 1912-063-1088).

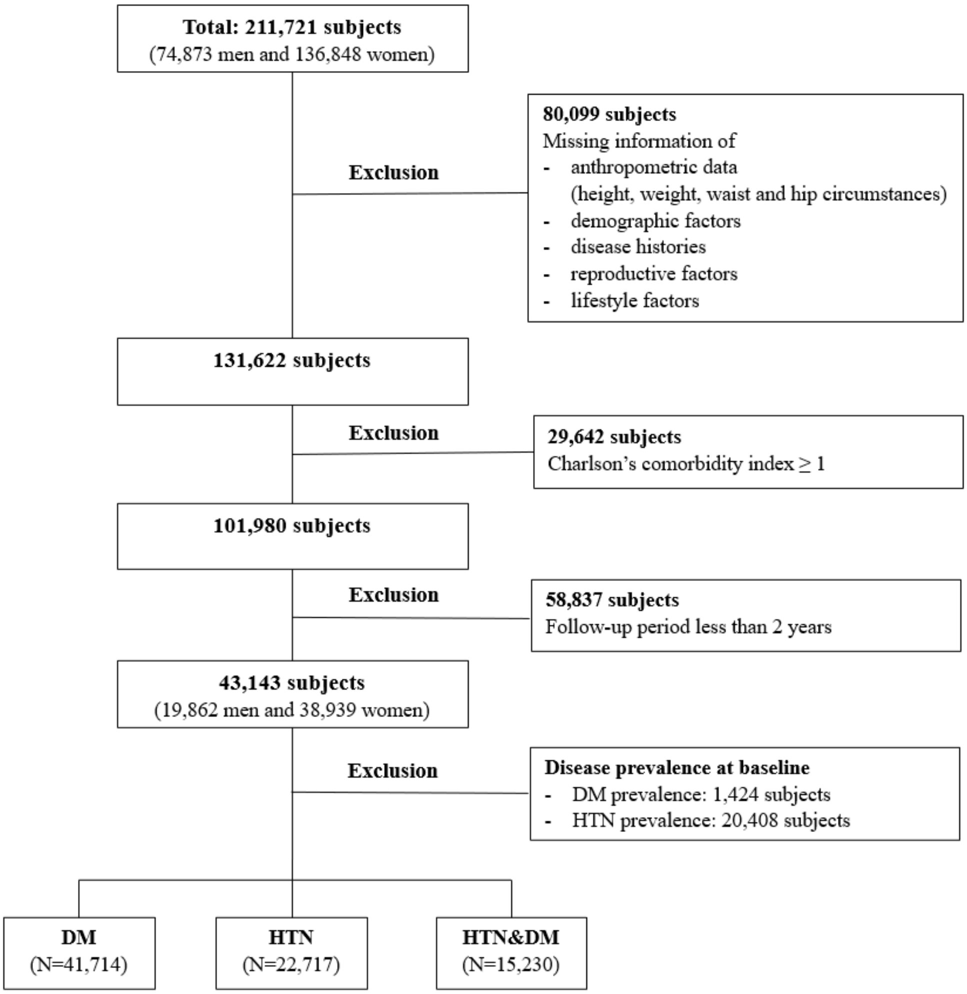


Figure 7. Flow chart of the study population selection from the Korean Genome and Epidemiology Study for development and validation of biological age

2.2.6. Prediction model study

A total of 211,721 individuals aged 40 to 79 years in integrated data, follow-up was conducted in 87,159 participants. After excluding 49,514 participants who were previously diagnosed with HTN, or missing information on demographic, lifestyle, blood test, history of CVD, and family history of CVD (N=19,404), a total of 30,110, 60,698, and 21,459 participants were included in this study for HTN, DM, and comorbidity of HTN and DM prediction model, respectively. The population selection after imputation for missing data were shown in Appendix 3. The study protocol was reviewed and approved by the IRB of the Seoul National University (No. 1912-063-1088).

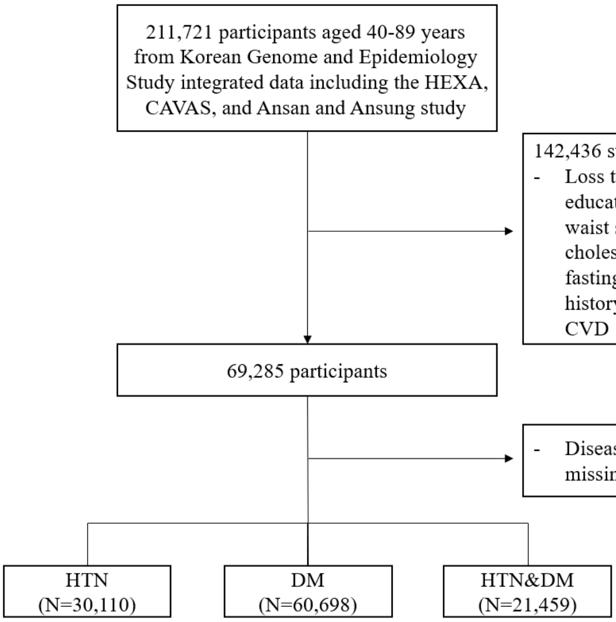


Figure 8. Flow chart of the study population selection from the Korean Genome and Epidemiology Study for disease prediction model

2.3. Key variables

2.3.1. Definition of metabolic disease

Hypertension

The American College of Cardiology and American Heart Association (ACC/AHA) 2017 guideline for high blood pressure in adults was used to classify HTN in this study [79]. We defined HTN as taking antihypertensive medication, having a systolic blood pressure of 130 mmHg or more, or having a diastolic blood pressure of 80 mmHg or above.

Diabetes mellitus

According to the WHO and International Diabetes Federation (IDF), DM is defined as using anti-diabetic drugs, having a fasting glucose level of 126 mg/dL or higher, or having an HbA1C level of 6.5% or higher [80]. On the other hand, DM was classified based on the plasma glucose level after 8 hours of fasting in KNHANES.

Dyslipidemia (hypercholesterolemia and hypertriglyceridemia)

The two kinds of LIP, hypercholesterolemia and hypertriglyceridemia were defined according to the National Cholesterol Education Program ' s Adult Treatment Panel III

(NCEP−ATP III) standards [81]. Total cholesterol of 240 mg/dL or higher was considered as hypercholesterolemia, and a TG level of 200 mg/dL or higher was regarded as hypertriglyceridemia.

Obesity

Based on the WHO obesity standards, different BMI definition was used between Korea and the US population [82]. Obesity was defined as a BMI of 25kg/m^2 or higher in the Korean population and 30kg/m^2 or higher in the US population.

Metabolic syndrome

According to the NCEP−ATP III criteria,[81] metabolic syndrome is classified as having three of the five conditions mentioned below: 1) A blood pressure of 130/85 mmHg or greater; 2) a fasting glucose level higher than 100 mg/dL; 3) HDL level less than 40 mg/dL for men and 50 mg/dL for women; 4) a TG level higher than 150 mg/dL; 5) a waist circumstance of 102 cm or greater for men and 88 cm or greater for women in the US and 90 cm or greater or for men and 85 cm or greater for women in the Korean population [83].

2.3.2. Exposure variables

Disease score

The disease score was calculated according to the presence of comorbidities (HTN, DM, and LIIP) at baseline.

Family history of CVD

A self-reported diagnosis of a first-degree family history of CVD was used to define a positive family history of CVD.

Cardiometabolic disease

We defined CMDs as self-reported history of HTN, DM, CHD, and stroke at baseline.

Healthy lifestyle score

Healthy lifestyle score (HLS) was generated using cigarette smoking (never, former, and current), alcohol drinking status (never, former, and current), and BMI level (<18.5, 18.5-22.9, 23.0-24.9, 25.0-27.4, 27.5-29.9, and ≥30kgm^2). The healthy lifestyle status was defined as 'nondrinker', 'nonsmoker', and 'BMI 18.5-27.4kg/m^2' in this study. One point was given if the alcohol drinking status was a 'nondrinker.' Another point was given if the smoking status was a 'nonsmoker. Lastly, a final point was given if the BMI was within the range of 18.5 to 27.4kg/m^2. The sum of these three points was defined as HLS, with increasing scores indicating healthier lifestyle.

Classification of change in lifestyle factors

The detailed classification of changes in lifestyle factors was shown in Table 2. Intensity of smoking was calculated using cigarettes per day. We categorized into maintenance of non-smoker (0), light/moderate smoker (<20), and heavy smoker (≥20 cigarettes per day). For alcohol consumption, the intensity was categorized into non-drinker (0), light/moderate drinker (<15g per day for women and <30g per day for men), and heavy drinker (≥15g per day for women and ≥30g per day for men). The change in BMI was categorized as underweight (<18.5), normal (18.5-23.0), and obesity (≥23kg/m^2) [36]. The waist size was categorized as normal (<85cm for women and <90cm for men), and abdominal obesity (≥85cm for women and ≥90cm for men)[83]. We also classified the each lifestyle factor' s trajectories over time. The detailed classification of lifestyle trajectories was shown in Appendix 4.

Table 2. Classification of change in lifestyle factors

	First examination	Second examination
Smoking status		
	Never	Never
	Past	Past
	Never/past	Current
	Current	Past
	Current	Current (dose decreased: $1^{st} > 2^{nd}$ dose)
	Current	Current (dose maintained: $1^{st} = 2^{nd}$ dose)
	Current	Current (dose increased: $1^{st} < 2^{nd}$ dose)
*Logical error	Past/current	Never → Past
Intensity of smoking cigarettes per day		
	Non-smoker	Non-smoker
	Non-smoker	Light/moderate smoker (<20 cigarettes/day)
	Non-smoker	Heavy smoker (≥20 cigarettes/day)
	Light/moderate smoker	Non-smoker
	Light/moderate smoker	Light/moderate smoker
	Light/moderate smoker	Heavy smoker
	Heavy smoker	Non-smoker
	Heavy smoker	Light/moderate smoker
	Heavy smoker	Heavy smoker
Alcohol drinking status		
	Never	Never
	Past	Past
	Never/past	Current
	Current	Past
	Current	Current (dose decreased: $1^{st} > 2^{nd}$ dose)
	Current	Current (dose maintained: $1^{st} = 2^{nd}$ dose)
	Current	Current (dose increased: $1^{st} < 2^{nd}$ dose)
*Logical error	Past/current	Never → Past
Intensity of alcohol consumption		
	Non-drinker	Non-drinker
	Non-drinker	Light/moderate drinker (<15g/day for women, <30g/day for men)
	Non-drinker	Heavy drinker (≥15g/day for women, ≥30g/day for men)
	Light/moderate drinker	Non-drinker
	Light/moderate drinker	Light/moderate drinker
	Light/moderate drinker	Heavy drinker
	Heavy drinker	Non-drinker
	Heavy drinker	Light/moderate drinker
	Heavy drinker	Heavy drinker

Table 2 (*Continued*). Classification of change in lifestyle factors

	First examination	Second examination
Physical activity status		
	Inactive	Inactive
	Inactive	Active
	Active	Inactive
	Active	Active
BMI status		
	Underweight	Underweight (BMI<18.5kg/m^2)
	Underweight	Normal weight (18.5≤BMI<25kg/m^2)
	Normal weight	Underweight
	Normal weight	Normal weight
	Normal weight	Obesity (25kg/m^2≤BMI)
	Obesity	Normal weight
	Obesity	Obesity
Waist size		
	Normal	Normal (<85cm for women, <90cm for men)
	Normal	Abdominal obesity (≥85cm for women, ≥90cm for men)
	Abdominal obesity	Normal
	Abdominal obesity	Abdominal obesity

Biological age

Among the 128 variables, those with missing rates of higher than 20% blood test and calculated dietary intake measurements, which needed to be measured by health professionals were excluded. This in this study, the BA was calculated using the following variables: (1) body measurement (height, weight, waist, and hip size); (2) demographic factors (income, education level, marital status, and occupation); (3) lifestyle behaviors (smoking duration [years], smoking consumption [packs per day], second-hand smoking [yes/no], drinking frequency [none, 1 time, 2–3 times, 4–6 times/week and daily], frequency of regular exercise [none, 1 time, 2–3 times, 4–6 times/week and daily]); and (4) history of disease (dyslipidemia, asthma, allergy, and thyroid disease).

To identify the definite impact of the BA on metabolic disease, we newly defined 'Age−difference (Age−Diff)', as the difference between CA and BA ('Age−Diff' = BA−CA). The categories of 'Age−Diff were classified into four groups: "Very young BA (BA−CA ≤ −5)"; "Young BA (−5 < BA−CA ≤ −1)"; "Same BA as CA (−1 < BA−CA ≤ 1);"; "Older BA (BA−CA > 1)", respectively.

2.3.3. Outcome variables

CVD incident

The primary outcome was a new diagnosis of CVD including MI and stroke. The endpoint of this study was the date of CVD diagnosis, or last date of follow−up. CVD was defined as positive response to medical history questionnaires during routine examinations. The date of the latest follow−up was February 2017.

CVD and premature CVD death

The outcomes of interest were all−cause, CVD death, and premature death, defined as death before 70 years of age. According to the ICD−9 and ICD−10 codes, the cause of death was classified as follows: all−cause (all ICD−9 or ICD−10 codes, except external causes of death), CVD (ICD−9 codes 390−459; ICD−10 codes I00−I99), ischemic heart disease (IHD ICD−9 code 410−414; ICD−10 code I20−I25), stroke, (ICD−9 code 430−438; ICD−10 code I60−69), ischemic stroke (ICD−9 code 434; ICD−10 code I63), and hemorrhagic stroke (ICD−9 code 431; ICD−10 code I60−I62).

Metabolic outcomes (HTN, DM, comorbidity of HTN and DM, and metabolic syndrome)

HTN incident was defined as systolic blood pressure more than 130 mmHg or diastolic blood pressure higher than 80 mmHg or taking any antihypertensive drugs during the follow−up period [84]. DM was defined as either a fasting plasma glucose level higher than 126 mg/dL, HbA1c level greater than 6.5%, or taking any anti−diabetic medications [85]. Comorbidity of HTN and DM was defined as co−occurrence of HTN and DM at the same time. And a new diagnosis of MetS during follow−up period.

2.4. Statistical analysis

2.4.1. Age−standardized prevalence

We estimated the age−standardized prevalence rate with the 95% confidence intervals (CIs) of HTN, DM, LIP, obesity, MetS, and the comorbidity in each study population. All prevalence rates was calculated based on direct age−standardized approaches[86] using the WHO 2000−2025 world standard population database (Table 3) [87].

The age−standardized rate (ASR) is calculated as follow:

$$\mathrm{ASR} = \frac{\sum_{i=1}^{A} a_i w_i}{\sum_{i=1}^{A} w_i}$$

w_i: Number of populations in the ith age group of the WHO world standard population.

a_i: Age−specific rate in the ith age group.

Equation 1. Calculation of direct age standardized rate[44]

The estimates were also subdivided by sex and the median survey years (before/after 2010).

Table 3. The World Health Organization 2000-2025 world standard population for each age group

Age group	**WHO world standard population**
0-4	8,860
5-9	8,690
10-14	8,600
15-19	8,470
20-24	8,220
25-29	7,930
30-34	7,610
35-39	7,150
40-44	6,590
45-59	6,040
50-54	5,370
55-59	4,550
60-64	3,720
65-69	2,960
70-74	2,210
75-79	1,520
80-84	910
85+	630
Total	100,000

2.4.2. Cardiovascular risk assessment

Family history of CVD and the risk of CVD study

The baseline characteristics were compared using Chi-squared test for categorical variables and Student's t-test for continuous variables. We performed multivariable Cox proportional hazards regression analysis to estimate hazard ratios (HRs) with 95% confidence intervals (CIs) for CVD outcomes including MI and stroke according to the family history of CVD and the baseline disease status. To assess the fitness of Cox proportional hazard model, proportional hazard assumption was tested with scaled Schoenfeld residuals. For primary analysis, we assessed a combined association between metabolic comorbidities and family history of CVD and CVD outcomes. Adjusted HRs and 95% CI for MI were calculated from adjusting the age, sex, BMI, waist to hip ratio (WHR), smoking status, alcohol drinking, regular exercise, and income level. In this analysis, we considered individuals with a negative family history of CVD and none of metabolic diseases as reference group.

Lifestyle factors and the risk of CVD death study

Baseline characteristics of 11 cohorts are presented as mean ±

standard deviation (SD) for continuous variables, and as numbers and percentages for categorical variables. For primary analysis, cox proportional hazard analysis was used to estimate hazard ratios (HRs) with 95% confidence intervals (95% CIs) of all−cause and CVD−specific death associated with HLS according to the combination of CMDs. In this analysis, we considered individuals with none of healthy lifestyle factors as reference group in each disease status.

As secondary analyses, we assessed the association of combination of healthy lifestyle factors with death and premature death from all−cause and CVD−specific death according to the number of CMDs at baseline. In this analysis, we categorized individuals into sixteen groups based on combination of healthy lifestyle factors: (1) none of healthy lifestyle factors, (2) non−smoking, (3) non−drinking, (4) healthy BMI, (5) non−smoking and non−drinking, (6) non−smoking and healthy BMI, (7) non−drinking and healthy BMI, (8) non−smoking, non−drinking, and healthy BMI.

Change in lifestyles study

After assessing the goodness of fit of proportional hazard assumption, we performed multivariable cox proportional hazard regression model to calculate the HR and 95% CIs for outcomes according to changes in lifestyle factors between the two biennial follow-up period (2001–2003 and 2003–2005). The HRs were adjusted for potential confounders of age, sex, education level, income level, smoking, alcohol drinking, physical activity, BMI, total cholesterol, and family history of CVD.

2.4.3. Prediction model

Biological age study

For continuous elements including body measurements and lifestyle information, we performed z-score standardization. To find the optimized coefficients for variables and calculate the BA, elastic net linear regression [88] was applied using the standardized elements. Then we validated it based on the 10-fold cross-validation [89] using the R (version 3.3.3) with the '*glmnet*' package. We estimated the correlation coefficients (r) to calculate the correlation between CA and BA. To estimate the odds ratios (ORs) of metabolic diseases according to CA (<50, 50–59, 60–59, and ≥70 years), BA (<50, 50–59, 60–59 and ≥70 years), and Age-Diff groups, we conducted logistic regression analyses. To assess the HR of BA and the risk of HTN, DM, and comorbidity of HTN and DM, we conducted Cox proportional hazards regression analyses. Moreover, further analyses were performed to assess the risk of disease within a 5-year follow-up.

Prediction model study

The cox proportional hazard regression model is a semi-parametric model for survival data that estimate the effect of covariates on the hazard rate [90]. Random Forest (RF) is a method based on the decision tree to identify complicated interactions and non-linearities of predictor effects for risk stratification with a lower prediction error than statistics-based modeling. In this study random survival forest (RSF), which calculate the cumulative hazard for each tree's terminal nodes and generate an ensemble cumulative hazard based on the RF model was used [91]. The gradient boosting machine (GBM) is an ensemble leaning algorithm, which develops a prediction model by additive extension of multiple models [92]. An elastic net regularized cox proportional hazards regression is another machine learning methods optimized a predictive model [88, 93]. In this study, we developed HTN, DM, and comorbidity of HTN and DM prediction model based on the cox proportional hazard regression model, and machine learning approaches (RSF, GBM, and elastic net). For imputation of missing data, we used multivariate data imputation methods.

According to the variable selection, the following variable sets were used as for prediction model: (i) all variable from previously published prediction model of each disease (Appendix 5-6) and variables with variance inflation factor (VIF)<5 were selected (Model 1); (ii) statistically significant variables from cox proportional hazard regression model (Model 2).

The prediction model based on the statistics and machine learning based modeling was developed in the training set (70% of the total population) and validated in the test set (30% of the total population). Predictive performance of the model was tested based on the concordance statistics (c-index), which showed the probability of the model to predict the developing disease risk

All statistical analyses in this study were done with SAS 9.4 software (SAS Institute, Cary, NC, USA) and R with the *mice*, *glmnet*, *gbm*, *randomForestSRC* packages (version 4.1.0).

III. Results

3.1. Prevalence study

This study was published in An et al. (2022) [S. An, C. Ahn, J. Jang, J. Lee, D. Kang, JK. Lee, SK. Park, "Comparison of the Prevalence of Cardiometabolic Disorders and Comorbidities in Korea and United States: Analysis of the National Health and Nutrition Examination Survey", "*Journal of Korean Medical Science*", 2022, 37 (18)].

General characteristics

Individuals' mean age was 54.0, 53.9, 52.7, and 56.5 years old in the NHANES, KNHANES, HEXA, and CAVAS study, respectively. The highest BMI was shown in NHANES at 29.3kg/m^2, while the lowest was found among HEXA at 23.9kg/m^2. Korean populations were less likely to smoke and drink alcohol than those in the US, and they were more likely to do physical activity. The prevalence of CVD and cancer was significantly greater in US individuals than those in KNHANES and KoGES (Table 4).

Table 4. Baseline characteristics of the four study groups (NHANES, KNHANES, HEXA, and CAVAS)

	NHANES	KNHANES	HEXA	CAVAS
No. of participants	15,872	26,492	139,345	24,994
Study Entry, yr	2003-2014	2007-2014	2004-2013	2005-2011
	Mean ± SD	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**
Age, yr	54.0 ± 8.61	53.9 ± 8.63	52.7 ± 7.99 *	56.5 ± 7.94 *
BMI, kg/m^2	29.3 ± 6.09 *	24.1 ± 3.12	23.9 ± 2.90 *	24.5 ± 3.14 *
	N (%)	**N (%)**	**N (%)**	**N (%)**
Women, %	8,101 (51.0) *	14,829 (56.0)	92,368 (66.3) *	15,551 (62.2) *
College or more, %	7,867 (49.6) *	8,385 (31.7)	58,484 (42.0) *	5,274 (21.1) *
Ever smokers, %	7,906 (49.8) *	9,873 (37.3)	37,266 (26.7) *	7,130 (28.5) *
Ever drinkers, %	12,172 (76.7) *	20,816 (78.6)	67,234 (48.3) *	12,223 (48.9) *
Regular exercise, %	2,259 (14.2) *	7,211 (27.2)	73,649 (52.9) *	8,115 (32.5) *
Stroke, %	598 (3.8) *	542 (2.1)	1,448 (1.0) *	574 (2.3) *
Myocardial infarction, %	668 (4.2) *	643 (2.1)	3,382 (2.4) *	521 (2.4) *
Cancer, %	1,348 (8.5) *	1,028 (3.9)	4,376 (3.1) *	566 (2.3) *

Abbreviation: NHANES, National Health and Nutrition Examination Survey; KNHANES, Korean National Health and Nutrition Examination Survey; HEXA-KoGES, Health Examinees study (an urban cohort study) in the Korean Epidemiology and Genome Study; CAVAS-KoGES, Cardiovascular disease association study (a rural cohort study) in the Korean Epidemiology and Genome Study
*P < 0.001 for the test for the difference between each group and the KNHANES

Prevalence of metabolic diseases

The prevalence of metabolic disease was greater in US adults, while lowest prevalence was shown in Korean urban population. Among the diseases, HTN had the highest age-standardized prevalence, with more than a half of the individuals in NHANES (56.8%) and Korea (KNHANES, 49.9 %; HEXA, 51.0 %; CAVAS, 60.3 %). HTN, obesity, and MetS were prevalent in the NHANES (56.8%, 38.6%, 36.5%) and CAVAS (60.3%, 40.9%, 33.2%) than in the KNHANES (49.9%, 36.2%, 29.4%). On the other hand, we found the lowest prevalence of LIP (hypercholesterolemia, 11.3%; hypertriglyceridemia, 12.8%), obesity (31.9%), MetS (18.8%) in the Koreans living in urban areas (Figure 9).

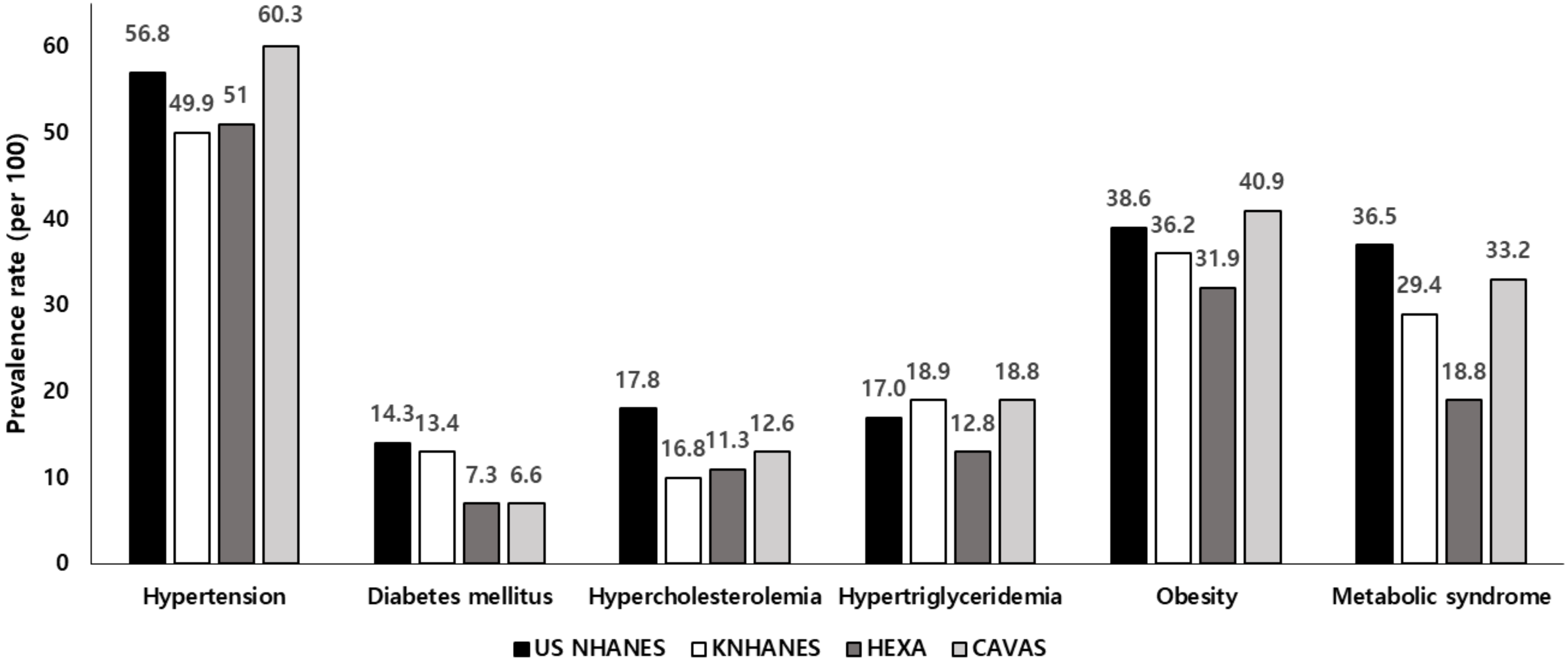


Figure 9. Age-standardized prevalence rates of metabolic disease

The gender difference in the age-standardized prevalence of metabolic diseases was likewise comparable across the two countries. In both US and Korea, we found that men have a larger prevalence of HTN (Men – NHANES, 60.0%; KNHANES, 58.7%; HEXA, 62.9%; CAVAS, 69.5%; Women – NAHENS, 53.7%; KNHANES, 41.0%; HEAX, 45.4%; CAVAS, 55.6%), DM (Men – NHANES, 16.1%; KNHANES, 16.1%; HEXA, 10.2%; CAVAS, 9.2%; Women – NAHENS, 12.6%; KNHANES, 10.7%; HEAX, 5.9%; CAVAS, 5.3%), hypertriglyceridemia (Men – NHANES, 21.5%; KNHANES, 26.1%; HEXA, 21.3%; CAVAS, 28.0%; Women – NAHENS, 12.7%; KNHANES, 12.3%; HEAX, 8.8%; CAVAS, 14.1%), and MetS (Men – NHANES, 37.4%; KNHANES, 33.5%; HEXA, 25.0%; CAVAS, 37.9%; Women – NAHENS, 35.5%; KNHANES, 25.1%; HEAX, 16.0%; CAVAS, 31.3%) than women, while women had a greater prevalence of hypercholesterolemia (Men – NHANES, 16.5%; KNHANES, 14.5%; HEXA, 9.6%; CAVAS, 10.8%; Women – NAHENS, 19.0%; KNHANES, 18.9%; HEAX, 12.4%; CAVAS, 13.9%). Obesity was more common in women (39.9%) in the US than in men (37.3%), whereas it was more prevalent in men in Korea (Men – KNHANES, 38.8%; HEXA, 40.1%; CAVAS, 41.9%; Women – KNHANES, 33.3%; HEAX, 28.3%; CAVAS, 40.8%) (Figure 10 & Appendix 7).

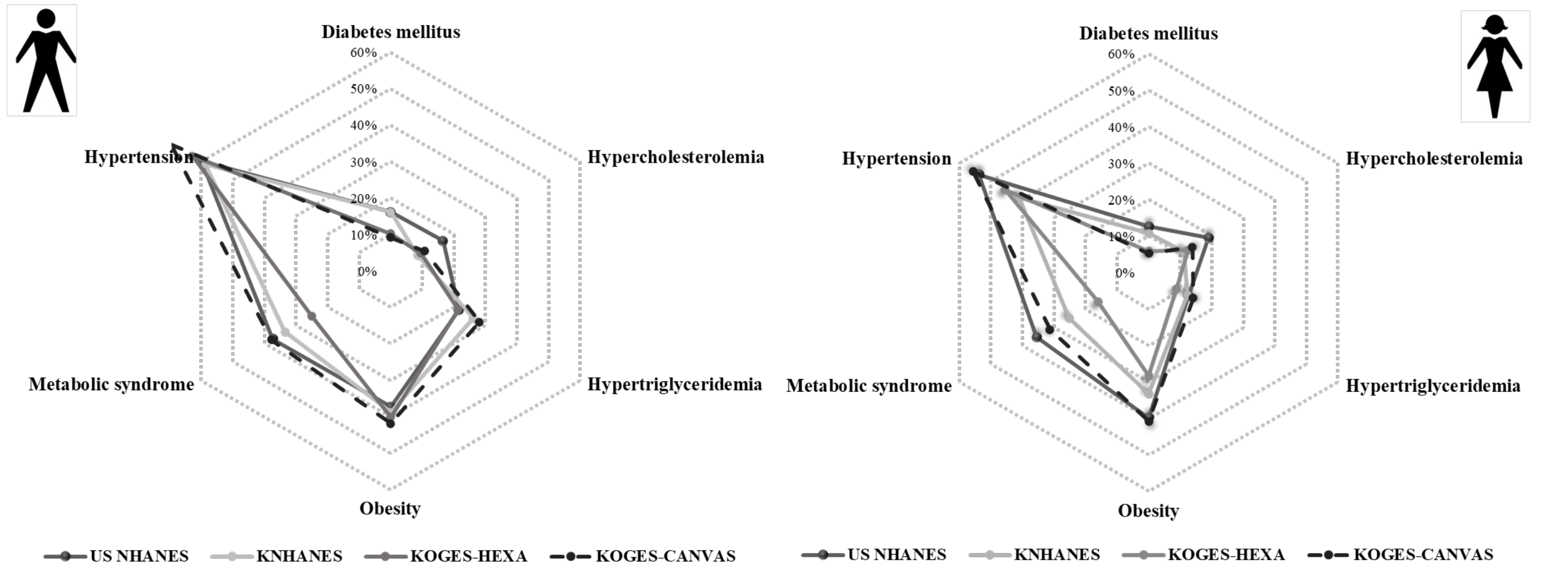


Figure 10. Sex-specific age-standardized prevalence rates (per 100 persons) of metabolic diseases

Comparison of the prevalence rates between before and after 2010

According to the median survey years, we subdivided the prevalence rate into before 2010 (from 2003 to 2009) and after 2010 (from 2010 to 2014). We found that more than a half of the population was diagnosed with HTN after 2010 in both Korea and the US. For DM, there is increasing prevalence rates after 2010 in Korea (KNHANES, from 11.8% to 14.1%; HEXA, from 5.7% to 9.7%; CAVAS, from 6.4% to 8.9%). Individuals in KNHANES had higher prevalence rates of hypercholesterolemia (from 13.8% to 18.2%), while individuals in NHANES had lower prevalence rates (hypercholesterolemia, from 19.2% to 15.4%; hypertriglyceridemia, from 18.2% to 15.1%) after 2010. Moreover, we found that individuals in KNHAENS had greater prevalence of hypercholesterolemia and hypertriglyceridemia (18.2% and 18.9%, respectively) compared to those in NHANES (15.4% and 14.4%). Prevalence of obesity increased in NHANES (from 37.8% to 40.3%) and CAVAS population (from 40.8% to 41.7%), while decreased prevalence was observed in KNHANES (from 37.1% to 35.8%) and HEXA (from 32.3% to 31.3%) population. For MetS, the decreased prevalence rates were observed in both US and Korea (NHANES, from 37.1% to 35.5%; KNHANES, from 32.7% to 28.3%; HEXA, from 19.4% to 17.8%; CAVAS, from 33.8% to 28.5%) (Figure 11).

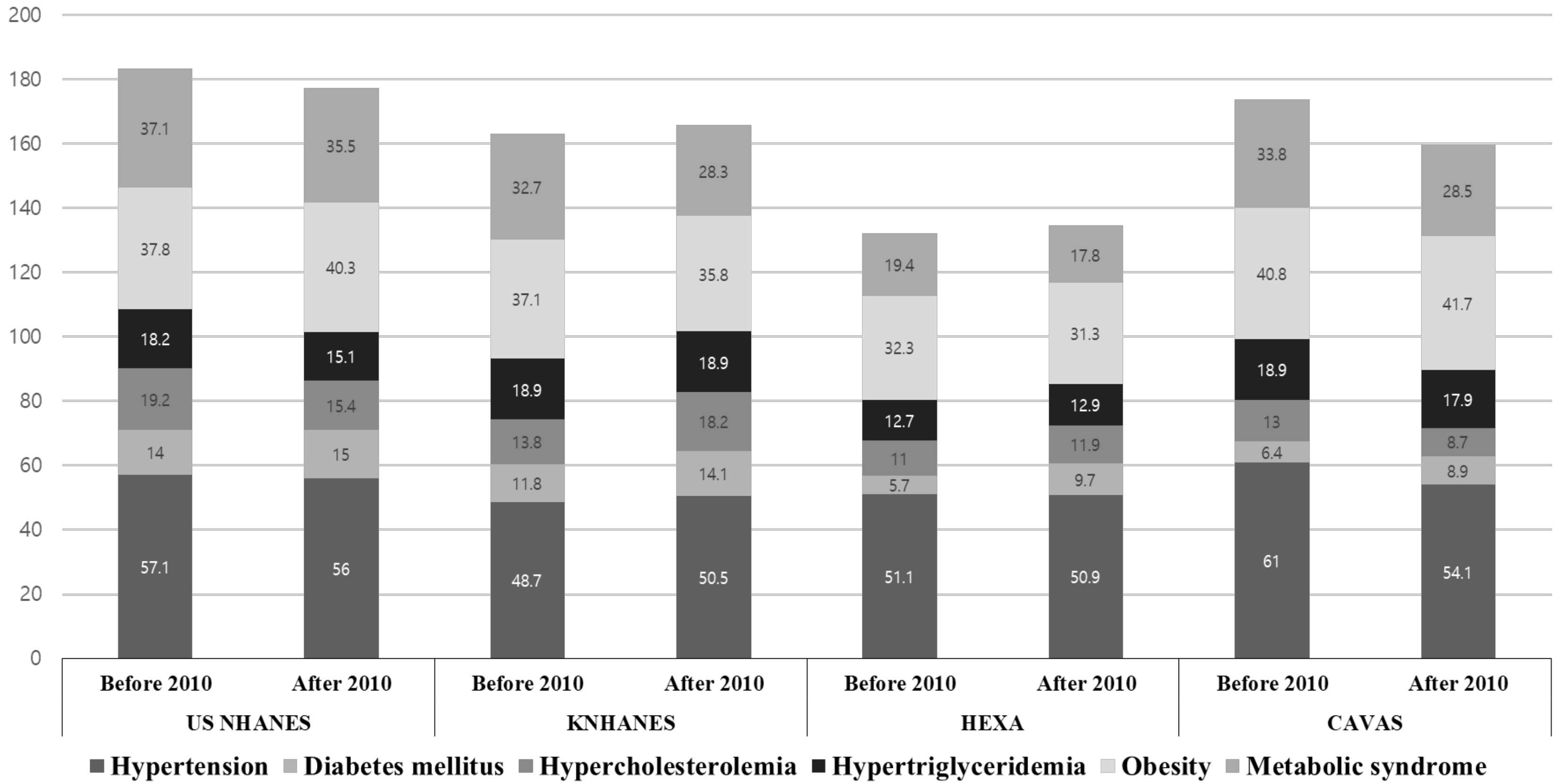


Figure 11. Age-standardized prevalence rates (per 100 persons) of metabolic disease according to the median survey year

Prevalence of metabolic comorbidity

The comorbidity was less common in Korea than in the US. Overall, 24.4% of the NHANES, 29.3% of the KNHANES, 30.9% of HEXA, and 19.5% of CAVAS were free of metabolic diseases. In the KNHANES, 31.1%, 23.2%, 11.8%, and 4.6% had one, two, three, and four diseases, respectively, whereas the NHANES had 31.5%, 26.0%, 13.4%, and 4.7%, respectively. Individuals living in rural areas are more likely to have comorbidities compared to those in urban areas (Figure 12).

In both Korean (KNHANES, 11.6%; HEXA, 12.7%; CAVAS, 17.2%) and US population (12.5%), the most common composition was HTN and obesity. In US population, the second most common composition was HTN, DM, and obesity (5.2%), while HTN, hypertriglyceridemia, and obesity in Korea (KNHENAS, 4.3%; HEXA, 3.1%; CAVAS, 5.6%) (Figure 12 and Appendix 8).

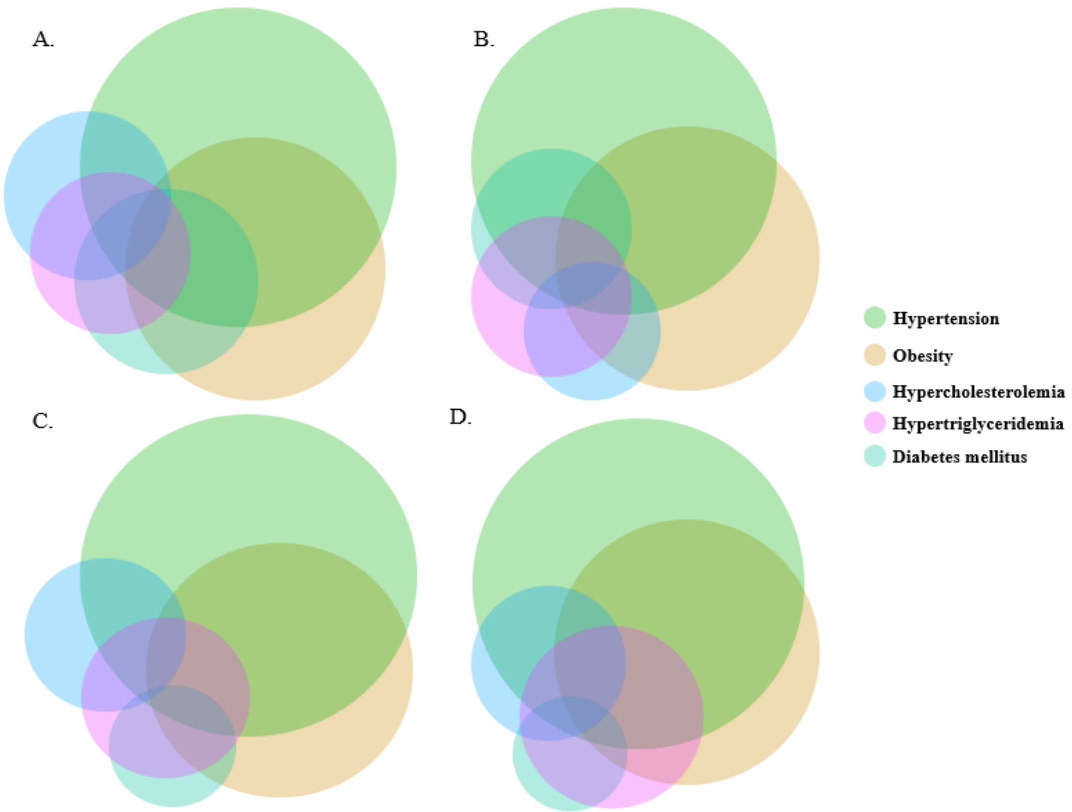


Figure 12. Combination of age-standardized prevalence for disease comorbidity according to each studies (A. NHANES; B. KNHANES; C. HEXA; D. CAVAS)

3.2. Family history of CVD and the risk of CVD study

General characteristics

Among 72,111 individuals (mean [SD] age, 54.3 [8.40] years; 24,605 [34.1%] men), 14,169 (19.6%) had a positive family history of CVD in first degree while 57,942 (80.4%) reported a negative family history of CVD. At baseline, individuals with a positive family history of CVD were more likely to be current alcohol drinker, have high income level, and have HTN and LIP compared to the those with a negative family history of CVD (Table 5). During a median follow-up of 5 years (range, 1–14 years), there were 983 (1.4%) and 559 (0.8%) cases of MI and stroke, respectively. Compared to the individuals with a negative family history of CVD, those with a positive family history showed a greater risk for CVD (HR 1.28, 95% CI: 1.13–1.44). Compared to the individuals with none of diseases, the risks of CVD were 1.46 (95% CI: 1.25–1.70) in participants with one disease, 2.00 (95% CI: 1.70–2.35) in those with two diseases, and 2.25 (95% CI: 1.78–2.84) in those with three diseases. Similarly, increase of disease score was associated with an increase in risk of MI and stroke (Appendix 9–10). Among individuals with a positive family history of CVD, current smoking, obesity, and physical inactive were significantly associated with increased risk of CVD (Appendix 11).

Table 5. Baseline characteristics of participants by family history of cardiovascular disease

	Negative family history of CVD (N=57,942)	**Positive family history of CVD (N=14,169)**	***p*-value**
Age, years	54.3 ± 8.49	54.0 ± 8.04	<.001
Male, N (%)	37,852 (65.3)	9,654 (68.1)	<.001
Monthly income ≥ ₩4,000K, N (%)	11,079 (19.1)	3,258 (23.0)	<.001
Current smoker, N (%)	6,445 (11.1)	1,452 (10.3)	0.024
Current alcohol drinker, N (%)	25,212 (43.5)	6,291 (44.4)	0.046
Regular exercise, N (%)	29.898 (51.6)	7,619 (53.8)	<.001
BMI ≥ 25 kg/m^2	19,161 (33.1)	4,782 (33.8)	0.057
WHR ≥ 0.90 for men, 0.85 for women	26,814 (46.3)	6,504 (45.9)	0.272
Hypertension, N (%)	31,165 (53.8)	81,29 (57.4)	<.001
Diabetes mellitus, N (%)	5,552 (9.6)	1,288 (9.1)	0.073
Dyslipidemia, N (%)	21,275 (36.7)	5,596 (39.5)	<.001

Abbreviation: CVD, cardiovascular disease; BMI, body mass index; WHR, waist to hip ratio;

Cardiometabolic disease, family history of CVD, and the risk of MI and stroke

The combined association of family history of CVD and the combination of metabolic diseases with the risk of CVD including MI and stroke is shown in Table 5. After adjustment for age, sex, body mass index, waist to hip ratio, smoking status, alcohol drinking, regular exercise, and income level, individuals with a positive family history and metabolic disease had a higher risk for CVD.

Among individuals with a negative family history of CVD, the HRs for CVD were 1.41 (95% CI: 0.88–2.27) for individuals with DM, 1.43 (95% CI: 1.19–1.72) in those with HTN, 1.32 (95% CI: 1.04–1.66) in those with LIP, 1.98 (95% CI: 1.48–2.64) in those with DM and HTN, 2.25 (95% CI: 1.51–3.37) in those with DM and LIP, 1.91 (95% CI: 1.59–2.30) in those with HTN and LIP, and 2.16 (95% CI: 1.66–2.81) in those with DM, HTN, and LIP compared to the individuals without family history of CVD and none of metabolic diseases (Table 6).

Among individuals having a positive family history of CVD, the HRs for CVD were 1.09 (95% CI: 0.78–1.53) for participants with none of diseases, 1.89 (95% CI: 0.70–5.10) in those with DM, 2.02 (95% CI: 1.59–2.57) in those with HTN, 1.48 (95% CI: 1.01–2.17) in those with LIP, 1.93 (95% CI: 1.07–3.47) in those with DM and

HTN, 2.28 (95% CI: 0.94−5.55) in those with DM and LIP, 2.56 (95% CI: 2.02−3.24) in those with HTN and LIP, and 2.88 (95% CI: 1.96−4.24) in those with DM, HTN, and LIP than the people with a negative family history of CVD and none of metabolic diseases (Table 6).

Table 6. Combined association of family history of cardiovascular disease and combination of metabolic disease with cardiovascular disease risk

Family history of CVD	Disease status at baseline	No. of cohorts	Cardiovascular disease	
			No. of CVD	Hazard Ratio[1] (95% CI)
Negative				
	None	18,019	184	1.00
	DM	832	19	1.41 (0.88-2.27)
	HTN	16,116	333	1.43 (1.19-1.72)
	LIP	7,166	120	1.32 (1.04-1.66)
	DM and HTN	17,00	65	1.98 (1.48-2.64)
	DM and LIP	760	28	2.25 (1.51-3.37)
	HTN and LIP	11,089	334	1.91 (1.59-2.30)
	DM, HTN, and LIP	2,260	87	2.16 (1.66-2.81)
Positive				
	None	4,037	41	1.09 (0.78-1.53)
	DM	157	4	1.89 (0.70-5.10)
	HTN	4,013	108	2.02 (1.59-2.57)
	LIP	1,709	30	1.48 (1.01-2.17)
	DM and HTN	366	12	1.93 (1.07-3.47)
	DM and LIP	137	5	2.28 (0.94-5.55)
	HTN and LIP	3,122	118	2.56 (2.02-3.24)
	DM, HTN, and LIP	628	31	2.88 (1.96-4.24)

Abbreviation: CVD, Cardiovascular disease; MI, Myocardial infarction; CI, Confidence interval; HTN, Hypertension; DM, Diabetes mellitus; LIP, Dyslipidemia

1. Cox proportional hazards regression model were adjusted by sex, age at baseline, body mass index, waist and hip ratio, current smoking status, current alcohol drinking, regular exercise, and income level

The combined association of family history of CVD and the combination of metabolic diseases with the risk of MI and stroke is shown in Table 7.

Among individuals with a positive family history of CVD, the HRs for MI were 1.28 (95% CI: 0.85−1.91) for individuals with none of diseases, 1.50 (95% CI: 0.37−6.09) in those with DM, 1.91 (95% CI: 1.39−2.60) in those with HTN, 1.78 (95% CI: 1.12−2.81) in those with LIP, 1.86 (95% CI: 0.86−4.01) in those with DM and HTN, 2.38 (95% CI: 0.75−7.50) in those with DM and LIP, 2.91 (95% CI: 2.17−3.89) in those with HTN and LIP, and 3.30 (95% CI: 2.06−5.30) in those with DM, HTN, and LIP than the people with a negative family history of CVD and none of metabolic diseases (Table 7). The HRs for stroke were 0.85 (95% CI: 0.46−1.58) for individuals with none of diseases, 2.65 (95% CI: 0.65−10.81) in those with DM, 2.34 (95% CI: 1.62−3.39) in those with HTN, 1.02 (95% CI: 0.49−2.12) in those with LIP, 2.48 (95% CI: 1.07−5.73) in those with DM and HTN, 2.06 (95% CI: 0.50−8.43) in those with DM and LIP, 2.15 (95% CI: 1.44−3.22) in those with HTN and LIP, and 2.54 (95% CI: 1.33−4.84) in those with DM, HTN, and LIP (Table 7).

Table 7. Combined association of family history of cardiovascular disease and combination of metabolic disease with risk of myocardial infarction and stroke

Family history of CVD	Disease status at baseline	No. of cohorts	Myocardial infarction		stroke	
			No. of MI	Hazard Ratio[1] (95% CI)	No. of stroke	Hazard Ratio[1] (95% CI)
Negative						
	None	18,019	112	1.00	72	1.00
	DM	832	13	1.65 (0.93-2.94)	6	1.06 (0.46-2.45)
	HTN	16,116	220	1.56 (1.23-1.96)	117	1.31 (0.97-1.77)
	LIP	7,166	87	1.60 (1.21-2.12)	35	0.96 (0.64-1.44)
	DM and HTN	17,00	41	2.09 (1.45-3.01)	26	1.98 (1.25-3.14)
	DM and LIP	760	21	2.91 (1.81-4.66)	7	1.32 (0.60-2.88)
	HTN and LIP	11,089	207	1.98 (1.56-2.51)	135	1.91 (1.42-2.58)
	DM, HTN, and LIP	2,260	54	2.29 (1.64-3.20)	34	1.99 (1.31-3.04)
Positive						
	None	4,037	30	1.28 (0.85-1.91)	12	0.85 (0.46-1.58)
	DM	157	2	1.50 (0.37-6.09)	2	2.65 (0.65-10.81)
	HTN	4,013	62	1.91 (1.39-2.60)	48	2.34 (1.62-3.39)
	LIP	1,709	22	1.78 (1.12-2.81)	8	1.02 (0.49-2.12)
	DM and HTN	366	7	1.86 (0.86-4.01)	6	2.48 (1.07-5.73)
	DM and LIP	137	3	2.38 (0.75-7.50)	2	2.06 (0.50-8.43)
	HTN and LIP	3,122	81	2.91 (2.17-3.89)	38	2.15 (1.44-3.22)
	DM, HTN, and LIP	628	21	3.30 (2.06-5.30)	11	2.54 (1.33-4.84)

Abbreviation: CVD, Cardiovascular disease; MI, Myocardial infarction; CI, Confidence interval; HTN, Hypertension; DM, Diabetes mellitus; LIP, Dyslipidemia

1. Cox proportional hazards regression model were adjusted by sex, age at baseline, body mass index, waist and hip ratio, current smoking status, current alcohol drinking, regular exercise, and income level

According to the disease score, among individuals who had a positive family history of CVD, the HRs for CVD were 1.87 (95% CI: 1.50−2.33) in individuals with one disease, 2.47 (95% CI: 1.97−3.10) in those with two diseases, and 2.88 (95% CI: 1.96−4.24) in those with three diseases compared to the people with a negative family history of CVD and none of metabolic diseases. For MI, the HRs were 1.86 (95% CI: 1.40−2.47) in individuals with a positive history of CVD and one disease, 2.77 (95% CI: 2.09−3.67) in those with two diseases, and 3.30 (95% CI: 2.06−5.39) in those with three diseases. For stroke, the HRs were 1.99 (95% CI: 1.40−2.82) in individuals with a positive history of CVD and one disease, 2.18 (95% CI: 1.49−3.18) in those with two diseases, and 2.52 (95% CI: 1.33−4.79) in those with three diseases. The risk for CVD, MI, and stroke significantly increased with increasing number of metabolic diseases (P−trend <.001). (Figure 13).

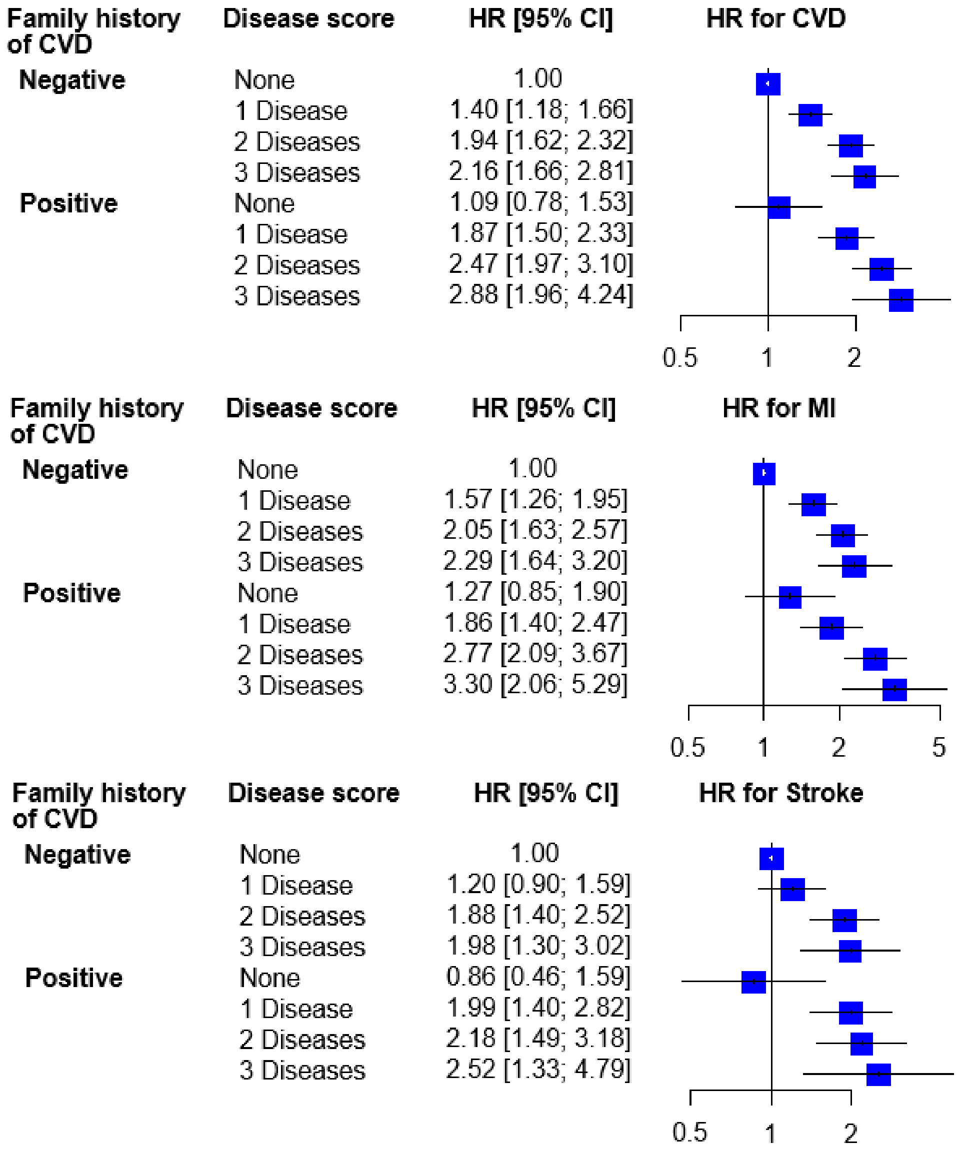


Figure 13. Combined association of family history of cardiovascular disease and disease score with risk for cardiovascular disease, myocardial infarction, and stroke

3.1. Lifestyle factors and the risk of CVD death study

General characteristics

A total of 403,852 Asian individuals (mean [SD] age, 53.7 [9.9] years; 51.5% female) from 11 multinational cohorts participated in this study. During a median of 15 follow-up years, 59,368 (14.7%) all-cause and 17,152 (4.2%) cardiovascular deaths occurred. Of all deaths combined, there were 22,557 all-cause premature deaths and 5,774 premature CVD deaths (Table 8).

Table 8. Baseline characteristics of participants in the Asian Cancer Consortium

Country	Cohort	Study entry	Follow-up years	Women	Age at enrollment	BMI (kg/m^2)	Current smokers	All-cause death	
								All-cause	CVD
	N	Year	Median	%	Mean (SD)	Mean (SD)	%	N	N
ACC, total	403,852	1984-2006	15	51.1	53.7 (9.9)	23.6 (3.17)	29.7	59,368	17,152
China									
SMHS	61,405	2001-2006	10	0	55.4 (9.7)	23.7 (3.1)	58.7	5,428	1,792
SWHS	74,912	1996-2000	15	100	52.6 (9.1)	24.0 (3.4)	2.4	7,640	2,486
Taiwan									
CBCSP	23,618	1991-1992	16	49.6	47.3 (10.0)	24.0 (3.4)	25.3	2,732	552
Japan									
JPHC1	42,600	1990-1992	23	52.2	49.6 (5.9)	23.6 (3.0)	28.4	7,357	1,822
JPHC2	55,398	1992-1995	20	52.6	54.2 (8.8)	23.5 (3.1)	27.6	12,411	3,122
Ohsaki	41,157	1995	13	46.9	59.7 (10.4)	23.5 (3.1)	33.1	7,016	2,349
Miyagi	36,056	1990	18	45.2	51.6 (7.6)	23.6 (3.0)	37.3	4,697	1,148
3 pref. Miyagi	21,540	1984	15	52.8	56.5 (11.3)	23.1 (3.2)	30.4	4,146	1,629
Takayama	19,213	1992	16	45.4	53.8 (11.8)	22.3 (2.8)	35.3	3,072	1,001
Korea									
KMCC	16,084	1993-2005	13	59.6	54.4 (13.2)	23.7 (3.2)	26.9	3,020	700
Singapore									
SCHS	11,869	1993-1999	12	26.7	55.2 (7.6)	22.9 (3.2)	34.6	1,849	551

Abbreviation: N, number; SD, standard deviation; BMI, body mass index; CVD, cardiovascular disease; SMHS, Shanghai Men's Health Study; SWHS, Shanghai Women's Health Study; CBCSP, Community-based Cancer Screening Project; JPHC1, Japan Public Health Center-based prospective Study; JPHC2, Japan Public Health Center-based prospective Study; Ohsaki, Ohsaki National Health Insurance Cohort Study; Miyagi, Miyagi Cohort; 3 pref. Miyagi, 3 prefecture Miyagi Study; Takayama, Takayama Study; KMCC, Korean Multi-center Cancer Cohort Study; SCHS, Singapore Chinese Health Study

Risk for death and premature death from CVD according to lifestyle factors and CMDs at baseline was shown in Table 9. Each element of HLS was independently associated with lower risk of death and premature death from cardiovascular disease. The ideal healthy status consisted of never smoking (HR 0.63, 95% CI: 0.60−0.65 for CVD death; HR 0.55, 95% CI: 0.51−0.59 for premature CVD death), never alcohol drinking (HR 0.93, 95% CI: 0.90−0.96 for CVD death; HR 0.84, 95% CI: 0.79−0.89 for premature CVD death), and BMI in the range of 20.0 to 27.4kg/m^2 (HR 0.78, 95% CI: 0.75−0.81 for CVD death; HR 0.73, 95% CI: 0.68−0.78 for premature CVD death) compared to the unhealthy status (ever smoking, ever alcohol drinking, and BMI in the rage of <18.5 or ≥27.5kg/m^2, respectively). In terms of CMDs, each of hypertension (HR=1.63, 95% CI: 1.58−1.69), DM (HR=1.63, 95% CI: 1.55−1.71), CHD (HR=1.67, 95% CI: 1.59−1.75), and stroke (HR=2.69, 95% CI: 2.54−2.86) was related to the increased risk of CVD death (Table 9 & Appendix12).

Table 9. Risk for total and premature cardiovascular death according to lifestyle factors and cardiometabolic diseases

Characteristics	**Cohort**	**CVD death (N=19,442)**		**Premature CVD death (N=5,774)**	
	N	**N**	**HR (95% CI)[2]**	**N**	**HR (95% CI)[1]**
Healthy lifestyle factors					
Cigarette smoking					
Never	243,481	8,617	1.00	2,491	1.00
Past	40,310	2,682	1.20 (1.14-1.26)	673	1.24 (1.12-1.37)
Current	120,061	5,853	1.77 (1.70-1.85)	2,610	2.01 (1.87-2.17)
[Unhealthy]: Ever	160,371	8,535	1.00	3,283	1.00
[Healthy]: Never	243,481	8,617	0.63 (0.60-0.65)	2,491	0.55 (0.51-0.59)
Alcohol drinking					
Never	233,625	8,829	1.00	2,629	1.00
Past	9,294	1,006	1.57 (1.46-1.68)	261	2.01 (1.76-2.29)
Current	160,933	7,317	1.03 (0.99-1.07)	2,884	1.17 (1.10-1.24)
[Unhealthy]: Ever	170,227	8,323	1.00	3,145	1.00
[Healthy]: Never	233,625	8,829	0.93 (0.90-0.96)	2,629	0.84 (0.79-0.89)
BMI (kg/m^2)					
<18.5	15,681	1,054	1.48 (1.39-1.58)	250	1.45 (1.27-1.66)
18.5-22.9	165,795	6,778	1.00	2,241	1.00
23.0-24.9	101,581	3,947	0.93 (0.89-0.97)	1,381	0.93 (0.87-0.99)
25.0-27.4	77,685	3,081	0.91 (0.87-0.95)	1,063	0.90 (0.84-0.97)
27.5-29.9	30,339	1,438	1.02 (0.96-1.08)	507	1.07 (0.97-1.18)
≥30.0	12,771	854	1.38 (1.28-1.48)	332	1.71 (1.52-1.92)
[Unhealthy]: <18.5 or ≥ 27.5	58,791	3,346	1.00	1,089	1.00
[Healthy]: 18.5-27.4	345,061	13,806	0.78 (0.75-0.81)	4,685	0.73 (0.68-0.78)
Prior cardiometabolic diseases at baseline					
Hypertension					
No	316,412	9,808	1.00	3,702	1.00
Yes	87,440	7,344	1.63 (1.58-1.68)	2,072	1.96 (1.85-2.09)
Diabetes mellitus					
No	383,363	15,068	1.00	5,094	1.00
Yes	20,489	2,084	1.63 (1.55-1.71)	680	2.17 (2.00-2.36)
Coronary heart disease					
No	388,605	15,165	1.00	5,318	1.00
Yes	15,247	1,987	1.67 (1.59-1.75)	456	1.95 (1.77-2.16)
Stroke					
No	397,968	15,847	1.00	5,441	1.00
Yes	5,884	1,305	2.69 (2.54-2.86)	333	3.75 (3.33-4.21)

Abbreviation: CVD, cardiovascular disease; N, number; HR, hazard ratio; BMI, body mass index

1. Adjusted for age, sex, cigarette smoking, alcohol drinking, BMI, hypertension, diabetes mellitus, chronic heart disease, and stroke, excluding each analysis variable.

HLS and cause−specific death according to the disease status

The association of HLS with CVD death according to the disease status at baseline was shown in Figure 14. We found that the increasing number of HLS was significantly associated with decreased risk of death from all−cause, CVD, and premature death regardless of CMDs at baseline (Figure 14 & Appendix 13). The HRs of CVD death according to each unit in the HLS were 0.75 (95% CI: 0.73−0.78) in individuals without CMDs, 0.78 (95% CI: 0.75−0.82) in those with HTN, 0.89 (95% CI: 0.81−0.99) in those with DM, 0.77 (95% CI: 0.70−0.86) in those with CHD, 0.87 (95% CI: 0.79−0.96) in those with HTN and DM, 0.73 (95% CI: 0.67−0.80) in those with HTN and CHD, and 0.86 (95% CI: 0.77−0.95) in those with HTN and stroke, and 0.76 (95% CI: 0.63−0.93) in those with HTN, DM, and CHD at baseline (Appendix 14).

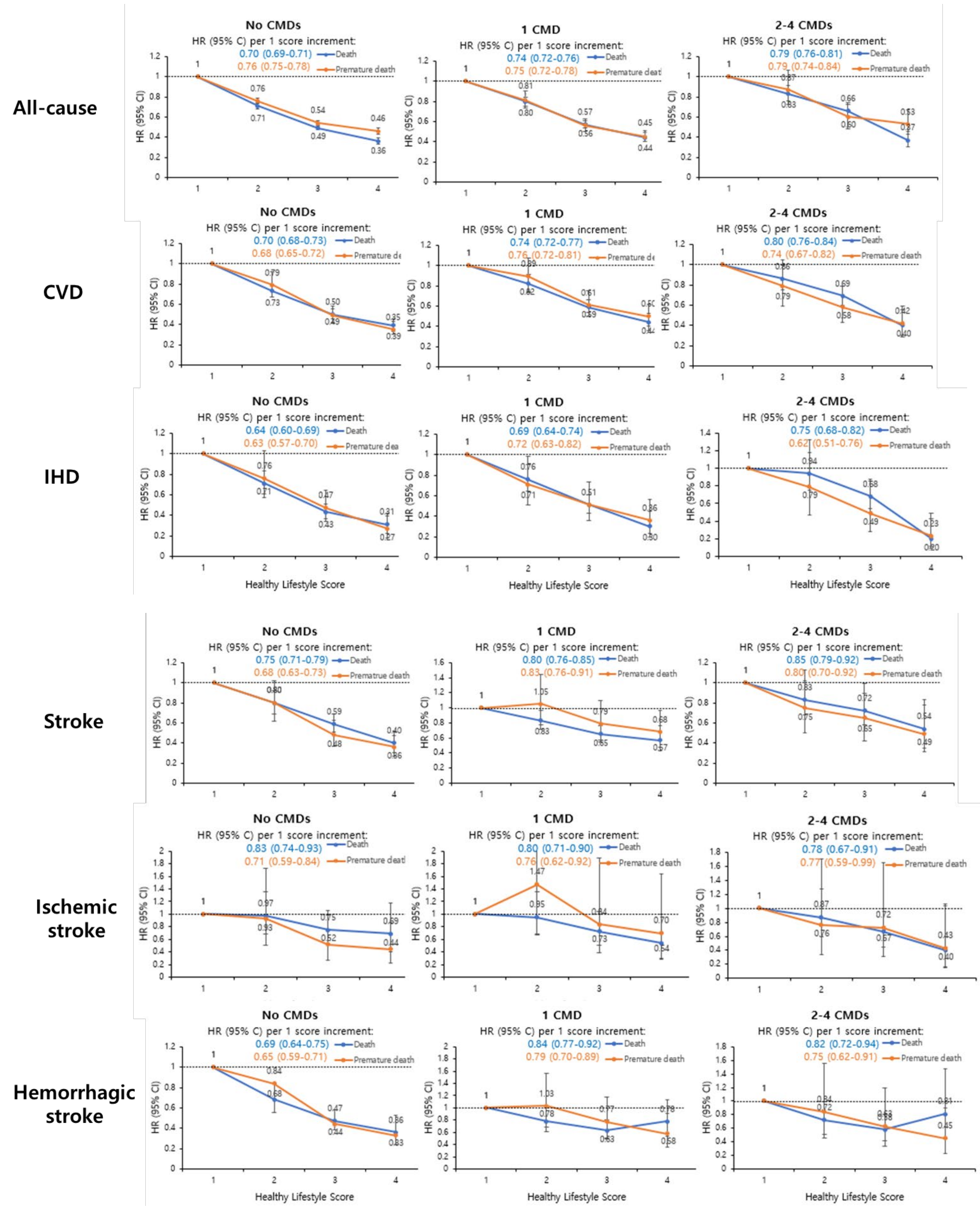


Figure 14. Association of healthy lifestyle score with all-cause and CVD-specific death according to disease status

Combination of healthy lifestyle factors with cause−specific death according to the number of CMDs

For individuals with one healthy lifestyle factors, non−smoking had the strongest association with decreasing the risk of all cause and CVD−specific death regardless of the number of CMDs at baseline. Compared to the individuals with none of healthy lifestyle factors, individuals who were non−drinking alcohol had a significant decrease in risk for death from all−cause (HR 0.84, 95% CI: 0.71−0.99), CVD (HR 0.73, 95% CI: 0.56−0.94), and especially death from stroke (HR 0.68, 95% CI: 0.46−0.99) among individuals with 2 or more CMDs at baseline. Among two healthy lifestyle factors, individuals who were non−smoking and had healthy BMI had the lowest risk of all−cause and CVD−specific death. When the impact of HLS was analyzed on individuals with multiple CMDs, at least two of healthy lifestyle factors were necessary to significantly decrease the risk of CVD−specific death. For individuals with cardiometabolic comorbidity, having three of healthy lifestyle factors was significantly associated with decrease in death from CVD (HR 0.51, 95% CI: 0.42−0.61), IHD (HR 0.47, 95% CI: 0.33−0.68), stroke (HR 0.54, 95% CI: 0.42−0.69), ischemic stroke (HR 0.53, 95% CI: 0.33−0.86), and hemorrhagic stroke (HR 0.39, 95% CI: 0.25−0.60) (Table 10).

Table 10. Association of combination of healthy lifestyle factors with all-cause and cardiovascular-specific death according to the number of cardiometabolic diseases at baseline

Healthy lifestyle factors	**Number of past cardiometabolic diseases at baseline**					
	No CMD		**1 CMD**		**2-4 CMDs**	
	Death	**HR (95% CI)[1]**	**Death**	**HR (95% CI)[1]**	**Death**	**HR (95% CI)[1]**
All-cause						
None	1,917	1.00	1,070	1.00	368	1.00
Non-smoking	602	0.56 (0.51-0.62)	422	0.57 (0.51-0.64)	111	0.53 (0.43-0.66)
Non-drinking	907	0.97 (0.90-1.05)	467	0.95 (0.85-1.06)	223	0.84 (0.71-0.99)
Healthy BMI	11,880	0.75 (0.71-0.78)	6,436	0.85 (0.79-0.90)	1,886	0.83 (0.74-0.92)
Non-smoking + Non-drinking	2,259	0.64 (0.60-0.68)	1,677	0.59 (0.55-0.65)	825	0.57 (0.50-0.65)
Non-smoking + Healthy BMI	3,376	0.42 (0.40-0.45)	1,778	0.48 (0.45-0.52)	461	0.50 (0.44-0.58)
Non-drinking + Healthy BMI	4,156	0.68 (0.64-0.71)	1,745	0.79 (0.73-0.85)	769	0.78 (0.69-0.88)
Non-smoking + Non-drinking + Healthy BMI	9,311	0.46 (0.43-0.48)	4,780	0.47 (0.44-0.51)	1,942	0.52 (0.46-0.59)
CVD						
None	417	1.00	329	1.00	163	1.00
Non-smoking	173	0.64 (0.54-0.7)	160	0.66 (0.55-0.80)	51	0.53 (0.39-0.73)
Non-drinking	214	1.01 (0.85-1.19)	154	0.97 (0.80-1.17)	90	0.73 (0.56-0.94)
Healthy BMI	2,603	0.78 (0.71-0.87)	1,951	0.84 (0.75-0.95)	756	0.76 (0.64-0.89)
Non-smoking + Non-drinking	577	0.61 (0.53-0.70)	630	0.64 (0.55-0.74)	388	0.57 (0.46-0.69)
Non-smoking + Healthy BMI	860	0.46 (0.40-0.51)	636	0.53 (0.46-0.60)	224	0.54 (0.44-0.66)
Non-drinking + Healthy BMI	908	0.67 (0.60-0.76)	595	0.84 (0.74-0.96)	355	0.78 (0.65-0.94)
Non-smoking + Non-drinking + Healthy BMI	2,264	0.43 (0.38-0.48)	1,761	0.51 (0.45-0.58)	893	0.51 (0.42-0.61)
Ischemic heart disease						
None	119	1.00	95	1.00	41	1.00
Non-smoking	36	0.55 (0.38-0.80)	36	0.57 (0.38-0.84)	15	0.68 (0.37-1.23)
Non-drinking	53	0.91 (0.66-1.26)	46	1.03 (0.72-1.47)	29	0.96 (0.60-1.55)
Healthy BMI	683	0.70 (0.58-0.85)	472	0.70 (0.56-0.88)	205	0.83 (0.59-1.16)
Non-smoking + Non-drinking	121	0.57 (0.43-0.75)	120	0.49 (0.37-0.66)	79	0.53 (0.35-0.80)
Non-smoking + Healthy BMI	177	0.37 (0.29-0.47)	121	0.38 (0.29-0.50)	48	0.50 (0.30-0.76)
Non-drinking + Healthy BMI	239	0.63 (0.51-0.78)	163	0.82 (0.63-1.05)	103	0.91 (0.63-1.31)
Non-smoking + Non-drinking + Healthy BMI	413	0.34 (0.27-0.43)	359	0.42 (0.32-0.54)	185	0.47 (0.33-0.68)

Table 10 (Continued). Association of combination of healthy lifestyle factors with all-cause and cardiovascular-specific death according to the number of cardiometabolic diseases at baseline

Healthy lifestyle factors	**Number of past cardiometabolic diseases at baseline**					
	No CMD		**1 CMD**		**2-4 CMDs**	
	Death	**HR (95% CI)[1]**	**Death**	**HR (95% CI)[1]**	**Death**	**HR (95% CI)[1]**
Stroke						
None	163	1.00	124	1.00	80	1.00
Non-smoking	74	0.68 (0.52-0.90)	75	0.81 (0.61-1.09)	20	0.43 (0.26-0.71)
Non-drinking	82	0.97 (0.74-1.27)	66	1.08 (0.80-1.46)	41	0.68 (0.46-0.99)
Healthy BMI	1,082	0.84 (0.71-0.99)	848	0.98 (0.81-1.18)	353	0.73 (0.57-0.93)
Non-smoking + Non-drinking	260	0.67 (0.54-0.83)	313	0.82 (0.65-1.02)	198	0.59 (0.45-0.78)
Non-smoking + Healthy BMI	379	0.50 (0.42-0.61)	298	0.65 (0.52-0.80)	118	0.59 (0.44-0.78)
Non-drinking + Healthy BMI	356	0.67 (0.56-0.81)	257	0.95 (0.77-1.18)	170	0.76 (0.58-0.99)
Non-smoking + Non-drinking + Healthy BMI	1,058	0.49 (0.41-0.59)	889	0.66 (0.54-0.81)	460	0.54 (0.42-0.69)
Ischemic Stroke						
None	28	1.00	35	1.00	23	1.00
Non-smoking	12	0.64 (0.33-1.28)	18	0.67 (0.38-1.19)	11	0.80 (0.38-1.65)
Non-drinking	19	1.35 (0.75-2.42)	17	0.99 (0.56-1.77)	10	0.55 (0.26-1.17)
Healthy BMI	234	1.10 (0.74-1.63)	234	0.96 (0.68-1.37)	109	0.80 (0.51-1.25)
Non-smoking + Non-drinking	54	0.86 (0.53-1.40)	77	0.70 (0.45-1.08)	56	0.56 (0.33-0.95)
Non-smoking + Healthy BMI	95	0.78 (0.50-1.19)	59	0.44 (0.29-0.67)	31	0.53 (0.30-0.91)
Non-drinking + Healthy BMI	93	1.05 (0.69-1.61)	64	0.84 (0.55-1.26)	50	0.76 (0.46-1.24)
Non-smoking + Non-drinking + Healthy BMI	262	0.77 (0.51-1.17)	241	0.62 (0.42-0.92)	137	0.53 (0.33-0.86)
Hemorrhagic Stroke						
None	87	1.00	51	1.00	29	1.00
Non-smoking	44	0.72 (0.50-1.04)	37	0.99 (0.64-1.51)	5	0.28 (0.11-0.72)
Non-drinking	33	0.70 (0.47-1.05)	24	0.93 (0.64-1.51)	17	0.73 (0.40-1.34)
Healthy BMI	523	0.74 (0.59-0.93)	355	1.01 (0.75-1.36)	139	0.83 (0.55-1.23)
Non-smoking + Non-drinking	99	0.42 (0.31-0.57)	126	0.80 (0.56-1.14)	69	0.47 (0.29-0.76)
Non-smoking + Healthy BMI	171	0.38 (0.29-0.50)	152	0.82 (0.60-1.14)	47	0.62 (0.39-1.00)
Non-drinking + Healthy BMI	132	0.45 (0.35-0.59)	92	0.83 (0.59-1.17)	49	0.61 (0.38-0.96)
Non-smoking + Non-drinking + Healthy BMI	442	0.33 (0.25-0.42)	331	0.60 (0.44-0.83)	140	0.39 (0.25-0.60)

Abbreviation: CMD, cardiometabolic disease; HR, hazard ratio; CVD, cardiovascular disease, BMI, body mass index

1. Adjusted for age and sex

The association of combination of HLS with premature all-cause and CVD-specific death according to the number of CMDs was presented in Table 11. For individuals with cardiometabolic comorbidity at baseline, at least two of healthy lifestyle factors were necessary to significantly decrease the risk of premature death from all-cause and CVD. Among individuals with 2 or more CMDs at baseline, non-smoking was significantly associated with lower risk of CVD death (HR 0.48, 95% CI: 0.27-0.87), especially death from stroke (HR 0.14, 95% CI: 0.03-0.57) compared to the people with none of healthy lifestyle factors. For combination of healthy lifestyle factors, similar stepwise decrease with increase in healthy lifestyle factors was observed with premature death from all-cause and CVD in individuals with varying number of CMDs (Table 11).

Table 11. Association of combination of healthy lifestyle factors with premature all-cause and cardiovascular-specific death according to the number of cardiometabolic diseases

Healthy lifestyle factors	Number of past cardiometabolic diseases at baseline					
	No CMD		1 CMD		2-4 CMDs	
	Pre death	HR (95% CI)[1]	Pre death	HR (95% CI)[1]	Pre death	HR (95% CI)[1]
All-cause						
None	821	1.00	402	1.00	117	1.00
Non-smoking	240	0.55 (0.47-0.63)	134	0.52 (0.43-0.64)	34	0.57 (0.39-0.84)
Non-drinking	396	0.94 (0.83-1.06)	178	0.90 (0.75-1.07)	68	0.87 (0.64-1.17)
Healthy BMI	5,520	0.76 (0.71-0.82)	2,116	0.83 (0.74-0.92)	571	0.90 (0.74-1.10)
Non-smoking + Non-drinking	927	0.67 (0.61-0.74)	468	0.55 (0.47-0.63)	185	0.53 (0.41-0.68)
Non-smoking + Healthy BMI	1,543	0.43 (0.39-0.47)	491	0.41 (0.36-0.46)	108	0.45 (0.35-0.59)
Non-drinking + Healthy BMI	1,719	0.61 (0.56-0.67)	582	0.74 (0.65-0.84)	188	0.75 (0.59-0.94)
Non-smoking + Non-drinking + Healthy BMI	4,009	0.45 (0.41-0.49)	1,302	0.42 (0.37-0.48)	438	0.49 (0.39-0.62)
CVD						
None	172	1.00	122	1.00	55	1.00
Non-smoking	70	0.69 (0.52-0.91)	51	0.66 (0.47-0.91)	14	0.48 (0.27-0.87)
Non-drinking	90	0.99 (0.77-1.29)	71	1.16 (0.86-1.55)	27	0.70 (0.44-1.11)
Healthy BMI	1,177	0.79 (0.67-0.92)	686	0.88 (0.73-1.07)	247	0.83 (0.62-1.11)
Non-smoking + Non-drinking	170	0.48 (0.38-0.61)	160	0.60 (0.47-0.78)	87	0.48 (0.33-0.70)
Non-smoking + Healthy BMI	339	0.41 (0.34-0.49)	163	0.44 (0.35-0.56)	49	0.42 (0.29-0.63)
Non-drinking + Healthy BMI	348	0.59 (0.49-0.70)	195	0.80 (0.64-1.00)	93	0.75 (0.53-1.04)
Non-smoking + Non-drinking + Healthy BMI	764	0.34 (0.28-0.41)	450	0.47 (0.37-0.59)	174	0.38 (0.27-0.54)
Ischemic heart disease						
None	50	1.00	43	1.00	17	1.00
Non-smoking	13	0.51 (0.28-0.95)	12	0.46 (0.24-0.88)	8	0.96 (0.41-2.25)
Non-drinking	26	1.01 (0.63-1.62)	34	1.61 (1.03-2.53)	9	0.76 (0.34-1.71)
Healthy BMI	344	0.78 (0.58-1.05)	183	0.66 (0.47-0.92)	73	0.79 (0.46-1.34)
Non-smoking + Non-drinking	36	0.47 (0.29-0.74)	33	0.40 (0.24-0.66)	20	0.41 (0.20-0.87)
Non-smoking + Healthy BMI	72	0.35 (0.24-0.50)	40	0.32 (0.21-0.50)	9	0.27 (0.12-0.61)
Non-drinking + Healthy BMI	101	0.58 (0.42-0.82)	66	0.77 (0.52-1.13)	28	0.72 (0.39-1.32)
Non-smoking + Non-drinking + Healthy BMI	132	0.27 (0.18-0.38)	101	0.33 (0.22-0.50)	31	0.25 (0.13-0.49)

Table 11 (Continued). Association of combination of healthy lifestyle factors with premature all-cause and cardiovascular-specific death according to the number of cardiometabolic diseases

Healthy lifestyle factors	**Number of past cardiometabolic diseases at baseline**					
	No CMD		**1 CMD**		**2-4 CMDs**	
	Pre death	**HR (95% CI)[1]**	**Pre death**	**HR (95% CI)[1]**	**Pre death**	**HR (95% CI)[1]**
Stroke						
None	71	1.00	44	1.00	27	1.00
Non-smoking	34	0.76 (0.50-1.15)	27	0.95 (0.58-1.54)	2	0.14 (0.03-0.57)
Non-drinking	34	0.90 (0.60-1.36)	23	1.02 (0.61-1.69)	12	0.63 (0.32-1.24)
Healthy BMI	489	0.79 (0.62-1.02)	295	1.06 (0.78-1.46)	120	0.82 (0.54-1.25)
Non-smoking + Non-drinking	85	0.52 (0.37-0.74)	87	0.87 (0.59-1.30)	49	0.50 (0.29-0.85)
Non-smoking + Healthy BMI	151	0.41 (0.31-0.55)	80	0.60 (0.41-0.87)	29	0.49 (0.29-0.84)
Non-drinking + Healthy BMI	134	0.54 (0.41-0.73)	86	0.96 (0.67-1.39)	50	0.82 (0.51-1.31)
Non-smoking + Non-drinking + Healthy BMI	367	0.35 (0.27-0.47)	240	0.67 (0.47-0.95)	103	0.42 (0.26-0.68)
Ischemic stroke						
None	11	1.00	7	1.00	7	1.00
Non-smoking	7	0.93 (0.36-2.45)	5	0.95 (0.30-3.04)	0	0.18 (0.01-2.63)
Non-drinking	6	0.9 (0.37-2.70)	5	1.30 (0.41-4.11)	3	0.58 (0.15-2.24)
Healthy BMI	87	0.92 (0.49-1.72)	69	1.54 (0.71-3.3)	33	0.86 (0.38-1.94)
Non-smoking + Non-drinking	10	0.35 (0.14-0.87)	13	0.60 (0.22-1.60)	12	0.44 (0.16-1.25)
Non-smoking + Healthy BMI	28	0.47 (0.23-0.96)	17	0.68 (0.27-1.66)	13	0.82 (0.32-2.11)
Non-drinking + Healthy BMI	25	0.65 (0.32-1.31)	18	1.19 (0.50-2.76)	13	0.78 (0.31-1.95)
Non-smoking + Non-drinking + Healthy BMI	72	0.40 (0.20-0.81)	47	0.61 (0.26-1.44)	24	0.35 (0.14-0.90)
Hemorrhagic stroke						
None	43	1.00	25	1.00	12	1.00
Non-smoking	25	0.89 (0.54-1.47)	17	1.03 (0.55-1.92)	2	0.27 (0.06-1.21)
Non-drinking	16	0.70 (0.39-1.24)	13	1.00 (0.51-1.97)	6	0.68 (0.25-1.81)
Healthy BMI	313	0.84 (0.61-1.16)	160	1.03 (0.55-1.92)	59	0.93 (0.50-1.72)
Non-smoking + Non-drinking	50	0.49 (0.31-0.75)	60	1.01 (0.61-1.69)	29	0.51 (0.24-1.10)
Non-smoking + Healthy BMI	92	0.40 (0.28-0.58)	45	0.58 (0.35-0.96)	9	0.31 (0.13-0.75)
Non-drinking + Healthy BMI	69	0.47 (0.32-0.68)	42	0.84 (0.51-1.38)	25	0.92 (0.46-1.84)
Non-smoking + Non-drinking + Healthy BMI	219	0.33 (0.23-0.47)	129	0.60 (0.38-0.97)	52	0.38 (0.19-0.77)

Abbreviation: CMD, cardiometabolic disease; HR, hazard ratio; CVD, cardiovascular disease, BMI, body mass index

1. Adjusted for age and sex

3.4. Change in lifestyle factors study

General characteristics

The general characteristics of the 4,638, 6,709, and 5,262 total population for HTN, DM, and MetS were presented respectively in Table 12. Among the 4,638 participants for HTN, the mean age was 50.1 years, 47.5% were men, and 1,414 HTN events (30.5%) occurred. Among the 6,709 participants for DM, the mean age was 51.8 years, 48.0% were men, and 732 DM events (10.9%) occurred. Among the 3,292 participants for MetS, the mean age was 49.5 years, 53.8% were men, and 1,060 MetS events (32.2%) occurred (Table 12).

Table 12. General characteristics of the study population for hypertension, diabetes mellitus, and metabolic syndrome in the Ansan and Ansung study

	Baseline population		
	Participants without HTN (N=4,638)	**Participants without DM (N=6,709)**	**Participants without MetS (N=3,292)**
Age, years, mean (SD)	50.1 (8.29)	51.8 (8.76)	49.5 (8.2)
Sex, n (%)			
Men	2,204 (47.5)	3,222 (48.0)	1,771 (53.8)
Women	2,434 (52.5)	3,488 (52.0)	1,521 (46.2)
Education, n (%)			
Elementary school	1,185 (25.6)	2,118 (31.6)	677 (20.6)
High school	2,727 (58.8)	3,637 (54.2)	2,006 (60.9)
College and more	706 (15.2)	921 (13.7)	601 (18.3)
Income, n (%)			
<1,000K/month	1,293 (27.9)	2,239 (33.4)	773 (23.5)
1,000-2,000K	1,402 (30.2)	1,969 (29.4)	974 (29.6)
2,000-4,000K	1,483 (32.0)	1,901 (28.3)	1,171 (35.6)
≥4,000K	407 (8.8)	513 (7.7)	346 (10.5)
BMI, kg/m^2, mean (SD)	24.1 (2.96)	24.4 (3.07)	23.6 (2.7)
Waist circumstance, cm, mean (SD)	80.8 (8.28)	82.2 (8.64)	78.7 (7.5)
Smoking, n (%)			
Never	2,760 (59.5)	4,008 (59.7)	1,866 (56.7)
Ever	678 (14.6)	1,062 (15.8)	557 (16.9)
Current	1,200 (25.9)	1,639 (24.4)	869 (26.4)
Alcohol drinking, n (%)			
Never	2,35 (46.0)	3,073 (45.8)	1,372 (41.7)
Ever	280 (6.1)	399 (6.0)	178 (5.4)
Current	2,223 (47.9)	3,237 (48.2)	1,742 (52.9)
Physical activity, n (%)			
No	2,778 (59.9)	3,808 (56.8)	2,160 (65.6)
Yes	1,860 (40.1)	2,901 (43.2)	1,132 (34.4)
Family history of CVD, n (%)			
No	4,397 (94.8)	6,363 (94.8)	3,119 (94.7)
Yes	241 (5.2)	346 (5.2)	173 (5.3)
Total-cholesterol, n (%)			
<240	4,295 (92.6)	6,179 (92.1)	3,038 (92.3)
≥240	343 (7.4)	529 (7.9)	254 (7.7)

Smoking

After adjustment for age, sex, education level, income level, alcohol drinking, physical activity, BMI, total cholesterol, and family history of CVD, individuals who were continuously maintained their dose of cigarette smoking had a significant increase in risk for HTN (HR 1.26, 95% CI: 1.01–1.59) compared to persistent never smokers. For DM, individuals who were continuously smoking had an increase in risk for DM (dose decreased, HR 1.82, 95% CI: 1.27–2.60; maintained, HR 2.06, 95% CI: 1.51–2.81; increased, HR 2.06, 95% CI: 1.51–2.19). For MetS, individuals who were continuously smoking had an increase in risk for MetS (dose decreased, HR 1.67, 95% CI: 1.26–2.21; maintained, HR 1.49, 95% CI: 1.15–1.93; increased, HR 1.49, 95% CI: 1.09–2.03) (Figure 15).

Figure 16 represented the results of the risk of HTN, DM, and MetS based on the intensity of smoking in cigarettes per day. Compared to the individuals who were persistent light/moderate smoker, the participants who increased their dose of smoking from light/moderate to heavy had a significantly increased risk for HTN (HR 1.65, 95% CI: 1.08–2.53) (Figure 16). Moreover, a significant increase in the risk of DM and MetS was observed in the “fall and rise in constantly smoker” trajectory compared to the “never smoker” trajectory (Appendix 15).

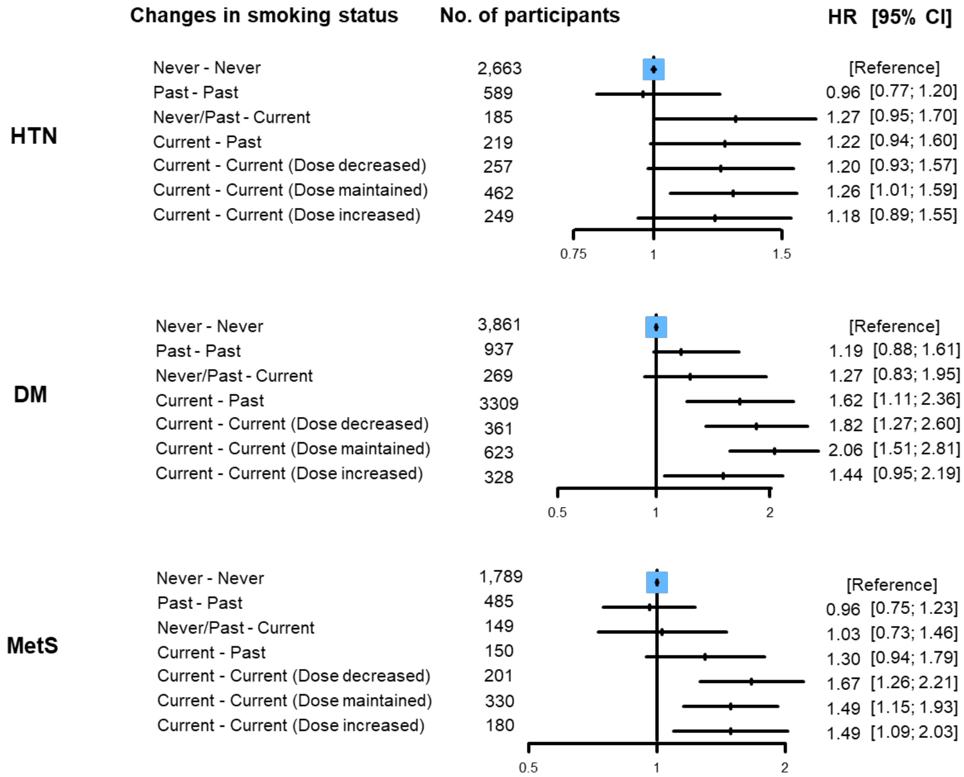


Figure 15. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to change in cigarette smoking status. (Hazard ratios are adjusted for age, sex, education, income, alcohol drink, physical activity, body mass index, total cholesterol level, and family history of cardiovascular disease)

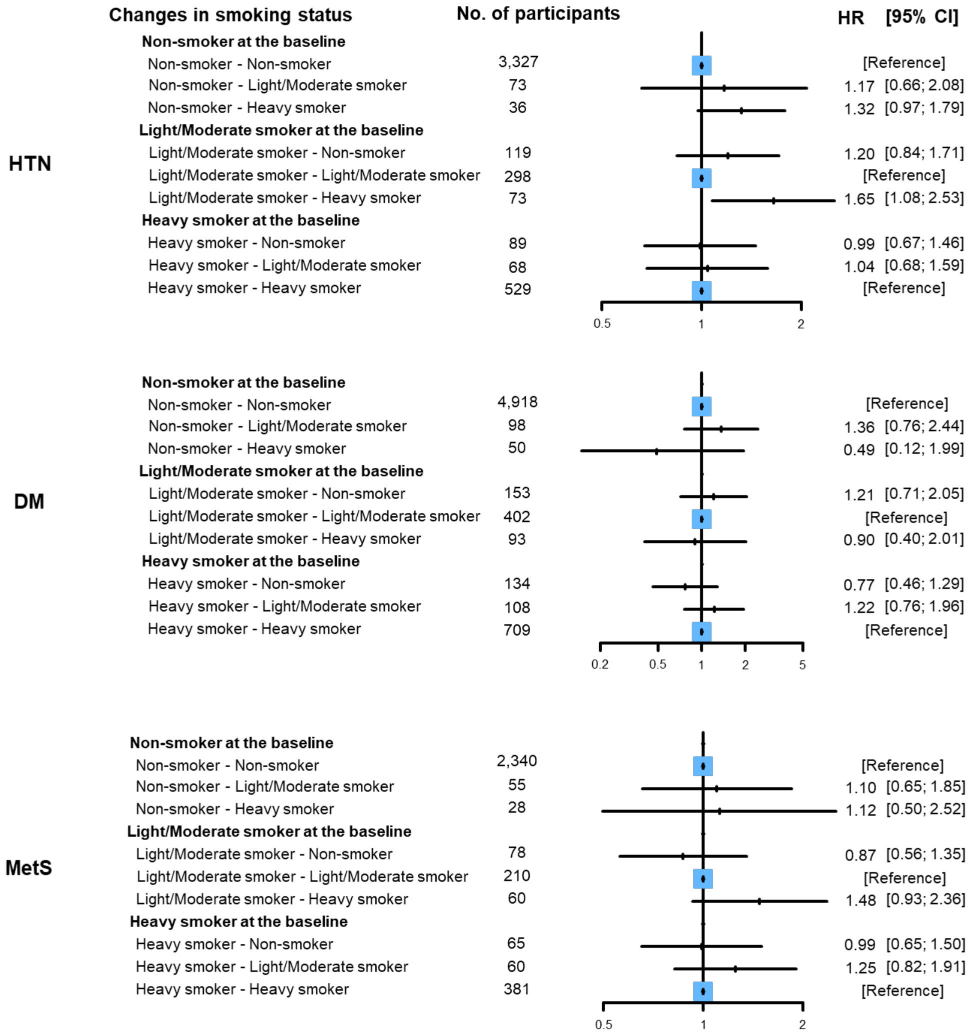


Figure 16. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to intensity of smoking in cigarettes per day (Hazard ratios are adjusted for age, sex, education, income, alcohol drink, physical activity, body mass index, total cholesterol level, and family history of cardiovascular disease)

Alcohol drinking

After multivariable adjustment for age, sex, education level, income level, smoking, physical activity, BMI, total cholesterol, and family history of CVD, individuals who were continuously maintained their dose of alcohol consumption had a significant increase in risk for HTN (HR 1.78, 95% CI: 1.28−2.46) and DM (HR 1.66, 95% CI: 1.08−2.56) compared to the never alcohol drinkers. While individuals who were decreased their dose of alcohol consumption had a significant increase in risk for MetS (HR 1.25, 95% CI: 1.04−1.51) compared to the never alcohol drinkers (Figure 17).

Compared to the individuals with continuously light/moderate alcohol consumption, the participants who increased their intensity of consumption from light/moderate to heavy had a significantly increased risk for HTN (HR 1.35, 95% CI: 1.06−1.72) DM (HR 1.43, 95% CI: 1.04−1.97), and MetS (HR 1.42, 95% CI: 1.10−1.84). Moreover, individuals who decreased their intensity of consumption from light/moderate to non−drinker had a significantly increased risk for DM (HR 1.62, 95% CI: 1.18−2.23) (Figure 18). Moreover, a significant increase in the risk of MetS was observed in the “fall and rise alcohol consumption in current drinker” trajectory compared to the never drinker trajectory (Appendix 16).

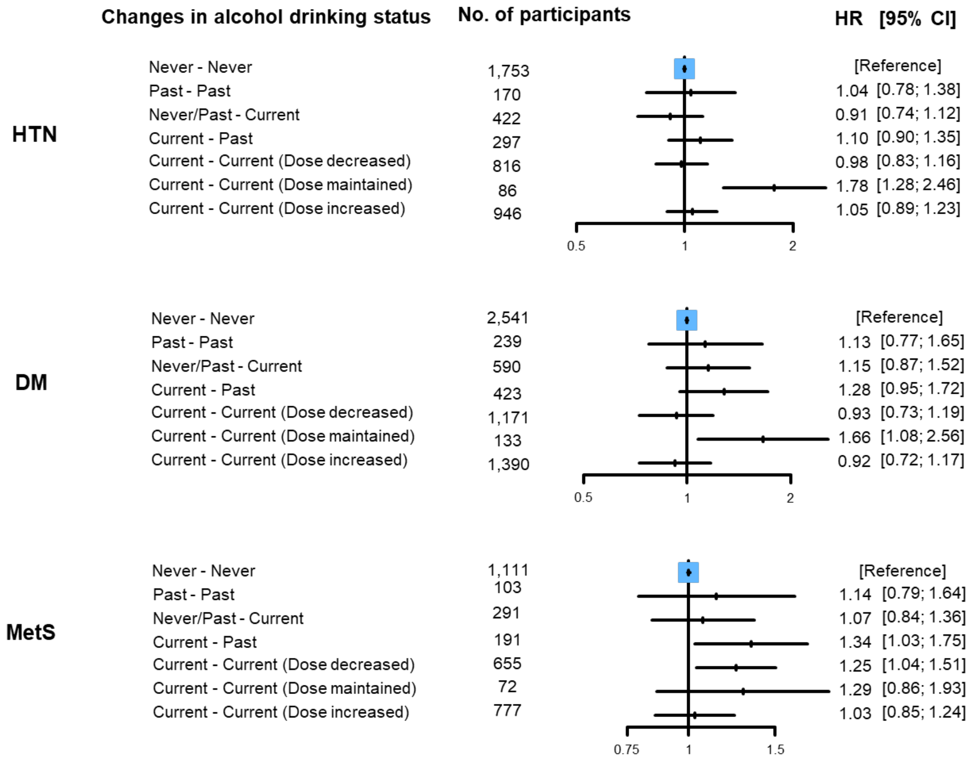


Figure 17. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to change in alcohol drinking status (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, physical activity, body mass index, total cholesterol level, and family history of cardiovascular disease)

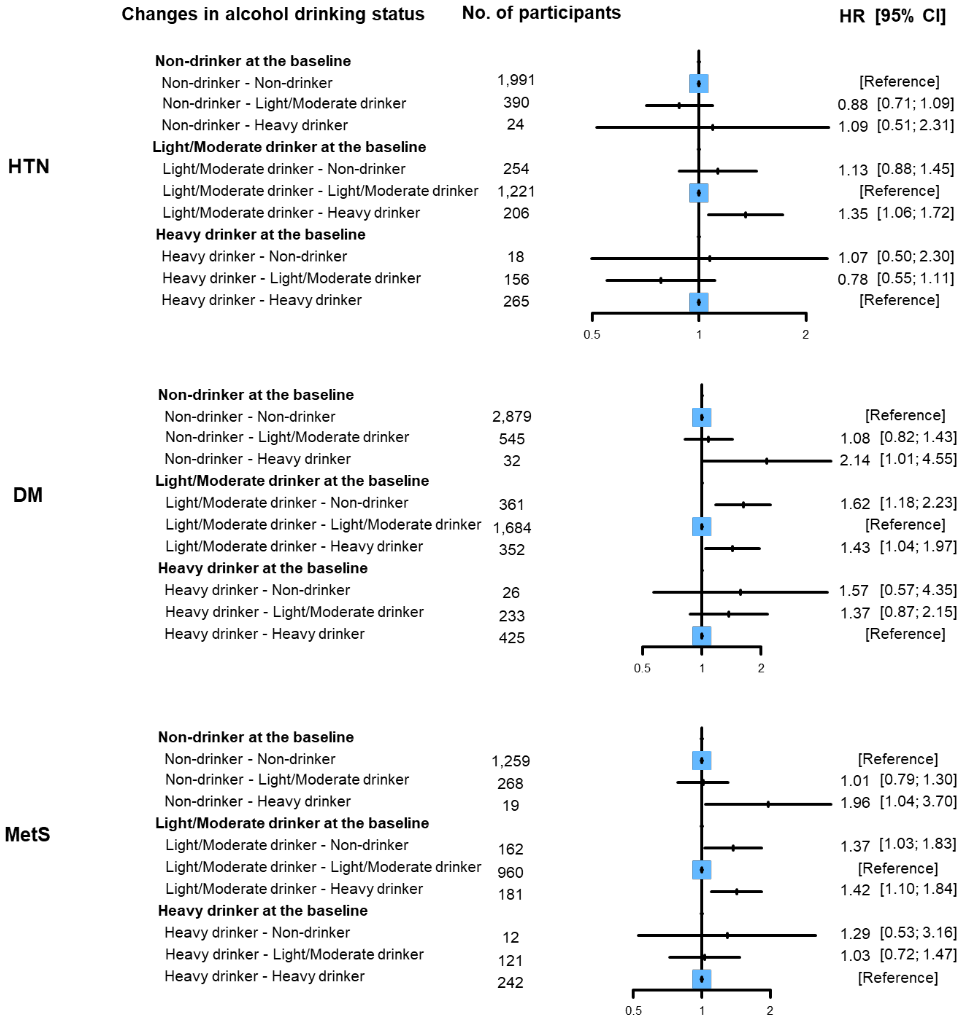


Figure 18. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to intensity of alcohol consumption per day (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, physical activity, body mass index, total cholesterol level, and family history of cardiovascular disease)

Physical activity

After adjustment for age, sex, education level, income level, smoking, alcohol drinking, BMI, total cholesterol, and family history of CVD, individuals became physically inactive in the second examination from the physically active in the first examination period had a significant increase in risk for DM (HR 1.25, 95% CI: 1.03–1.51) and MetS (HR 1.29, 95% CI: 1.08–1.54) compared to the individuals with persistent physically inactive (Figure 19). We also found inverse associations for decreasing physical activity trajectory against DM and MetS (Appendix 17).

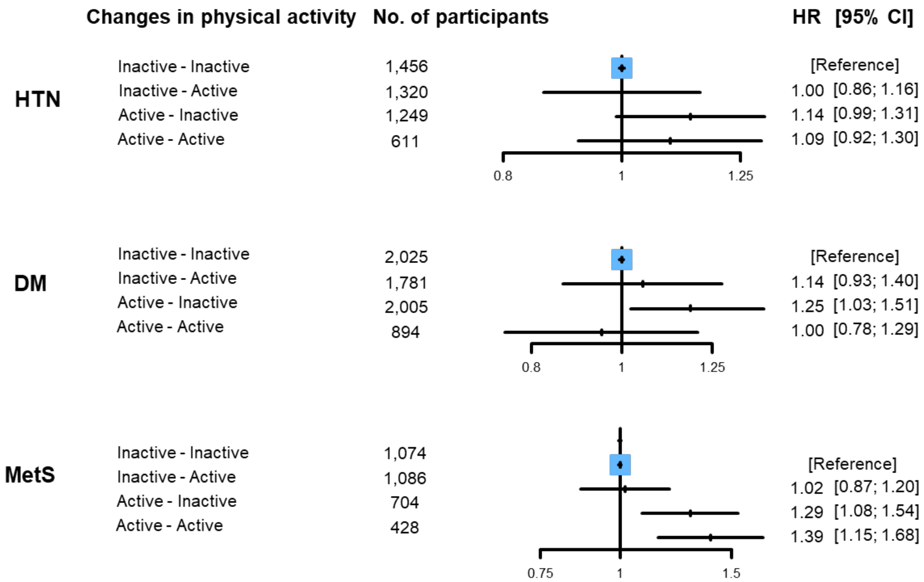


Figure 19. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to change in physical activity (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, alcohol drink, body mass index, total cholesterol level, and family history of cardiovascular disease)

Obesity

After adjustment for age, sex, education level, income level, smoking, alcohol drinking, physical activity, total cholesterol, and family history of CVD, participants who newly became BMI ≥ 25kg/m^2 in the second examination from the normal BMI range of 18.5–25kg/m^2 in the first examination period had a significant increase in risk for HTN (HR 1.34, 95% CI: 1.04–1.72), DM (HR: 2.00, 95% CI: 1.45–2.75), and MetS (HR 1.88, 95% CI: 1.44–2.45) compared to the individuals continuously had normal BMI. On the other hand, compared to the participants with continuously BMI≥ 25kg/m^2, individuals became normal BMI in the second examination from the BMI ≥ 25kg/m^2 in the first examination period had a significant decrease in risk for HTN (HR 1.57, 95% CI: 0.43–0.75), DM (HR: 0.45, 95% CI: 0.30–0.67), and MetS (HR 0.52, 95% CI: 0.39–0.70) (Figure 20). Similar results were observed in BMI trajectories over time (Appendix 18).

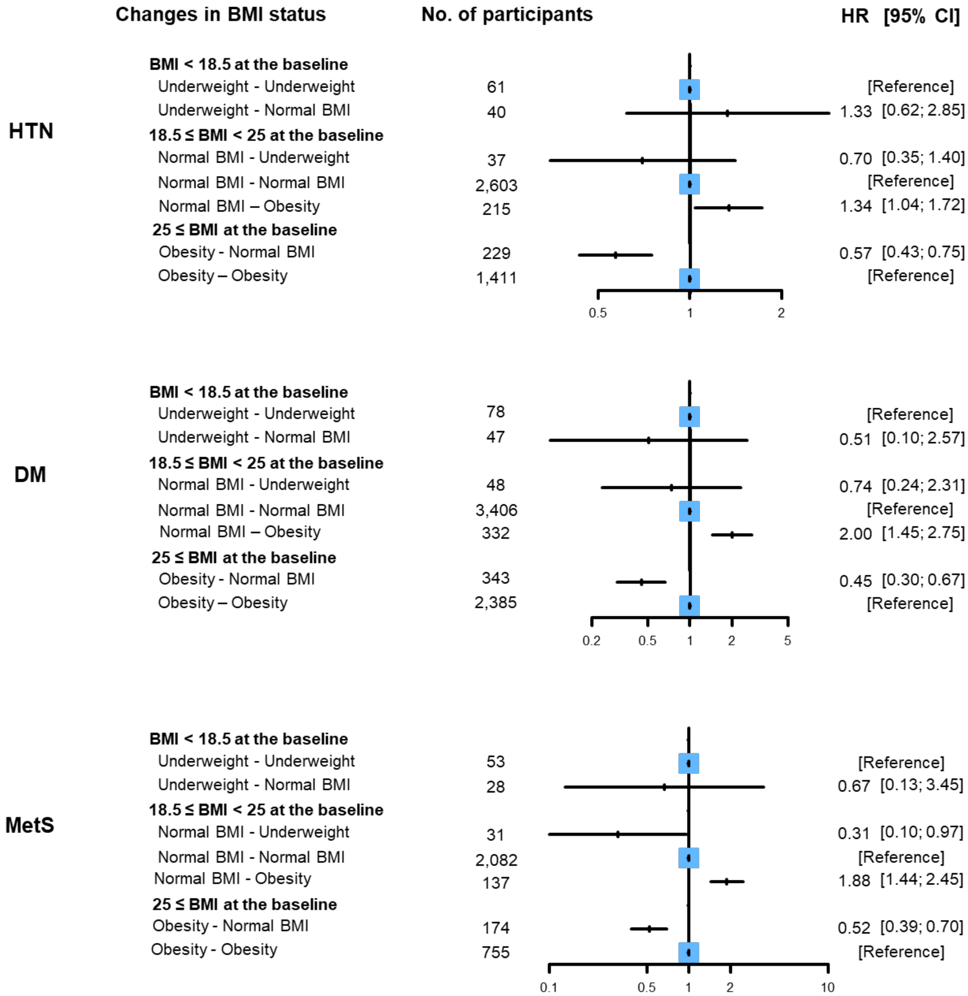


Figure 20. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to change in BMI status (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, alcohol drink, physical activity, total cholesterol level, and family history of cardiovascular disease)

Compared to the individuals with continuously normal waist size, those who newly became abnormal obesity had a significant increase in risk for HTN (HR 1.41, 95% CI: 1.18−1.69), and DM (HR: 2.48, 95% CI: 2.00−3.07). On the other hand, individuals became normal waist size from the abdominal obesity had a significant decrease in risk for HTN (HR 1.57, 95% CI: 0.43−0.75), and DM (HR: 0.45, 95% CI: 0.30−0.67) compared to those with continuously abdominal obesity (Figure 21). Similar results were observed in trajectory of waist size over time (Appendix 19).

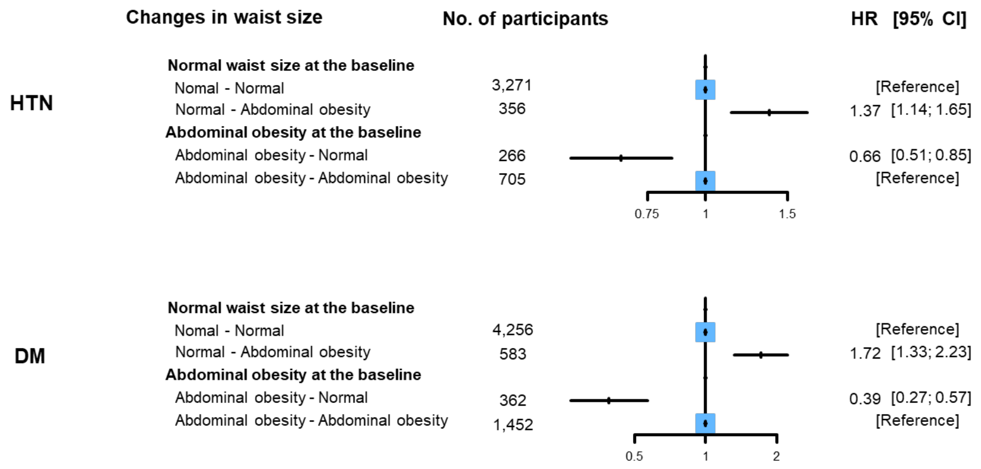


Figure 21. Adjusted hazard ratios for hypertension and diabetes mellitus according to change in waist size (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, alcohol drink, physical activity, total cholesterol level, and family history of cardiovascular disease)

3.5. Biological age study

This study was published in An et al. (2022) [S. An, C. Ahn, S. Moon, EJ Sim, SK. Park, "Individualized Biological Age as a Predictor of Disease: Korean Genome and Epidemiology Study (KoGES) Cohort", "*Journal of Personalized Medicine*", 2022, 12 (3), 505].

General characteristics

A total of 101,980 healthy participants (Charlson's comorbidity index of '0') aged 40–89 years were included to calculate the BA. More than a half (65.4%) was women and the mean age at baseline was 53.0 and 51.9 years for men and women, respectively (Table 13). Among them, 58,801 individuals had repeated measurements after a median 5 years of follow-up of 5 (range: 2–13). Among them, 2,474 subjects, 7,274 subjects, and 535 subjects were newly identified having DM, HTN, and combination of DM and HTN, respectively.

Table 13. Baseline characteristics of healthy participants at the baseline in the Korean Genome and Epidemiology Study

Variables	**Cohort participants with CCI=0 at baseline (n=101,980)**		**Non-diabetes cohort participants at baseline (n=41,714)***		**Non-hypertension cohort participants at baseline (n=22,717)***	
	Men (n=35,331)	**Women (n=66,649)**	**Men (n=13,693)**	**Women (n=28,021)**	**Men (n=5,733)**	**Women (n=16,984)**
	Mean ± SD	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**	**Mean ± SD**
Age, years	53.0 ± 8.58	51.9 ± 7.97	53.9 ± 8.41	52.1 ± 7.78	53.5 ± 8.35	50.7 ± 7.50
Height, cm	168.7 ± 5.84	156.3 ± 5.43	168.6 ± 5.78	156.3 ± 5.38	168.7 ± 5.75	156.7 ± 5.30
Weight, kg	69.5 ± 9.33	57.9 ±7.74	69.4 ± 9.01	57.8 ± 7.63	67.9 ± 8.76	56.9 ± 7.24
Waist size, cm	85.5 ± 7.50	78.5 ± 8.29	85.3 ± 7.33	78.3 ± 8.31	84.0 ± 7.29	76.9 ± 7.89
Hip size, cm	95.7 ± 5.69	93.5 ± 5.75	95.7 ± 5.53	93.3 ± 5.65	95.0 ± 5.49	92.8 ± 5.49
	No. (%)	**No. (%)**	**No. (%)**	**No. (%)**	**No. (%)**	**No. (%)**
College or more	12,673 (35.9)	12,939 (19.4)	5,115 (37.4)	5,501 (19.6)	2,282 (39.8)	3,934 (23.2)
Have occupation	30,281 (85.7)	29,523 (44.3)	11,334 (82.8)	11,667 (41.6)	4,808 (83.9)	7,306 (43.0)
Income ≥ $4,000	9,471 (26.8)	15,120 (22.7)	3,634 (26.5)	6,437 (23.0)	1,616 (28.2)	4,532 (26.7)
Current smokers	11,801 (33.4)	1,496 (2.3)	3,885 (28.4)	438 (1.6)	1,779 (31.0)	308 (1.8)
Current drinkers	26,321 (74.5)	22,299 (33.5)	10,114 (73.9)	8,788 (31.4)	4,000 (69.8)	5,483 (32.3)
Regular exercise	18,928 (53.6)	32,295 (48.5)	7,825 (57.2)	14,456 (51.6)	3,260 (56.9)	8,810 (51.9)

Calculation of biological age

Based on the differences between men and women, we calculated a sex−specific BA. In this study, we calculated the BA using self−assessed questionnaire (Appendix 20−25). According to the elastic net regression variable selection process, a total of 20 and 23 predictors were selected for men and women, respectively. Among them, we found that waist size, alcohol consumption, and the smoking duration were positively associated with BA (Appendix 23). We also confirmed that the BA was significantly correlated with CA for men ($r = 0.709$, R−square = 0.502, $p < 0.001$) and women ($r = 0.688$, R−square = 0.473, $p < 0.001$), respectively (Figure 22).

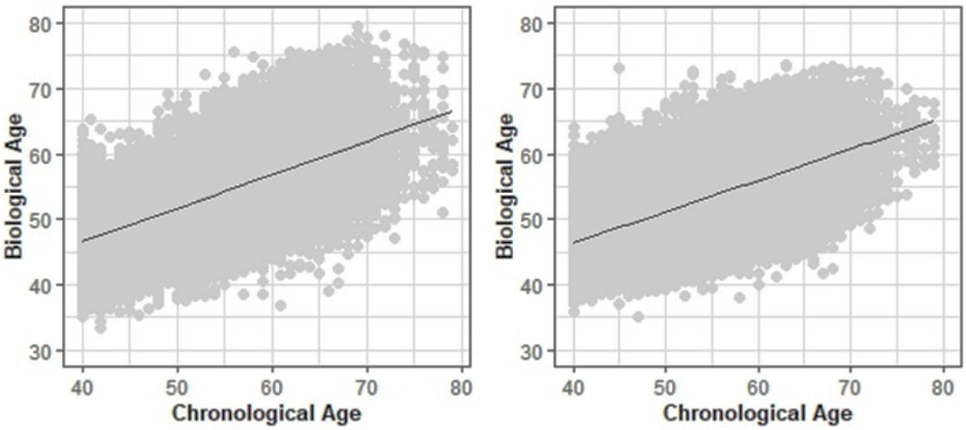


Figure 22. Relation of biological age and chronological age for men and women

Assessment of Biological age

We found that individuals in oldest CA group (≥70 years) had greater odds of DM (OR: 2.48, 95% CI: 1.93–3.17), HTN (OR: 2.66, 95% CI: 2.36–3.00), and comorbidity of DM and HTN (OR: 3.42, 95% CI: 2.44–4.80) compared to individuals in the youngest CA group (<50 years). As the BA increased by 1 year, the odds were increased by 6% for DM (OR: 1.06, 95% CI: 1.06–1.07), 7% for HTN (OR: 1.07, 95% CI: 1.07–1.08), and 10% for comorbidity of DM and HTN (OR: 1.10, 95% CI: 1.10–1.11). According to the Age–Diff, we found that Very young BA" group had the lowest odds of DM (OR: 0.72, 95% CI: 0.65–0.81), HTN (OR: 0.7, 95% CI: 0.68–0.75), and comorbidity of DM and HTN (OR: 0.65, 95% CI: 0.56–0.76) than those in "Same BA as CA" group (Table 14).

Table 14. Association of chronological age, biological age, and age-difference on the prevalence of diabetes mellitus, hypertension, and comorbidity of diabetes mellitus and hypertension

	Total Cohort N	Chronological Age (CA)		Biological Age (BA) [1]		Age-Diff (BA-CA) [2]		
		Cases N	OR (95% CI) [3]	Cases N	OR (95% CI) [3]	Age-Diff [2]	Cases N	OR (95% CI) [4]
DM								
<50	41,156	915	1.00	854	1.00	Very young BA	759	0.72 (0.65–0.81)
50–59	38,767	1405	1.65 (1.52–1.80)	1970	1.70 (1.57–1.85)	Young BA	788	0.88 (0.79–0.98)
60–69	20,821	1085	2.32 (2.12–2.54)	634	2.76 (2.48–3.06)	Same BA as CA	676	1.00
≥70	1236	72	2.48 (1.93–3.17)	19	2.76 (1.72–4.43)	Older BA	1254	1.17 (1.06–1.29)
Per 1-year increment	101,980	3477	1.04 (1.04–1.05)	3477	1.06 (1.06–1.07)		3477	*p–trend* < 0.001
HTN								
<50	41,156	15,977	1.00	14,915	1.00	Very young BA	10,037	0.72 (0.68–0.75)
50–59	38,767	19,873	1.68 (1.63–1.73)	27,683	1.78 (1.73–1.83)	Young BA	10,876	0.89 (0.85–0.92)
60–69	20,821	12,908	2.50 (2.41–2.59)	6777	2.95 (2.81–3.08)	Same BA as CA	9496	1.00
≥70	1236	804	2.66 (2.36–3.00)	187	2.99 (2.27–3.93)	Older BA	19,153	1.24 (1.19–1.29)
Per 1-year increment	101,980	49,562	1.05 (1.04–1.05)	49,562	1.07 (1.07–1.08)		49,562	*p–trend* < 0.001
Comorbidity of DM and HTN								
<50	41,156	576	1.00	521	1.00	Very young BA	535	0.65 (0.56–0.76)
50–59	38,767	944	2.13 (1.92–2.37)	1350	2.36 (2.12–2.61)	Young BA	537	0.86 (0.75–0.99)
60–69	20,821	772	3.85 (3.44–4.31)	449	5.12 (4.48–5.85)	Same BA as CA	431	1.00
≥70	1236	40	3.42 (2.44–4.80)	12	5.14 (2.75–9.62)	Older BA	903	1.39 (1.21–1.61)
Per 1-year increment	101,980	2332	1.07 (1.06–1.07)	2332	1.10 (1.10–1.11)		2332	*p–trend* < 0.001

Abbreviations: CA, chronological age; BA, biological age; DM, diabetes mellitus; HTN, hypertension; KOGES, Korean Genome and Epidemiology Study
1. BA using sex-specific Elastic net model**;** 2. BA-CA difference was classified into four groups: [Very young BA] BA was at least 5-year younger than CA; [Young BA] BA was between 1-year and < 5-year younger than CA; [Same BA as CA] BA-CA difference was between -1 year and 1 year; [Older BA] BA was at least 1 year older than CA (> 1 year); 3. Adjusted for sex; 4. Adjusted for sex and chronological age.

We found that individuals in the highest CA group, the risk was 1.88−fold for DM (95% CI: 1.28–2.76), 1.57−fold for HTN (95% CI: 1.19–2.07), and 2.21−fold for comorbidity of DM and HTN (95% CI: 0.82–5.99), while those in the "Older BA" group, the risk was 2.68−fold for DM (95% CI: 1.44–5.02), 2.48−fold for HTN (95% CI: 1.49–4.11), and 5.98−fold for the comorbidity of DM and HTN (95% CI: 0.83–43.01). Compared to the reference group, "Very young BA" group had the lowest risk of DM (HR: 0.63, 95% CI: 0.55–0.72), HTN (HR: 0.74, 95% CI: 0.68–0.81), and comorbidity of DM and HTN (HR: 0.65, 95% CI: 0.47–0.91). On the other hand, the "Older BA" group showed the highest risk of DM (HR: 1.20, 95% CI: 1.07 – 1.35), HTN (HR: 1.15, 95% CI: 1.07 – 1.23), and comorbidity of DM and HTN (HR: 1.32, 95% CI: 1.01–1.74) (Table 15). We also confirmed a consistent association within 5 follow−up years. The "Very young BA" group showed a significantly lower risk of DM (HR: 0.66, 95% CI: 0.54–0.80), HTN (HR: 0.74, 95% CI: 0.67–0.82), and comorbidity of DM and HTN (HR: 0.77, 95% CI: 0.47–1.26) compared to the reference group (Table 16).

Table 15. Association of chronological age, biological age, and age-difference on the risk for diabetes mellitus, hypertension, and comorbidity of diabetes mellitus and hypertension over total follow-up period

	Total Cohort N	Chronological Age (CA)		Biological Age (BA) [1]		Age-Diff (BA-CA) [2]		
		Cases N	HR (95% CI) [3]	Cases N	HR (95% CI) [3]	Age-Diff [2]	Cases N	HR (95% CI) [4]
DM								
<50	15,548	735	1.00	698	1.00	Very young BA	491	0.63 (0.55–0.72)
50–59	16,661	1035	1.61 (1.47–1.78)	1429	1.63 (1.49–1.79)	Young BA	591	0.93 (0.83–1.06)
60–69	9127	677	1.81 (1.63–2.01)	337	2.37 (2.08–2.70)	Same BA as CA	473	1.00
≥70	378	27	1.88 (1.28–2.76)	10	2.68 (1.44–5.02)	Older BA	919	1.20 (1.07–1.35)
Per 1-year increment	41,714	2474	1.03 (1.03–1.04)	2474	1.06 (1.05–1.06)		2474	*p–trend* < 0.001
HTN								
<50	9977	2822	1.00	2765	1.00	Very young BA	1405	0.74 (0.68–0.81)
50–59	8746	2840	1.38 (1.30–1.45)	3867	1.51 (1.44–1.59)	Young BA	1512	0.86 (0.80–0.93)
60–69	3863	1561	1.73 (1.63–1.84)	627	1.99 (1.82–2.17)	Same BA as CA	1473	1.00
≥70	131	51	1.57 (1.19–2.07)	15	2.48 (1.49–4.11)	Older BA	2884	1.15 (1.07–1.23)
Per 1-year increment	22,717	7274	1.03 (1.02–1.03)	7274	1.05 (1.04–1.05)		7274	*p–trend* < 0.001
Comorbidity of DM and HTN								
<50	7107	193	1.00	183	1.00	Very young BA	1047	0.65 (0.47–0.91)
50–59	5796	208	1.63 (1.32–2.02)	303	1.95 (1.59–2.38)	Young BA	135	1.10 (0.82–1.46)
60–69	2250	130	2.36 (1.84–3.03)	48	3.03 (2.15–4.27)	Same BA as CA	100	1.00
≥70	77	4	2.21 (0.82–5.99)	1	5.98 (0.83–43.01)	Older BA	196	1.32 (1.01–1.74)
Per 1-year increment	15,230	535	1.05 (1.03–1.06)	535	1.07 (1.06–1.09)		535	*p–trend* < 0.001

Abbreviations: CA, chronological age; BA, biological age; DM, diabetes mellitus; HT, hypertension; KOGES, Korean Genome and Epidemiology Study
1. BA using sex-specific Elastic net model**;** 2. BA-CA difference was classified into four groups: [Very young BA] BA was at least 5-year younger than CA; [Young BA] BA was between 1-year and < 5-year younger than CA; [Same BA as CA] BA-CA difference was between -1 year and 1 year; [Older BA] BA was at least 1 year older than CA (> 1 year); 3. Adjusted for sex; 4. Adjusted for sex and chronological age.

Table 16. Association of chronological age, biological age, and age-difference on the risk for diabetes mellitus, hypertension, and comorbidity of diabetes mellitus and hypertension on short-term follow-up period

	Total Cohort N	Chronological Age (CA)		Biological Age (BA) [1]		Age-Diff (BA-CA) [2]		
		Cases N	**HR (95% CI) [3]**	**Cases N**	**HR (95% CI) [3]**	**Age-Diff [2]**	**Cases N**	**HR (95% CI) [4]**
DM								
<50	15,548	292	1.00	295	1.00	Very young BA	251	0.66 (0.54–0.80)
50–59	16,661	514	1.67 (1.45–1.93)	692	1.72(1.50–1.97)	Young BA	310	1.02 (0.86–1.22)
60–69	9,127	373	2.18 (1.87–2.54)	202	2.81 (2.35–3.37)	Same BA as CA	219	1.00
≥70	378	18	2.43 (1.51–3.92)	8	3.76 (1.86–7.61)	Older BA	417	1.32 (1.11–1.56)
Per 1-year increment	41,714	1,197	1.04 (1.03–1.05)	1,197	1.06 (1.05–1.06)		1,197	*p–trend* < 0.001
HTN								
<50	9,977	1,586	1.00	1,516	1.00	Very young BA	966	0.74 (0.67–0.82)
50–59	8746	1,909	1.38 (1.29–1.48)	2,667	1.66 (1.56–1.77)	Young BA	980	0.84 (0.77–0.92)
60–69	3,863	1,140	1.86 (1.73–2.01)	478	2.19 (1.97–2.43)	Same BA as CA	931	1.00
≥70	131	36	1.68 (1.21–2.35)	10	2.26 (1.21–4.22)	Older BA	1,794	1.21 (1.11–1.32)
Per 1-year increment	22,717	4,671	1.03 (1.03–1.04)	4,671	1.05 (1.05–1.06)		4,671	*p–trend* < 0.001
Comorbidity of DM and HTN								
<50	7,107	87	1.00	80	1.00	Very young BA	217	0.77 (0.47–1.26)
50–59	5,796	106	1.87 (1.33–2.63)	166	2.29 (1.66–3.17)	Young BA	115	1.23 (0.80–1.91)
60–69	2,250	79	3.18 (2.17–4.64)	29	3.83 (2.37–6.20)	Same BA as CA	78	1.00
≥70	77	4	5.67 (2.04–15.77)	1	9.63 (1.32–70.28)	Older BA	115	1.47 (1.10–1.98)
Per 1-year increment	15,230	276	1.06 (1.04–1.08)	276	1.08 (1.06–1.11)		525	*p–trend* = 0.002

Abbreviations: CA, chronological age; BA, biological age; DM, diabetes mellitus; HT, hypertension; KOGES, Korean Genome and Epidemiology Study
1. BA using sex-specific Elastic net model**;** 2. BA-CA difference was classified into four groups: [Very young BA] BA was at least 5-year younger than CA; [Young BA] BA was between 1-year and < 5-year younger than CA; [Same BA as CA] BA-CA difference was between -1 year and 1 year; [Older BA] BA was at least 1 year older than CA (> 1 year); 3. Adjusted for sex; 4. Adjusted for sex and chronological age.

3.6. Prediction model study

Hypertension

A total of 30,110 individuals were included for HTN prediction model. During the median follow-up period of 4 years, 7,744 individuals (25.7%) had a newly diagnosed of HTN. The general characteristics of the study population from the KoGES used for statistical variable selection is shown in Table 17.

Prior to conduct statistical variable selection method, we tested multiple collinearities between variables based on the VIF and confirmed that there is no evidence of multiple collinearity (VIF<5) (Appendix 26-28).

The model adjusting for all of the variables (model 1) and adjusting for selected variables based on stepwise variable selection method (method 2) were presented in Table 18. According to the Model 1, family history of CVD (HR 1.13, 95% CI: 1.06-1.19), current alcohol drinking (HR 1.12, 95% CI: 1.06-1.18), and more than 240 mg/dL of total cholesterol level (HR 1.12, 95% CI: 1.04-1.21) were the remarkable predictors associated with incident HTN (Table 18).

Table 17. General characteristics of the study population for hypertension prediction model in the Korean Genome and Epidemiology Study

	Total (N=30,110)	Training set (N=21,077)	Test set (N=9,033)
Age, years, mean (SD)	51.6 (7.82)	51.6 (7.84)	51.6 (7.76)
Sex, n (%)			
Male	8,305 (27.6)	5,865 (27.8)	2,440 (27.0)
Female	21,805 (72.4)	15,212 (72.2)	6,593 (73.0)
Education, n (%)			
Elementary school	4,259 (14.10	3,029 (14.4)	1,230 (13.6)
High school	16,729 (55.6)	11,618 (55.1)	5,111 (56.6)
College and more	91,22 (30.3)	6,430 (30.5)	2,692 (29.8)
Income, n (%)			
<1,000K/month	3,468 (11.5)	2,439 (11.6)	1,029 (11.4)
1,000-2,000K	5,710 (19.0)	3,960 (18.8)	1,750 (19.4)
2,000-4,000K	12,963 (43.1)	9,119 (43.3)	3,844 (42.6)
≥4,000K	7,969 (26.5)	5,559 (26.4)	2,410 (26.7)
BMI, kg/m^2, mean (SD)	23.2 (2.68)	23.2 (2.68)	23.2 (2.67)
Waist circumstance, cm, mean (SD)	78.6 (8.22)	78.6 (8.19)	78.5 (8.27)
Total calorie intake, mean (SD)	1,764.0 (571.42)	1,762 (572.4)	1,767.6 (569.15)
Smoking, n (%)			
Never	23,376 (77.6)	16,341 (77.5)	7,035 (77.9)
Past	3,462 (11.5)	2,432 (11.5)	1,030 (11.4)
Current	3,272 (10.9)	2,304 (10.9)	968 (10.7)
Alcohol drinking, n (%)			
Never	16,626 (55.2)	11,602 (55.1)	5,024 (55.6)
Past	1,011 (3.4)	713 (3.4)	298 (3.3)
Current	12,473 (41.4)	8,762 (41.6)	3,711 (41.1)
Physical activity, n (%)			
No	14,092 (46.8)	9,885 (46.9)	4,207 (46.6)
Yes	16,018 (53.2)	11,192 (53.1)	4,826 (53.4)
Diabetes mellitus, n (%)			
No	28,408 (94.3)	19,877 (94.3)	8,531 (94.4)
Yes	1,702 (5.7)	1,200 (5.7)	502 (5.6)
Cardiovascular disease, n (%)			
No	29,374 (97.6)	20,565 (97.6)	8,809 (97.5)
Yes	736 (2.4)	512 (2.4)	224 (2.5)
Family history of CVD, n (%)			
No	24,764 (82.3)	17,363 (82.4)	7,401 (81.9)
Yes	5,346 (17.8)	3,714 (17.6)	1,632 (18.1)
HDL-cholesterol, n (%)			
Men≥40 and women≥50	20,760 (69.0)	14,527 (68.9)	6,233 (69.0)
Men<40 and women<50	9,350 (31.0)	6,550 (31.1)	2,800 (31.0)
Total-cholesterol, n (%)			
<200	17,596 (58.4)	12,258 (58.2)	5,338 (59.1)
200-240	9,506 (31.6)	6,733 (31.9)	2,773 (30.7)
≥240	3,008 (10.0)	2,086 (9.9)	922 (10.2)
Triglyceride level, n (%)			
<150	27,432 (91.1)	19,199 (91.1)	8,233 (91.1)
≥150	2,678 (8.9)	1,878 (8.9)	800 (8.9)
SBP, mmHg, mean (SD)	111.5 (9.39)	111.5 (9.43)	111.6 (9.29)
DBP, mmHg, mean (SD)	68.9 (6.04)	68.9 (6.05)	68.9 (6.00)
Albumin/creatinine ratio, mean (SD)	4.6 (0.27)	6.0 (1.17)	6.0 (1.18)

Table 18. Multivariable analysis for the association of risk factors and incident hypertension

	Model 1		Model 2	
	aHR (95% CI)	***P*-value**	**aHR (95% CI)**	***P*-value**
Sex	0.91 (0.83-0.99)	0.031	1.00 (0.95-1.06)	0.966
Age	1.02 (1.02 (1.02)	<.001	1.02 (1.02-1.02)	<.001
Education				
Elementary school	1	1.000	-	-
High school	0.99 (0.92-1.06)	0.701	-	-
College and more	0.99 (0.91-1.07)	0.727	-	-
Income				
<1,000K/month	1	1.000	-	-
1,000-2,000K	0.98 (0.91-1.06)	0.654	-	-
2,000-4,000K	1.03 (0.96-1.11)	0.406	-	-
≥4,000K	1.03 (0.94-1.12)	0.530	-	-
BMI	1.03 (1.02-1.05)	<.001	1.03 (1.02-1.04)	<.001
Waist circumstance	1.01 (1.01-1.02)	<.001	1.01 (1.01-1.02)	<.001
Smoking				
Never	1	1.000	-	-
Past	0.95 (0.88-1.04)	0.296	-	-
Current	1.03 (0.94-1.13)	0.505	-	-
Drinking				
Never	1	1.000	1	1.000
Past	1.05 (0.93-1.19)	0.449	1.06 (0.94-1.20)	0.326
Current	1.12 (1.06-1.18)	<.001	1.13 (1.08-1.19)	<.001
Physical activity	0.95 (0.91-1.00)	0.035	0.95 (0.91-0.99)	0.025
Total calorie intake	1.00 (1.00-1.00)	0.005	1.00 (1.00-1.00)	0.005
SBP, mmHg	1.05 (1.05-1.06)	<.001	1.05 (1.05-1.06)	<.001
DBP, mmHg	1.02 (1.01-1.02)	<.001	1.02 (1.01-1.02)	<.001
Diabetes mellitus	1.08 (0.99-1.18)	0.094	-	-
Total cholesterol				
<200	1	1.000	1	1.000
200-240	1.09 (1.04-1.15)	0.001	1.08 (1.03-1.14)	0.002
≥240	1.12 (1.04-1.21)	0.004	1.11 (1.03-1.19)	0.008
HDL-cholesterol	1.04 (0.99-1.10)	0.135	-	-
Triglyceride	1.00 (0.93-1.08)	0.974	-	-
Albumin/creatinine ratio	1.03 (1.01-1.06)	0.012	-	-
Cardiovascular disease	0.99 (0.86-1.14)	0.895	-	-
Family history of CVD	1.13 (1.06-1.19)	<.001	1.13 (1.06-1.19)	<.001

We constructed the predictive models for hypertension based on before and after imputation data, respectively. Before imputation, the predictive performances of for models using Cox proportional hazard (Cox PH) model were 0.7017 (Model 1) and 0.7024 (Model 2), respectively. For model 2, the c−statistics of the RSF, GBM, and elastic net were 0.7005, 0.7015, and 0.7025, respectively. After imputation, the c−statistics using Cox PH, RSF, GBM, and elastic net were 0.7013, 0.7025, 0.7040, and 0.7016, respectively (Model 2) (Table 19).

Table 19. Predictive performance of the models for hypertension based on statistical and machine learning-based models

	C-index (95% CI)	
	Model 1	**Model 2**
Before imputation		
CoxPH	0.7017 (0.7015-0.7020)	0.7024 (0.7022-0.7027)
RSF	**0.7024 (0.7021-0.7026)**	0.7005 (0.7003-0.7008)
GBM	0.7010 (0.7008-0.7013)	0.7015 (0.7013-0.7018)
ElasticNet	0.7021 (0.7018-0.7024)	**0.7025 (0.7023-0.7028)**
After imputation		
CoxPH	0.7161 (0.7159-0.7163)	0.7013 (0.7011-0.7015)
RSF	0.7152 (0.7150-0.7154)	0.7025 (0.7024-0.7028)
GBM	**0.7182 (0.7180-0.7184)**	**0.7040 (0.7038-0.7042)**
ElasticNet	0.7163 (0.7161-0.7164)	0.7016 (0.7014-0.7018)

Diabetes mellitus

A total of 60,698 participants were included for DM prediction model. During the median follow−up period of 4 years, 3,221 individuals (5.3%) had a newly diagnosed of DM. The general characteristics of the study population from the KoGES used for statistical variable selection is shown in Table 20.

Prior to conduct statistical variable selection method, we tested multiple collinearities between variables based on the VIF and confirmed that there is no evidence of multiple collinearity (VIF<5).

The model adjusting for all of the variables (model 1) and adjusting for selected variables based on stepwise variable selection method (method 2) were presented in Table 21. According to the Model 1, history of CVD (HR 1.53, 95% CI: 1.32−1.78), HTN (HR 1.45, 95% CI: 1.34−1.57) and more than 200mg/dL of triglyceride level (HR 1.42, 95% CI: 1.30−1.55) were the remarkable predictors associated with incident DM (Table 21).

Table 20. General characteristics of the study population for diabetes mellitus prediction model in the Korean Genome and Epidemiology Study

	Total (N=60,698)	Training set (N=42,489)	Test set (N=18,209)
Age, years, mean (SD)	53.3 (8.11)	53.3 (8.09)	53,3 (8.14)
Sex, n (%)			
Male	21,392 (35.2)	15,035 (35.4)	6,358 (34.9)
Female	39,305 (64.8)	27,454 (64.6)	11,851 (65.1)
Education, n (%)			
Elementary school	10,479 (17.3)	7,322 (17.2)	3,157 (17.3)
High school	33,307 (54.9)	23,348 (55.0)	9,960 (54.7)
College and more	16,911 (27.9)	11,819 (27.8)	5.092 (28.0)
Income, n (%)			
<1,000K/month	8,497 (14.0)	5,921 (13.9)	2,576 (14.2)
1,000-2,000K	12,704 (20.9)	8,887 (20.9)	3,817 (21.0)
2,000-4,000K	25,336 (41.7)	17,840 (42.0)	7,497 (41.2)
≥4,000K	14,160 (23.2)	9,841 (23.2)	4,319 (23.7)
BMI, kg/m^2, mean (SD)	23.8 (2.85)	23.8 (2.85)	23.8 (2.84)
Waist circumstance, cm, mean (SD)	80.7 (8.58)	80.7 (8.58)	80.6 (8.57)
Total calorie intake, mean (SD)	1,770.3 (573.52)	1,770.8 (576.4)	1,769.2 (566.6)
Smoking, n (%)			
Never	44,293 (73.0)	30,978 (73.1)	13,315 (73.1)
Past	9,267 (1.3)	6,522 (15.4)	2,745 (15.1)
Current	7,137 (11.8)	4,989 (11.7)	2,149 (11.8)
Alcohol drinking, n (%)			
Never	31,171 (51.4)	21,824 (51.4)	9,347 (51.3)
Past	2,261 (3.7)	1,585 (3.7)	676 (3.7)
Current	27,265 (44.9)	19,080 (44.9)	8,186 (45.0)
Physical activity, n (%)			
No	27,970 (46.1)	19,619 (46.2)	8,352 (45.9)
Yes	32,727 (53.9)	22,870 (53.8)	9,857 (54.1)
Hypertension, n (%)			
No	29,018 (47.8)	20,237 (47.6)	8,781 (48.2)
Yes	31,679 (52.2)	22,252 (52.4)	9,428 (51.8)
Cardiovascular disease, n (%)			
No	58,694 (96.7)	41,102 (96.7)	17,593 (96.6)
Yes	2,003 (3.3)	1,387 (3.3)	616 (3.4)
Family history of CVD, n (%)			
No	49,135 (81.0)	34,336 (80.8)	14,800 (81.3)
Yes	11,562 (19.0)	8,153 (19.2)	3,409 (18.7)
HDL-cholesterol, n (%)			
Men≥40 and women≥50	41,471 (68.3)	28,980 (68.2)	12,492 (68.6)
Men<40 and women<50	19,226 (31.7)	13,509 (31.8)	5,717 (31.4)
Total-cholesterol, n (%)			
<200	33,509 (55.2)	23,439 (55.2)	10,071 (55.3)
200-240	20,267 (33.4)	14,257 (33.6)	6,010 (33.0)
≥240	6,921 (11.4)	4,793 (11.3)	2,128 (11.7)
Triglyceride level, n (%)			
<150	53,021 (87.4)	37,091 (87.3)	15,931 (87.5)
≥150	7,676 (12.6)	5,398 (12.7)	2,278 (12.5)
SBP, mmHg, mean (SD)	121.9 (15.15)	121.9 (15.15)	121.7 (15.15)
DBP, mmHg, mean (SD)	76.3 (10.05)	76.3 (10.03)	76.2 (10.10)

Table 21. Multivariable analysis for the association of risk factors and incident diabetes mellitus

	Model 1		Model 2	
	aHR (95% CI)	***P*-value**	**aHR (95% CI)**	***P*-value**
Sex	0.98 (0.87-1.11)	0.780	1.12 (1.01-1.26)	0.043-
Age	1.03 (1.02-1.04)	<.001	1.03 (1.02-1.03)	<.001
Education				
Elementary school	1	1.000	-	-
High school	1.06 (0.96-1.17)	0.259	-	-
College and more	0.96 (0.85-1.09)	0.539	-	-
Income				
<1,000K/month	1	1.000	-	-
1,000-2,000K	1.07 (0.96-1.20)	0.233	-	-
2,000-4,000K	1.08 (0.96-1.20)	0.191	-	-
≥4,000K	1.11 (0.97-1.26)	0.123	-	-
BMI	1.11 (1.09-1.13)	<.001	1.11 (1.09-1.13)	<.001
Waist circumstance	1.02 (1.01-1.03)	<.001	1.02 (1.01-1.03)	<.001
Smoking				
Never	1	1.000	1	1.000
Past	1.19 (1.05-1.35)	0.007	1.19 (1.03-1.31)	<.001
Current	1.55 (1.37-1.76)	<.001	1.55 (1.40-1.70)	<.001
Drinking				
Never	1	1.000	-	-
Past	1.05 (0.89-1.24)	0.585	-	-
Current	0.96 (0.88-1.05)	0.333	-	-
Physical activity	0.99 (0.92-1.06)	0.739	-	-
Total calorie intake	1.00 (1.00-1.00)	<.001	1.00 (1.00-1.00)	<.001
hypertension	1.45 (1.34-1.57)	<.001	1.45 (1.34-1.57)	<.001
Total cholesterol				
<200	1	1.000	1	1.000
200-240	1.06 (0.98-1.14)	0.167	1.06 (0.98-1.15)	0.149
≥240	1.35 (1.22-1.50)	<.001	1.35 (1.22-1.50)	<.001
HDL-cholesterol	1.25 (1.16-1.35)	<.001	1.28 (1.19-1.39)	<.001
Triglyceride	1.42 (1.3-1.55)	<.001	1.42 (1.30-1.55)	<.001
Cardiovascular disease	1.53 (1.32-1.78)	<.001	1.55 (1.33-1.79)	<.001
Family history of CVD	1.22 (1.12-1.34)	<.001	1.23 (1.12-1.34)	<.001

The predictive performances of the models for DM using Cox proportional hazard models and machine learning-based models including RSF, GBM, and elastic net were shown in Table 22. Before imputation, the predictive performances of for models using Cox PH, RSF, GBM, and elastic net were 0.7272, 0.7022, 0.7248, and 0.7273, respectively (Model 2). After imputation, the c-statistics using Cox PH, RSF, GBM, and elastic net were 0.7225, 0.7016, 0.7220, and 0.7226, respectively (Model 2) (Table 22).

Table 22. Predictive performance of the models for diabetes mellitus based on statistical and machine learning-based models

	C-index (95% CI)	
	Model 1	**Model 2**
Before imputation		
CoxPH	**0.7282 (0.7279-0.7285)**	0.7272 (0.7269-0.7275)
RSF	0.7112 (0.7110-0.7115)	0.7022 (0.7019-0.7025)
GBM	0.7252 (0.7249-0.7255)	0.7248 (0.7245-0.7250)
ElasticNet	0.7280 (0.7277-0.7282)	**0.7273 (0.7270-0.7276)**
After imputation		
CoxPH	**0.7335 (0.7333-0.7337)**	0.7225 (0.7222-0.7227)
RSF	0.7193 (0.7191-0.7195)	0.7016 (0.7013-0.7018)
GBM	0.7295 (0.7293-0.7297)	0.7220 (0.7217-0.7222)
ElasticNet	**0.7335 (0.7333-0.7337)**	**0.7226 (0.7224-0.7229)**

Comorbidity of hypertension and Diabetes mellitus

A total of 21,459 participants were included for comorbidity of HTN and DM prediction model. During the median follow−up period of 4 years, 338 individuals (1.6%) had a newly diagnosed of comorbidity of HTN and DM. The general characteristics of the study population from the KoGES used for statistical variable selection is shown in Table 23.

Prior to conduct statistical variable selection method, we tested multiple collinearities between variables based on the VIF and confirmed that there is no evidence of multiple collinearity (VIF<5).

The model adjusting for all of the variables (model 1) and adjusting for selected variables based on stepwise variable selection method (method 2) were presented in Table 24. According to the Model 1, current smoking (HR 1.79, 95% CI: 1.20−2.65), history of CVD (HR 1.80, 95% CI: 1.10−2.95) and more than 150mg/dL of triglyceride level (HR 1.62, 95% CI: 1.21−2.16) were the remarkable predictors associated with comorbidity of HTN and DM (Table 24).

Table 23. General characteristics of the study population for comorbidity of hypertension and diabetes mellitus prediction model in the Korean Genome and Epidemiology Study

	Total (N=21,459)	Training set (N=15,022)	Test set (N=6,437)
Age, years, mean (SD)	50.9 (7.6)	50.9 (7.6)	50.9 (7.7)
Sex, n (%)			
Male	5,116 (23.8)	3,604 (24.0)	1,512 (23.5)
Female	16,343 (76.2)	11,418 (76.0)	4,925 (76.5)
Education, n (%)			
Elementary school	2,663 (12.4)	1,856 (12.4)	807 (12.5)
High school	11,931 (55.6)	8,352 (55.6)	3,579 (55.6)
College and more	6,865 (32.0)	4,814 (32.0)	2,051 (31.9)
Income, n (%)			
<1,000K/month	2,147 (10.0)	1,488 (9.9)	659 (10.2)
1,000-2,000K	3,997 (18.6)	2,833 (18.9)	1,164 (18.1)
2,000-4,000K	9,381 (43.7)	6,512 (43.3)	2,869 (44.6)
≥4,000K	5,934 (27.7)	4,189 (27.9)	1,745 (27.1)
BMI, kg/m^2, mean (SD)	22.9 (2.6)	22.9 (2.6)	22.8 (2.6)
Waist circumstance, cm, mean (SD)	77.4 (8.0)	77.4 (8.0)	77.3 (7.9)
Total calorie intake, mean (SD)	1,762.7 (574.8)	1,761.8 (568.0)	1,764.6 (590.2)
Smoking, n (%)			
Never	17,254 (80.4)	12,039 (80.1)	5,215 (81.0)
Past	2,134 (9.9)	1,499 (10.0)	635 (9.9)
Current	2,071 (9.7)	1,484 (9.9)	576 (9.1)
Alcohol drinking, n (%)			
Never	12,205 (56.9)	8,552 (56.9)	3,653 (56.7)
Past	657 (3.1)	452 (3.0)	205 (3.2)
Current	8,597 (40.0)	6,018 (40.1)	2,579 (40.1)
Physical activity, n (%)			
No	10,091 (47.0)	6,990 (46.5)	3,101 (48.2)
Yes	11,368 (53.0)	8,032 (53.5)	3,336 (51.8)
Cardiovascular disease, n (%)			
0	21,020 (98.0)	14,713 (97.9)	6,307 (98.0)
1	439 (2.0)	309 (2.1)	130 (2.0)
Family history of CVD, n (%)			
0	17,670 (82.3)	12,367 (82.3)	5,303 (82.4)
1	3,789 (17.7)	2,655 (17.7)	1,134 (17.6)
HDL-cholesterol, n (%)			
Men≥40 and women≥50	15,087 (70.3)	10,575 (70.4)	4,512 (70.1)
Men<40 and women<50	6,372 (29.7)	4,447 (29.6)	1,925 (29.9)
Total-cholesterol, n (%)			
<200	12,737 (59.4)	8,919 (59.4)	2,818 (59.3)
200-240	6,666 (31.1)	4,643 (30.9)	2,023 (31.4)
≥240	2,056 (9.6)	1,460 (9.7)	596 (9.3)
Triglyceride level, n (%)			
<150	19,885 (92.7)	13,923 (92.7)	5,962 (92.6)
≥150	1,574 (7.3)	1,099 (7.3)	475 (7.4)
SBP, mmHg, mean (SD)	110.0 (9.3)	110.0 (9.3)	110.0 (9.2)
DBP, mmHg, mean (SD)	68.1 (6.1)	68.1 (6.1)	68.1 (6.1)
Albumin/creatinine ratio, mean (SD)	6.0 (1.2)	6.0 (1.2)	6.0 (1.2)

Table 24. Multivariable analysis for the association of risk factors and the risk of comorbidity of hypertension and diabetes mellitus

	Model 1		Model 2	
	aHR (95% CI)	***P*-value**	**aHR (95% CI)**	***P*-value**
Sex	1.16 (0.74-1.80)	0.672	1.00 (0.67-1.50)	0.999
Age	1.05 (1.04-1.07)	<.001	1.05 (1.03-1.06)	<.001
Education				
Elementary school	1	1.000	-	-
High school	1.32 (0.96-1.82)	0.059	-	-
College and more	1.05 (0.70-1.57)	0.610	-	-
Income				
<1,000K/month	1	1.000	-	-
1,000-2,000K	0.73 (0.50-1.05)	0.18	-	-
2,000-4,000K	0.88 (0.62-1.23)	0.993	-	-
≥4,000K	0.81 (0.54-1.21)	0.696	-	-
BMI	1.13 (1.06-1.20)	<.001	1.13 (1.07-1.21)	<.001
Waist circumstance	1.05 (1.03-1.08)	<.001	1.05 (1.02-1.07)	<.001
Smoking				
Never	1	1.000	1	1
Ever	1.00 (0.65-1.53)	0.901	1.04 (0.68-1.58)	0.875
Current	1.79 (1.20-2.65)	0.010	1.78 (1.20-2.64)	0.004
Drinking				
Never	1	1.000	-	-
Ever	1.29 (0.77-2.15)	0.280	-	-
Current	1.09 (0.84-1.41)	0.622	-	-
Physical activity	1.09 (0.88-1.37)	0.423	-	-
SBP, mmHg	1.07 (1.05-1.09)	<.001	1.07 (1.05-1.09)	<.001
DBP, mmHg	1.03 (1.01-1.05)	0.031	1.03 (1.01-1.05)	<.001
Total calorie intake	1.00 (1.00-1.00)	0.005	-	-
Total cholesterol				
0	1	1.000	-	-
1	1.13 (0.89-1.45)	0.319	-	-
2	1.61 (1.15-2.23)	0.005	-	-
HDL-cholesterol	1.30 (1.02-1.66)	0.004	-	-
Triglyceride	1.62 (1.21-2.16)	0.001	1.84 (1.40-2.42)	<.001
Albumin/creatinine ratio	1.17 (1.03-1.32)	0.013	1.15 (1.02-1.29)	0.027
Cardiovascular disease	1.80 (1.10-2.95)	0.020	1.82 (1.12-2.98)	0.016
Family history of CVD	1.39 (1.06-1.83)	0.019	1.40 (1.07-1.84)	0.015

The predictive performances of the models for comorbidity of HTN and DM using Cox proportional hazard models and machine learning-based models including RSF, GBM, and elastic net were shown in Table 25. Before imputation, the predictive performances of for models using Cox PH, RSF, GBM, and elastic net were 0.7809, 0.7780, 0.7759, and 0.7826, respectively (Model 2). After imputation, the c-statistics using Cox PH, RSF, GBM, and elastic net were 0.8170, 0.7907, 0.8089, and 0.8165, respectively (Model 2) (Table 25).

Table 25. Predictive performance of the models for comorbidity of hypertension and diabetes mellitus based on statistical and machine learning-based models

	C-index (95% CI)	
	Model 1	**Model 2**
Before imputation		
CoxPH	0.7810 (0.7799-0.7822)	0.7809 (0.7798-0.7821)
RSF	**0.7860 (0.7848-0.7872)**	0.7780 (0.7768-0.7692)
GBM	0.7802 (0.7789-0.7814)	0.7759 (0.7746-0.7771)
ElasticNet	0.7830 (0.7818-0.7841)	**0.7826 (0.7814-0.7838)**
After imputation		
CoxPH	0.8180 (0.8171-0.8188)	**0.8170 (0.8161-0.8179)**
RSF	0.8135 (0.8126-0.8143)	0.7907 (0.7897-0.7917)
GBM	0.8096 (0.8087-0.8104)	0.8089 (0.8080-0.8097)
ElasticNet	**0.8192 (0.8183-0.8201)**	0.8165 (0.8156-0.8174)

IV. Discussion

4.1. Key findings

In this study, we highlighted the importance of metabolic comorbidity and suggested the machine learning-based disease prediction models for metabolic comorbidity prevention and management in Korean population.

First, we found that Korea had a lower prevalence of metabolic comorbidity compared to the US. In Korea, individuals living in urban areas had the lower prevalence of comorbidity than those living in rural areas. Second study evaluated the combined effects of metabolic comorbidity with a first-degree family history of CVD on the risk of CVD in KoGES database. We found that individuals with DM, HTN, LIP, and with a family history of CVD had a 2.88-fold increased risk of CVD, a 3.30-fold increased risk of MI, and a 2.52-fold increased risk of stroke compared to the people with a negative family history of CVD and none of metabolic diseases. Third study investigated the impact of lifestyle factors (cigarette smoking, alcohol consumption, and obesity) with CMDs on CVD death in Asian multi-center cohort studies. The results showed that as the HLS increased by 1 score, the risk was decreased by 13% in those with HTN and DM, 27% in those with HTN and CHD, and 14% in those with HTN and stroke, and 24% in those with

HTN, DM, and CHD at baseline. Among the three lifestyle factors, non-smoking had the strongest association with decreasing risk of CVD-specific death regardless of the number of CMDs. Moreover, among individuals with cardiometabolic comorbidity, having three of healthy lifestyle factors was significantly associated with decrease in overall and premature CVD-specific death. Forth, based on the repeated measurements for assessing changes in lifestyle factors in Ansan and Ansung Study, we found that unhealthy lifestyle change including increased intensity of cigarette smoking, alcohol consumption, and BMI was associated with a significantly elevated risk of HTN, DM, and MetS. Finally, for improving the individualized health status, we developed a self-assessed BA as a predictor for metabolic comorbidity. Individuals with a lower BA compared to the CA have a decreased risk of HTN and DM comorbidity and the risk decreased rapidly within 5 years of follow-up. For disease prediction study, predictive models based on machine learning approaches achieved a high discriminatory ability for co-occurrence of HTN and DM. The predictive ability of machine learning approaches is promising, especially elastic net algorithm. We also found that prediction models using multiple imputations showed a better prediction accuracy (Figure 23).

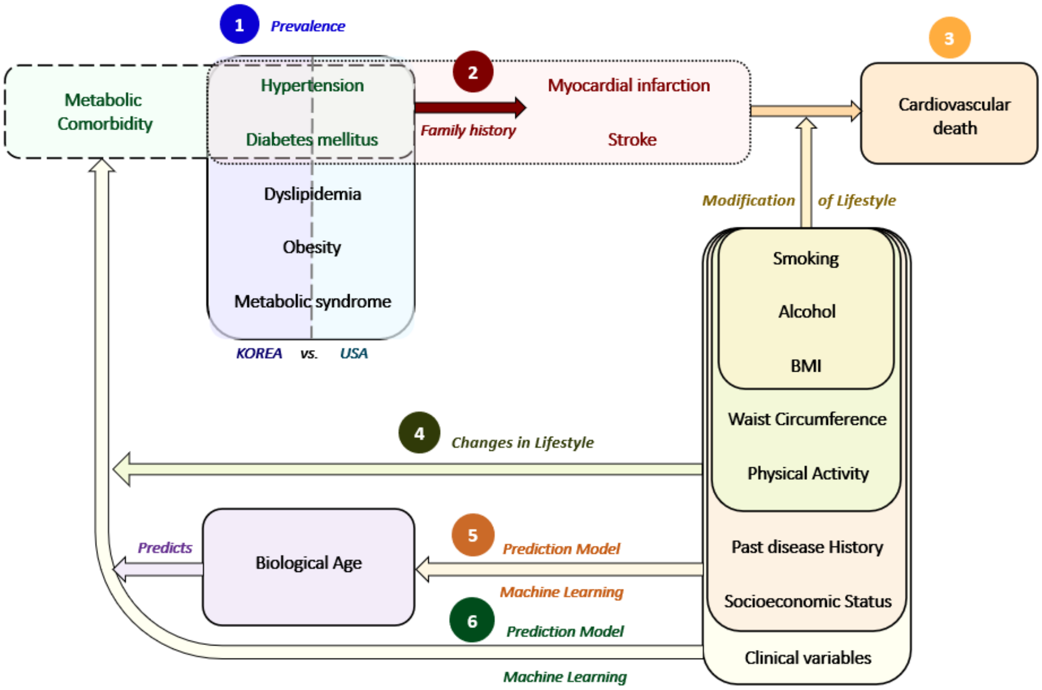


Figure 23. Summary of the results

4.2. Comparison to previous studies

4.2.1. Prevalence study

In this study, we estimated and compared the prevalence of metabolic disease and comorbidity between Korea and the US using four data sources. We found that Korea showed the lower prevalence of metabolic disease and comorbidity compared to the US. This disparity might be associated with the variances in lifestyle behaviors and dietary patterns between two countries [94]. Due to economic growth and westernization, however, the increasing prevalence of metabolic comorbidity is a major problem in Korea [13–16]. Several studies have estimated that the prevalence rates of metabolic diseases in the Korea and the US, respectively;[18, 19] but few studies compared the differences. Comparing the prevalence of metabolic disease and comorbidity in Korea with the US may be an important role for prevention strategies to reduce the future CVD risk in Korea.

We also found that Korean rural residents had a higher prevalence of comorbidity than those living in urban areas. This disparity might be based on the difference on SES, unhealthy lifestyle behaviors, and limited access to health care system [95–97]. This emphasizes the necessary of public health program in rural areas to prevent the further risk of CVD mortality.

4.2.2. Family history of CVD and the risk of CVD study

The prevalence of metabolic diseases was associated with an increased risk of MI, as demonstrated in this study. With the aging population, the prevalence of metabolic comorbidities is constantly increasing, and a continued increase in CVD is inevitable [22, 27, 28]. The risk of MI in diabetic patients with high blood pressure has been reported to be more than 2-fold higher than that in patients without these conditions [23]. In another study, patients with a prevalence of DM and LIP had a 1.3-fold increase in CVD risk [24]. Moreover, the prevalence of cardiometabolic comorbidities increases the risk of overall and CVD-related mortality [47, 98, 99].

A family history of CVD is another major risk factor for MI [100]. Previous studies have reported that family history represents a genetic predisposition that contributes to an increased risk of MI [101]. Moreover, parental CVD is associated with a greater prevalence of metabolic disease [26]. However, no prior study has found a relationship between metabolic comorbidities and MI events in patients with a family history of CVD. This study identified individuals with cardiometabolic comorbidities who were at a high risk of MI based on their genetic background. These results demonstrate, for the first time, that metabolic comorbidities contribute to hereditary aggregation of MI and stroke. We also found that adherence to a healthy lifestyle was important even among individuals with a positive family history of CVD.

4.2.3. Lifestyle and the risk of CVD death study

The cardiometabolic comorbidity, lifestyle factors, and the risk of CVD death study is in line with a Japanese cohort study that showed an inverse association between healthy lifestyle factors and the risk of CVD death [34]. Other cohort studies have also shown that multiple healthy lifestyle factors significantly reduced the risk of all-cause and CVD death [38-41]. However, previous studies are limited in that they included individuals without cardiometabolic diseases at baseline. Only a few studies have shown that a healthier lifestyle can consistently prolong life expectancy irrespective of the multiple chronic diseases, but these studies primarily focused on the Western population [42, 43]. To our knowledge, this study is the first study to examine the association of HLS with CVD-specific death according to the combination of CMDs in Asian population.

In our HLS, never smokers are given a point as cigarette smoking is a strong independent risk factor for all-cause and CVD death [38, 102, 103]. The relation between alcohol consumption and mortality, however, is controversial. Previous studies suggested J-shape associations of alcohol consumption with death [104, 105] while other studies showed linear association [106-109]. Based on these findings and the potential adverse effects of moderate alcohol consumption, only never-drinkers are given a point when calculating the HLS. Lastly, individuals with BMI ranging from 18.5 to 27.4 are given a point in the score as BMI demonstrated a U-shaped relationship with all-cause and CVD death in the Asian population [37, 74, 110].

The mechanism whereby the three elements of the HLS reduce all-cause and CVD death may be by facilitating the control of CMDs. Firstly, smoking cessation can reduce blood pressure and arterial stiffness, which then lowers the risk of coronary heart disease and stroke [111-113]. Secondly, heavy alcohol consumption is associated with increasing blood pressure as well as blood glucose levels [114]. Thirdly, people with healthy BMI have fewer comorbidities as abdominal obesity is associated with a 1.48- and 1.65-times higher probability of high blood pressure and blood sugar level, respectively [115]. As these conditions are also associated with higher risk of CVDs, so by better controlling these conditions with higher HLS, the risk of all-cause and CVD death can be multiplicatively reduced [116-121].

4.2.4. Change in lifestyles study

In change in lifestyle factors and MetS study, we found that individuals who were continuously smoking, increased their intensity of alcohol consumption from light/moderate to heavy, became physically inactive from the physically active, or newly became obesity in the second examination from the normal BMI in the first examination had a significantly increased risk of MetS. According to the meta-analyses, heavy smokers [122], heavy alcohol consumption [123], low levels of physical activity [124], and overweight/obesity [125] are associated with increasing risk of MetS. These previous studies examined the relationship between lifestyle at baseline and risk of MetS [122-125]. A recent study found the association between change in drinking alcohol and MetS in Korean population; however, it could not identify whether changes in drinking alcohol occurred before/after the MetS [45]. As our study is based on the multiple repeated measurements of lifestyle factors, our results suggest the evidence for the causal relationship between lifestyle behaviors and MetS. Our results also suggest a dose-response association between the change in lifestyle factors and MetS by specifying the dose of cigarette smoking per day, amount of alcohol consumption per day, and overweight/obesity definition.

4.2.5. Biological age study

In self-assessed biological age and metabolic comorbidity study, we developed a BA using self-assessed measurement. There were several studies to suggest BA as an indicator for health status, however, they used clinical biomarkers, including laboratory blood tests [53, 126], physical tests (grip strength and vertical jump) [52], physiological factors (body mass index and percent body fat mass) [53, 126], metabolomics [51], and DNA methylation [57] to calculate a BA. Thus, it is difficult to generalize for public health due to the restrictions on information collection. Among the predictors for BA in this study, we found a positive association between waist size and BA. This association was confirmed by previous studies that abdominal obesity is associated with metabolic diseases [127-129]. We also found that smoking and drinking were significantly related to the BA. This association was in line with the J-shaped relationship between alcohol consumption and all-cause mortality in Korea [130]. The association between smoking duration and aging was also supported that smoking increased oxidative stress which accelerated aging [131–133]. These findings support the evidence that lifestyle is related to the biological aging.

Prior study generally used principal component analysis or multiple linear regression to calculate the BA, however, these methods had overfitting problem and low interpretability [134]. Thus, in this study, the elastic net regression with 10-fold cross-validation methods were used to calculate the BA that reduce overfitting and minimize bias [88, 89]. Furthermore, previous studies were limited to find association between the BA and disease prevalence [52] or estimate the risk of mortality [51, 53]. In this study, we developed and validated it as a useful index of the risk of developing metabolic disease.

4.2.6. Prediction model study

In the machine learning-based prediction model of metabolic comorbidity study, we developed prediction models using statistical and machine learning approaches (RSF, GBM, and elastic net) to predict the comorbidity of HTN and DM based on the common risk factors and evaluated its accuracy. In recent years, previous studies have developed risk prediction model for DM, and HTN [135, 136], however, the existing models focus on predicting only a single disease at a time. Since increasing number of people may suffer from metabolic comorbidity, these models are inadequate for predicting the comorbidity of HTN and DM simultaneously. Although there are many studies using machine learning approaches for developing a predictive model [137–141], a few studies have developed a disease prediction model based on the machine learning algorithms for analyzing time–series data [69, 142].

Both HTN and DM share common risk factors. Age, BMI, SES status, lifestyle factors, high lipid profiles, history of CVD, and family history of CVD are significant predictors of HTN and DM and used in previous prediction models [135, 136, 141]. further, blood pressure and albumin/creatinine ratio are common risk factors for HTN [136, 143]. Our prediction model showed that current cigarette smoking, high blood pressure, high lipid profiles, albumin/creatinine ratio, history of CVD, and family history of CVD were at risk of developing both HTN and DM. The high accuracy of our machine learning–based prediction models for HTN and DM comorbidity made it potential for early–detection and health management.

The machine learning algorithms including RSF, GBM, and elastic net showed high predictive ability in prediction of comorbidity of HTN and DM. However, there is heterogeneity among those methods. Elastic net is a regularization algorithm designed for shrinking the regression parameter estimates towards zero to select variables and obtain optimal estimates [88, 93]. While both random forest and boosting model can detect and predict non-linear associations and interactions among the variables [144]. To this end, we suggested to select the machine learning algorithms base on the study outcomes and the data characteristics [145, 146].

4.3. Strengths and limitations

There are several strengths in this study. First, we estimated and compared the nationally representative age−standardized prevalence of metabolic disease and comorbidity based on the NHANES and KNHANES data, which were the representative population data from the US and Korea, respectively. The Korean studies are also subdivided into an urban and rural cohort to compare the difference of prevalence between rural and urban in Korea. Second, the metabolic comorbidity, family history of CVD, and the risk of CVD study is based on the large sample size in prospective study design with a long follow−up period. To our knowledge, this is the first study to estimate the impact of metabolic comorbidity on CVD among individuals with a family history of CVD in Korean population. This study highlights the necessity of accounting to metabolic comorbidity among individuals with a family history of CVD to reduce the risk of CVD. Moreover, this study provides evidence of interaction between family history of CVD and lifestyle factors in the development of CVD. Future genetic and lifestyle risk factors interactions studies are important as supporting our findings and providing individualized lifestyle prevention strategies. Third, the impact of cardiometabolic comorbidity and healthy lifestyle factors on CVD death study is the largest multicenter cohort study that examine the association between the combination of healthy lifestyle factors with CVD−specific death according to the number of CMDs. Is also examine the impact of HLS on CVD death stratified to the combination of CMDs at baseline. These strengths allowed us to comprehensively examine the impact of healthy lifestyle factors on all−cause and CVD death with sufficient statistical power. Fourth, change in lifestyle factors and MetS study found the effects of change in lifestyle factors on MetS using population−based cohort study with multiple repeated measurements. Fifth, the self−assessed biological age and metabolic comorbidity study is the first study to develop the BA prediction model using the self−assessed measurements that are well−measured, well−understood, and easily collected based on machine learning approaches. As BA was calculated based on the modifiable factors, it could be useful to suggest healthy lifestyle guidelines for prevention. At last, we developed machine learning−based predictive models for predicting HTN and DM co−occurrence with high predictive ability. We also compared the prognostic performance among models fitted to the imputed data and missing data and found that applying imputation of missing values can improve the predictive accuracy of models. This might help in early identification of individuals with HTN and DM comorbidity to reduce further healthcare burden in Korea.

However, the findings of our studies should be interpreted in the context of several limitations. First, differences between each of the studies from study design, measurement techniques, and misclassification bias could be introduced in prevalence study. For nationally representative dataset, relevant published weights for survey sample were used to analysis. Second, as this study used self-reported history of disease, family history of CVD, and lifestyle factors, there was response bias, which could have underestimated or overestimated the values [100, 147]. Previous validation studies, however, reported that the accuracy of both self-reported CVD and family history was over 80% [148, 149]. Moreover, due to the follow-up loss in this study population, the reduced effective sample size and differential rates of risk factors between the comparison groups with different follow-up rates might cause selection bias [150, 151]. Third, an explanation for past drinkers having greater risk of CVD death despite their alcohol cessation is that their comorbidities may have motivated them to stop drinking [152]. Moreover, our study was only able to include 3 lifestyle factors (smoking, drinking, and BMI level) due to the lack of physical and dietary data. WHO enacted a global guideline with 5 lifestyle factors, namely smoking cessation, alcohol abstinence, healthy BMI, regular exercise, and a healthy diet, to prevent the risk of CVDs. Further research including all the 5 lifestyle factors recommended by WHO should be conducted to produce more comprehensive results on the benefits of a healthy lifestyle. Fourth, due to the limited number of study populations, we could not examine the association between change in lifestyle factors and metabolic comorbidity. Otherwise, we investigated the dose-response relationship between change in lifestyle factor and MetS, which is a cluster of metabolic conditions. Future studies investigating whether change in lifestyle factors is associated with the risk of metabolic comorbidity are needed. Also, there is a need to create a risk prediction model based on change in lifestyle factors. Fifth, future research investigating lifestyle-based BA in different populations with diverse lifestyle behaviors, will need to be undertaken to generalize the BA. Further study should attempt to investigate the association between the BA and the risk of CVD and CVD-related mortality. At last, the c-statistics may not be optimal in assessing prediction models due to difficulty in representing the small changes in coefficients and limited clinical relevance [153]. Thus, other measure of model performance such as net reclassification index, which represents the improvement in model reclassification, is recommended in the future study [154].

V. Conclusions

This study highlights the necessity of accounting to metabolic comorbidity to reduce the risk of CVD outcomes in Korean population. Although individuals already have had cardiometabolic comorbidity, healthy lifestyles (smoking cessation, abstaining from alcohol, and maintaining BMI) are effective to reduce the further risk of CVD death. Moreover, lifestyle changes help to decrease the risk of a cluster of metabolic conditions. At last, machine learning-based self-assessed of BA and disease prediction model may be an effective tool for identifying the high-risk group and decreasing burden of metabolic comorbidities in Korea through health promotion.

References

1. Tran J, Norton R, Conrad N, Rahimian F, Canoy D, Nazarzadeh M, et al. Patterns and temporal trends of comorbidity among adult patients with incident cardiovascular disease in the UK between 2000 and 2014: A population-based cohort study. PLoS Med. 2018;15(3):e1002513. Epub 2018/03/07. doi: 10.1371/journal.pmed.1002513. PubMed PMID: 29509757; PubMed Central PMCID: PMCPMC5839540 on the journal's editorial board. KR also served as a guest editor on PLOS Medicine’s Cardiovascular Disease Special Issue.
2. Koton S, Schneider AL, Rosamond WD, Shahar E, Sang Y, Gottesman RF, et al. Stroke incidence and mortality trends in US communities, 1987 to 2011. Jama. 2014;312(3):259-68. Epub 2014/07/17. doi: 10.1001/jama.2014.7692. PubMed PMID: 25027141.
3. Schnabel RB, Yin XY, Gona P, Larson MG, Beiser AS, McManus DD, et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. Lancet (London, England). 2015;386(9989):154-62. doi: 10.1016/S0140-6736(14)61774-8. PubMed PMID: WOS:000357742200032.
4. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet (London, England). 2017;390(10100):1211-59. Epub 2017/09/19. doi: 10.1016/s0140-6736(17)32154-2. PubMed PMID: 28919117; PubMed Central PMCID: PMCPMC5605509.
5. Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017;390(10100):1151-210. doi: 10.1016/s0140-6736(17)32152-9.
6. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. Circulation. 2019;139(10):e56-e528. Epub 2019/02/01. doi: 10.1161/cir.0000000000000659. PubMed PMID: 30700139.
7. Shin H-Y, Kim J, Lee S, Park MS, Park S, Huh S. Cause-of-death statistics in 2018 in the Republic of Korea. Journal of the Korean Medical Association. 2020;63(5):286-97. doi: 10.5124/jkma.2020.63.5.286.
8. Helfand M, Buckley D, Fleming C, Fu R, Freeman M, Humphrey L, et al. U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews. Screening for Intermediate Risk Factors for Coronary Heart Disease. Rockville (MD): Agency for Healthcare Research and Quality (US); 2009.

9. Kokubo Y, Iwashima Y. Higher blood pressure as a risk factor for diseases other than stroke and ischemic heart disease. Hypertension (Dallas, Tex : 1979). 2015;66(2):254-9. Epub 2015/06/17. doi: 10.1161/hypertensionaha.115.03480. PubMed PMID: 26077565.
10. Laakso M. Cardiovascular disease in type 2 diabetes from population to man to mechanisms: the Kelly West Award Lecture 2008. Diabetes Care. 2010;33(2):442-9. Epub 2010/01/28. doi: 10.2337/dc09-0749. PubMed PMID: 20103560; PubMed Central PMCID: PMCPMC2809299.
11. Wang X, Dong Y, Qi X, Huang C, Hou L. Cholesterol levels and risk of hemorrhagic stroke: a systematic review and meta-analysis. Stroke. 2013;44(7):1833-9. Epub 2013/05/25. doi: 10.1161/strokeaha.113.001326. PubMed PMID: 23704101.
12. Yaghi S, Elkind MS. Lipids and Cerebrovascular Disease: Research and Practice. Stroke. 2015;46(11):3322-8. Epub 2015/10/10. doi: 10.1161/STROKEAHA.115.011164. PubMed PMID: 26451029; PubMed Central PMCID: PMCPMC4624572.
13. Moran AE, Forouzanfar MH, Roth GA, Mensah GA, Ezzati M, Murray CJ, et al. Temporal trends in ischemic heart disease mortality in 21 world regions, 1980 to 2010: the Global Burden of Disease 2010 study. Circulation. 2014;129(14):1483-92. Epub 2014/02/28. doi: 10.1161/circulationaha.113.004042. PubMed PMID: 24573352; PubMed Central PMCID: PMCPMC4181359.
14. Kim HJ, Kim Y, Cho Y, Jun B, Oh KW. Trends in the prevalence of major cardiovascular disease risk factors among Korean adults: results from the Korea National Health and Nutrition Examination Survey, 1998-2012. Int J Cardiol. 2014;174(1):64-72. Epub 2014/04/20. doi: 10.1016/j.ijcard.2014.03.163. PubMed PMID: 24742812.
15. Suh S, Lee MK. Metabolic syndrome and cardiovascular diseases in Korea. J Atheroscler Thromb. 2014;21 Suppl 1:S31-5. Epub 2014/01/24. PubMed PMID: 24452115.
16. Lee SW, Kim HC, Lee HS, Suh I. Thirty-year trends in mortality from cardiovascular diseases in Korea. Korean Circ J. 2015;45(3):202-9. Epub 2015/05/30. doi: 10.4070/kcj.2015.45.3.202. PubMed PMID: 26023308; PubMed Central PMCID: PMCPMC4446814.
17. Yoon YS, Oh SW. Recent Shift of Body Mass Index Distribution in Korea: a Population-based Korea National Health Insurance Database, 2002-2013. J Korean Med Sci. 2017;32(3):434-8. Epub 2017/02/02. doi: 10.3346/jkms.2017.32.3.434. PubMed PMID: 28145646; PubMed Central PMCID: PMCPMC5290102.
18. Kim H, Kim S, Han S, Rane PP, Fox KM, Qian Y, et al. Prevalence and incidence of atherosclerotic cardiovascular disease and its risk factors in Korea: a nationwide population-based study. BMC Public Health. 2019;19(1):1112. Epub 2019/08/16. doi: 10.1186/s12889-019-7439-0. PubMed PMID: 31412823; PubMed Central PMCID: PMCPMC6694551.

19. Muntner P, Hardy ST, Fine LJ, Jaeger BC, Wozniak G, Levitan EB, et al. Trends in Blood Pressure Control Among US Adults With Hypertension, 1999-2000 to 2017-2018. Jama. 2020;324(12):1190-200. Epub 2020/09/10. doi: 10.1001/jama.2020.14545. PubMed PMID: 32902588; PubMed Central PMCID: PMCPMC7489367 consulting fees from Amgen Inc. Dr Levitan reported receiving grant funding from and serving on advisory boards for Amgen Inc; and serving as a consultant to Novartis. Dr Colantonio reported receiving grant funding from Amgen Inc. No other disclosures were reported.
20. Cui J, Liu Y, Li Y, Xu F, Liu Y. Type 2 Diabetes and Myocardial Infarction: Recent Clinical Evidence and Perspective. Front Cardiovasc Med. 2021;8:644189. Epub 2021/03/16. doi: 10.3389/fcvm.2021.644189. PubMed PMID: 33718461; PubMed Central PMCID: PMCPMC7943438.
21. Reindl M, Reinstadler SJ, Feistritzer HJ, Theurl M, Basic D, Eigler C, et al. Relation of Low-Density Lipoprotein Cholesterol With Microvascular Injury and Clinical Outcome in Revascularized ST-Elevation Myocardial Infarction. J Am Heart Assoc. 2017;6(10). Epub 2017/10/12. doi: 10.1161/jaha.117.006957. PubMed PMID: 29018020; PubMed Central PMCID: PMCPMC5721881.
22. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. J Am Coll Cardiol. 2020;76(25):2982-3021. Epub 2020/12/15. doi: 10.1016/j.jacc.2020.11.010. PubMed PMID: 33309175; PubMed Central PMCID: PMCPMC7755038.
23. Böhm M, Schumacher H, Teo KK, Lonn EM, Mahfoud F, Mann JFE, et al. Cardiovascular outcomes and achieved blood pressure in patients with and without diabetes at high cardiovascular risk. Eur Heart J. 2019;40(25):2032-43. Epub 2019/03/29. doi: 10.1093/eurheartj/ehz149. PubMed PMID: 30919899.
24. Kaze AD, Santhanam P, Musani SK, Ahima R, Echouffo-Tcheugui JB. Metabolic Dyslipidemia and Cardiovascular Outcomes in Type 2 Diabetes Mellitus: Findings From the Look AHEAD Study. J Am Heart Assoc. 2021;10(7):e016947. Epub 2021/03/18. doi: 10.1161/jaha.120.016947. PubMed PMID: 33728932; PubMed Central PMCID: PMCPMC8174364.
25. Andresdottir MB, Sigurdsson G, Sigvaldason H, Gudnason V. Fifteen percent of myocardial infarctions and coronary revascularizations explained by family history unrelated to conventional risk factors. The Reykjavik Cohort Study. Eur Heart J. 2002;23(21):1655-63. Epub 2002/10/26. doi: 10.1053/euhj.2002.3235. PubMed PMID: 12398822.
26. Dallongeville J, Grupposo MC, Cottel D, Ferrières J, Arveiler D, Bingham A, et al. Association between the metabolic syndrome and parental history of premature cardiovascular disease. Eur Heart J. 2006;27(6):722-8. Epub 2006/01/13. doi: 10.1093/eurheartj/ehi717. PubMed PMID: 16401673.
27. Park JJ, Lee CJ, Park S-J, Choi J-O, Choi S, Park S-M, et al. Heart Failure Statistics in Korea, 2020: A Report from the Korean Society of Heart Failure. Int J Heart Fail. 2021;3(4):224-36.
28. Hirode G, Wong RJ. Trends in the Prevalence of Metabolic Syndrome in

the United States, 2011-2016. Jama. 2020;323(24):2526-8. Epub 2020/06/24. doi: 10.1001/jama.2020.4501. PubMed PMID: 32573660; PubMed Central PMCID: PMCPMC7312413 Gilead Sciences and Abbvie and being part of the advisory board and speaker' s bureau for Gilead Sciences. No other disclosures were reported.
29. Kong MG, Jang SY, Jang J, Cho HJ, Lee S, Lee SE, et al. Impact of diabetes mellitus on mortality in patients with acute heart failure: a prospective cohort study. Cardiovasc Diabetol. 2020;19(1):49. Epub 2020/05/04. doi: 10.1186/s12933-020-01026-3. PubMed PMID: 32359358; PubMed Central PMCID: PMCPMC7196232.
30. Lee JY, Park JT, Joo YS, Lee C, Yun HR, Chang TI, et al. Association of Blood Pressure with Cardiovascular Outcome and Mortality: Results from the KNOW-CKD Study. Nephrol Dial Transplant. 2021. Epub 2021/09/03. doi: 10.1093/ndt/gfab257. PubMed PMID: 34473286.
31. World Health Organization. Global action plan for the prevention and control of noncommunicable diseases 2013-2020. https://apps.who.int/iris/handle/10665/94384. Published 31 May, 2013.
32. Rippe JM. Lifestyle Strategies for Risk Factor Reduction, Prevention, and Treatment of Cardiovascular Disease. Am J Lifestyle Med. 2019;13(2):204-12. Epub 2019/02/26. doi: 10.1177/1559827618812395. PubMed PMID: 30800027; PubMed Central PMCID: PMCPMC6378495.
33. Tsai MC, Lee CC, Liu SC, Tseng PJ, Chien KL. Combined healthy lifestyle factors are more beneficial in reducing cardiovascular disease in younger adults: a meta-analysis of prospective cohort studies. Sci Rep. 2020;10(1):18165. Epub 2020/10/25. doi: 10.1038/s41598-020-75314-z. PubMed PMID: 33097813; PubMed Central PMCID: PMCPMC7584648.
34. Eguchi E, Iso H, Tanabe N, Wada Y, Yatsuya H, Kikuchi S, et al. Healthy lifestyle behaviours and cardiovascular mortality among Japanese men and women: the Japan collaborative cohort study. Eur Heart J. 2012;33(4):467-77. Epub 2012/02/16. doi: 10.1093/eurheartj/ehr429. PubMed PMID: 22334626.
35. Meschia JF, Bushnell C, Boden-Albala B, Braun LT, Bravata DM, Chaturvedi S, et al. Guidelines for the primary prevention of stroke: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2014;45(12):3754-832. Epub 20141028. doi: 10.1161/str.0000000000000046. PubMed PMID: 25355838; PubMed Central PMCID: PMCPMC5020564.
36. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004;363(9403):157-63. doi: 10.1016/s0140-6736(03)15268-3. PubMed PMID: 14726171.
37. Zheng W, McLerran DF, Rolland B, Zhang X, Inoue M, Matsuo K, et al. Association between body-mass index and risk of death in more than 1 million Asians. N Engl J Med. 2011;364(8):719-29. Epub 2011/02/25. doi: 10.1056/NEJMoa1010679. PubMed PMID: 21345101; PubMed Central PMCID: PMCPMC4008249.

38. van Dam RM, Li T, Spiegelman D, Franco OH, Hu FB. Combined impact of lifestyle factors on mortality: prospective cohort study in US women. Bmj. 2008;337:a1440. Epub 2008/09/18. doi: 10.1136/bmj.a1440. PubMed PMID: 18796495; PubMed Central PMCID: PMCPMC2658866.
39. Veronese N, Li Y, Manson JE, Willett WC, Fontana L, Hu FB. Combined associations of body weight and lifestyle factors with all cause and cause specific mortality in men and women: prospective cohort study. Bmj. 2016;355:i5855. Epub 2016/11/26. doi: 10.1136/bmj.i5855. PubMed PMID: 27884868; PubMed Central PMCID: PMCPMC5122318
www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organization for the submitted work other than those detailed above; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.
40. Li Y, Pan A, Wang DD, Liu X, Dhana K, Franco OH, et al. Impact of Healthy Lifestyle Factors on Life Expectancies in the US Population. Circulation. 2018. Epub 2018/05/02. doi: 10.1161/circulationaha.117.032047. PubMed PMID: 29712712; PubMed Central PMCID: PMCPMC6207481.
41. Behrens G, Fischer B, Kohler S, Park Y, Hollenbeck AR, Leitzmann MF. Healthy lifestyle behaviors and decreased risk of mortality in a large prospective study of U.S. women and men. Eur J Epidemiol. 2013;28(5):361-72. Epub 2013/03/28. doi: 10.1007/s10654-013-9796-9. PubMed PMID: 23532745.
42. Chudasama YV, Khunti K, Gillies CL, Dhalwani NN, Davies MJ, Yates T, et al. Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: A longitudinal cohort study. PLoS Med. 2020;17(9):e1003332. Epub 2020/09/23. doi: 10.1371/journal.pmed.1003332. PubMed PMID: 32960883; PubMed Central PMCID: PMCPMC7508366 following competing interests. FZ is funded with an unrestricted educational grant from the NIHR CLAHRC East Midlands to the University of Leicester. KK has acted as a consultant and speaker for Novartis, Novo Nordisk, Sanofi-Aventis, Lilly, Servier, and Merck Sharp & Dohme. He has received grants in support of investigator and investigator-initiated trials from Novartis, Novo Nordisk, Sanofi-Aventis, Lilly, Pfizer, Boehringer Ingelheim, and Merck Sharp & Dohme. KK has received funds for research and honoraria for speaking at meetings and has served on advisory boards for Lilly, Sanofi-Aventis, Merck Sharp & Dohme, and Novo Nordisk. MJD has acted as consultant, advisory board member, and speaker for Novo Nordisk, Sanofi-Aventis, Lilly, Merck Sharp & Dohme, Boehringer Ingelheim, AstraZeneca, and Janssen; an advisory board member for Servier and Gilead Sciences Ltd; and as a speaker for NAPP, Mitsubishi Tanabe Pharma Corporation, and Takeda Pharmaceuticals International Inc. TY has received funding from the Leicester NIHR Leicester BRC. All other authors have declared that no competing interests exist.

43. Rizzuto D, Orsini N, Qiu C, Wang HX, Fratiglioni L. Lifestyle, social factors, and survival after age 75: population based study. Bmj. 2012;345:e5568. Epub 2012/09/01. doi: 10.1136/bmj.e5568. PubMed PMID: 22936786; PubMed Central PMCID: PMCPMC3431442 www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work.
44. An S, Ahn C, Jang J, Lee J, Kang D, Lee JK, et al. Comparison of the Prevalence of Cardiometabolic Disorders and Comorbidities in Korea and the United States: Analysis of the National Health and Nutrition Examination Survey. J Korean Med Sci. 2022;37(18):e149. Epub 20220509. doi: 10.3346/jkms.2022.37.e149. PubMed PMID: 35535376.
45. Choi S, Kim K, Lee JK, Choi JY, Shin A, Park SK, et al. Association between Change in Alcohol Consumption and Metabolic Syndrome: Analysis from the Health Examinees Study. Diabetes Metab J. 2019;43(5):615-26. Epub 20190423. doi: 10.4093/dmj.2018.0128. PubMed PMID: 31237129; PubMed Central PMCID: PMCPMC6834843.
46. Saklayen MG. The Global Epidemic of the Metabolic Syndrome. Curr Hypertens Rep. 2018;20(2):12. Epub 20180226. doi: 10.1007/s11906-018-0812-z. PubMed PMID: 29480368; PubMed Central PMCID: PMCPMC5866840.
47. Gami AS, Witt BJ, Howard DE, Erwin PJ, Gami LA, Somers VK, et al. Metabolic syndrome and risk of incident cardiovascular events and death: a systematic review and meta-analysis of longitudinal studies. J Am Coll Cardiol. 2007;49(4):403-14. Epub 2007/01/30. doi: 10.1016/j.jacc.2006.09.032. PubMed PMID: 17258085.
48. Steptoe A, Zaninotto P. Lower socioeconomic status and the acceleration of aging: An outcome-wide analysis. Proc Natl Acad Sci U S A. 2020;117(26):14911-7. Epub 2020/06/17. doi: 10.1073/pnas.1915741117. PubMed PMID: 32541023; PubMed Central PMCID: PMCPMC7334539.
49. Bosnes I, Nordahl HM, Stordal E, Bosnes O, Myklebust T, Almkvist O. Lifestyle predictors of successful aging: A 20-year prospective HUNT study. PLoS One. 2019;14(7):e0219200. Epub 2019/07/12. doi: 10.1371/journal.pone.0219200. PubMed PMID: 31295289; PubMed Central PMCID: PMCPMC6622492.
50. Frenzel A, Binder H, Walter N, Wirkner K, Loeffler M, Loeffler-Wirth H. The aging human body shape. NPJ Aging Mech Dis. 2020;6:5. Epub 2020/03/29. doi: 10.1038/s41514-020-0043-9. PubMed PMID: 32218988; PubMed Central PMCID: PMCPMC7093543.
51. Hertel J, Friedrich N, Wittfeld K, Pietzner M, Budde K, Van der Auwera S, et al. Measuring Biological Age via Metabonomics: The Metabolic Age Score. J Proteome Res. 2016;15(2):400-10. Epub 2015/12/15. doi: 10.1021/acs.jproteome.5b00561. PubMed PMID: 26652958.

52. Jee H, Jeon BH, Kim YH, Kim HK, Choe J, Park J, et al. Development and application of biological age prediction models with physical fitness and physiological components in Korean adults. Gerontology. 2012;58(4):344-53. Epub 2012/03/22. doi: 10.1159/000335738. PubMed PMID: 22433233.
53. Yoo J, Kim Y, Cho ER, Jee SH. Biological age as a useful index to predict seventeen-year survival and mortality in Koreans. BMC geriatrics. 2017;17(1):7. Epub 2017/01/07. doi: 10.1186/s12877-016-0407-y. PubMed PMID: 28056846; PubMed Central PMCID: PMCPMC5217268.
54. Berezina TN, Rybtsova NN, Rybtsov SA. Comparative Dynamics of Individual Ageing among the Investigative Type of Professionals Living in Russia and Russian Migrants to the EU Countries. Eur J Investig Health Psychol Educ. 2020;10(3):749-62. Epub 2020/07/26. doi: 10.3390/ejihpe10030055. PubMed PMID: 34542509; PubMed Central PMCID: PMCPMC8314298.
55. Lara J, Cooper R, Nissan J, Ginty AT, Khaw KT, Deary IJ, et al. A proposed panel of biomarkers of healthy ageing. BMC Med. 2015;13:222. Epub 2015/09/17. doi: 10.1186/s12916-015-0470-9. PubMed PMID: 26373927; PubMed Central PMCID: PMCPMC4572626.
56. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr. 2008;8:24. Epub 2008/10/02. doi: 10.1186/1471-2318-8-24. PubMed PMID: 18826625; PubMed Central PMCID: PMCPMC2573877.
57. Dugue PA, Bassett JK, Joo JE, Baglietto L, Jung CH, Wong EM, et al. Association of DNA Methylation-Based Biological Age With Health Risk Factors and Overall and Cause-Specific Mortality. Am J Epidemiol. 2018;187(3):529-38. Epub 2017/10/12. doi: 10.1093/aje/kwx291. PubMed PMID: 29020168.
58. Berezina TN, Rybtsov S. Acceleration of Biological Aging and Underestimation of Subjective Age Are Risk Factors for Severe COVID-19. Biomedicines. 2021;9(8). Epub 2021/08/28. doi: 10.3390/biomedicines9080913. PubMed PMID: 34440116; PubMed Central PMCID: PMCPMC8389586.
59. Wu JW, Yaqub A, Ma Y, Koudstaal W, Hofman A, Ikram MA, et al. Biological age in healthy elderly predicts aging-related diseases including dementia. Sci Rep. 2021;11(1):15929. Epub 2021/08/07. doi: 10.1038/s41598-021-95425-5. PubMed PMID: 34354164; PubMed Central PMCID: PMCPMC8342513.
60. Jung CH, Son JW, Kang S, Kim WJ, Kim HS, Kim HS, et al. Diabetes Fact Sheets in Korea, 2020: An Appraisal of Current Status. Diabetes Metab J. 2021;45(1):1-10. Epub 20210113. doi: 10.4093/dmj.2020.0254. PubMed PMID: 33434426; PubMed Central PMCID: PMCPMC7850879.
61. Chi JH, Lee BJ. Risk factors for hypertension and diabetes comorbidity in a Korean population: A cross-sectional study. PLoS One. 2022;17(1):e0262757. Epub 20220119. doi: 10.1371/journal.pone.0262757. PubMed PMID: 35045123; PubMed Central PMCID: PMCPMC8769319.
62. Kim MK, Han K, Koh ES, Kim ES, Lee MK, Nam GE, et al. Blood Pressure and Development of Cardiovascular Disease in Koreans With Type 2

Diabetes Mellitus. Hypertension. 2019;73(2):319-26. PubMed PMID: 30624985.
63. Shin JI, Oh J, Kim HC, Choi D, Yoon YS. Current State of Cardiovascular Research in Korea. Circ Res. 2019;125(12):1141-5. Epub 20191205. doi: 10.1161/circresaha.119.310859. PubMed PMID: 31804914; PubMed Central PMCID: PMCPMC7008968.
64. Lee HH, Cho SMJ, Lee H, Baek J, Bae JH, Chung WJ, et al. Korea Heart Disease Fact Sheet 2020: Analysis of Nationwide Data. Korean Circ J. 2021;51(6):495-503. doi: 10.4070/kcj.2021.0097. PubMed PMID: 34085422; PubMed Central PMCID: PMCPMC8176075.
65. Kim HC. Epidemiology of cardiovascular disease and its risk factors in Korea. Glob Health Med. 2021;3(3):134-41. doi: 10.35772/ghm.2021.01008. PubMed PMID: 34250288; PubMed Central PMCID: PMCPMC8239378.
66. Barbieri S, Mehta S, Wu B, Bharat C, Poppe K, Jorm L, et al. Predicting cardiovascular risk from national administrative databases using a combined survival analysis and deep learning approach. Int J Epidemiol. 2021. Epub 20211215. doi: 10.1093/ije/dyab258. PubMed PMID: 34910160.
67. Sung JM, Cho IJ, Sung D, Kim S, Kim HC, Chae MH, et al. Development and verification of prediction models for preventing cardiovascular diseases. PLoS One. 2019;14(9):e0222809. Epub 20190919. doi: 10.1371/journal.pone.0222809. PubMed PMID: 31536581; PubMed Central PMCID: PMCPMC6752799.
68. Dietrich S, Floegel A, Troll M, Kühn T, Rathmann W, Peters A, et al. Random Survival Forest in practice: a method for modelling complex metabolomics data in time to event analysis. Int J Epidemiol. 2016;45(5):1406-20. Epub 20160901. doi: 10.1093/ije/dyw145. PubMed PMID: 27591264.
69. Hathaway QA, Yanamala N, Budoff MJ, Sengupta PP, Zeb I. Deep neural survival networks for cardiovascular risk prediction: The Multi-Ethnic Study of Atherosclerosis (MESA). Comput Biol Med. 2021;139:104983. Epub 20211029. doi: 10.1016/j.compbiomed.2021.104983. PubMed PMID: 34749095.
70. https://wwwn.cdc.gov/nchs/nhanes.
71. Kweon S, Kim Y, Jang MJ, Kim Y, Kim K, Choi S, et al. Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES). International journal of epidemiology. 2014;43(1):69-77. Epub 2014/03/04. doi: 10.1093/ije/dyt228. PubMed PMID: 24585853; PubMed Central PMCID: PMCPMC3937975.
72. Shin S, Lee HW, Kim CE, Lim J, Lee JK, Lee SA, et al. Egg Consumption and Risk of Metabolic Syndrome in Korean Adults: Results from the Health Examinees Study. Nutrients. 2017;9(7). Epub 2017/07/04. doi: 10.3390/nu9070687. PubMed PMID: 28671590; PubMed Central PMCID: PMCPMC5537802.
73. Kim Y, Han BG. Cohort Profile: The Korean Genome and Epidemiology Study (KoGES) Consortium. International journal of epidemiology. 2017;46(4):1350. Epub 2017/09/25. doi: 10.1093/ije/dyx105. PubMed PMID:

28938752.
74. Chen Y, Copeland WK, Vedanthan R, Grant E, Lee JE, Gu D, et al. Association between body mass index and cardiovascular disease mortality in east Asians and south Asians: pooled analysis of prospective data from the Asia Cohort Consortium. BMJ (Clinical research ed). 2013;347:f5446. Epub 2014/01/30. doi: 10.1136/bmj.f5446. PubMed PMID: 24473060; PubMed Central PMCID: PMCPMC3788174.
75. Boffetta P, Hazelton WD, Chen Y, Sinha R, Inoue M, Gao YT, et al. Body mass, tobacco smoking, alcohol drinking and risk of cancer of the small intestine--a pooled analysis of over 500,000 subjects in the Asia Cohort Consortium. Annals of oncology : official journal of the European Society for Medical Oncology. 2012;23(7):1894-8. Epub 2011/12/08. doi: 10.1093/annonc/mdr562. PubMed PMID: 22147734; PubMed Central PMCID: PMCPMC3493138.
76. Chen Y, Wu F, Saito E, Lin Y, Song M, Luu HN, et al. Association between type 2 diabetes and risk of cancer mortality: a pooled analysis of over 771,000 individuals in the Asia Cohort Consortium. Diabetologia. 2017;60(6):1022-32. Epub 2017/03/08. doi: 10.1007/s00125-017-4229-z. PubMed PMID: 28265721; PubMed Central PMCID: PMCPMC5632944.
77. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373-83. Epub 1987/01/01. PubMed PMID: 3558716.
78. An S, Ahn C, Moon S, Sim EJ, Park SK. Individualized Biological Age as a Predictor of Disease: Korean Genome and Epidemiology Study (KoGES) Cohort. J Pers Med. 2022;12(3). Epub 20220321. doi: 10.3390/jpm12030505. PubMed PMID: 35330504; PubMed Central PMCID: PMCPMC8955355.
79. Whelton PK, Carey RM, Aronow WS, Casey DE, Jr., Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2018;138(17):e484-e594. Epub 2018/10/26. doi: 10.1161/cir.0000000000000596. PubMed PMID: 30354654.
80. World Health Organization & International Diabetes Federation. (2006). Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia : report of a WHO/IDF consultation. Geneva : World Health Organization. http://www.who.int/iris/handle/10665/43588.
81. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation. 2002;106(25):3143-421. Epub 2002/12/18. PubMed PMID: 12485966.
82. World Health Organization. Regional Office for the Western Pacific. (

2000). The Asia-Pacific perspective : redefining obesity and its treatment. Sydney : Health Communications Australia. http://www.who.int/iris/handle/10665/206936.
83. Lee SY, Park HS, Kim DJ, Han JH, Kim SM, Cho GJ, et al. Appropriate waist circumference cutoff points for central obesity in Korean adults. Diabetes Res Clin Pract. 2007;75(1):72-80. Epub 20060602. doi: 10.1016/j.diabres.2006.04.013. PubMed PMID: 16735075.
84. Whelton PK, Carey RM, Aronow WS, Casey DE, Jr., Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2018;71(19):e127-e248. Epub 2017/11/18. doi: 10.1016/j.jacc.2017.11.006. PubMed PMID: 29146535.
85. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2011;34 Suppl 1:S62-9. Epub 2011/01/14. doi: 10.2337/dc11-S062. PubMed PMID: 21193628; PubMed Central PMCID: PMCPMC3006051.
86. Boyle P, Parkin DM. Cancer registration: principles and methods. Statistical methods for registries. IARC Sci Publ. 1991;(95):126-58. Epub 1991/01/01. PubMed PMID: 1894318.
87. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJ, Lozano R, Inoue M. Age standardization of rates: a new who standard. World Health Organization. 2001;GPE Discussion Paper series: No.31.
88. Zou H, Hastie T. Regularization and Variable Selection via the Elastic Net. Journal of the Royal Statistical Society Series B (Statistical Methodology). 2005;67(2):301-20.
89. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. Bioinformatics. 2005;21(15):3301-7. Epub 2005/05/21. doi: 10.1093/bioinformatics/bti499. PubMed PMID: 15905277.
90. Zhang Z. Semi-parametric regression model for survival data: graphical visualization with R. Ann Transl Med. 2016;4(23):461. doi: 10.21037/atm.2016.08.61. PubMed PMID: 28090517; PubMed Central PMCID: PMCPMC5220043.
91. Wongvibulsin S, Wu KC, Zeger SL. Clinical risk prediction with random forests for survival, longitudinal, and multivariate (RF-SLAM) data analysis. BMC Med Res Methodol. 2019;20(1):1. Epub 20191231. doi: 10.1186/s12874-019-0863-0. PubMed PMID: 31888507; PubMed Central PMCID: PMCPMC6937754.
92. Chen Y, Jia Z, Mercola D, Xie X. A gradient boosting algorithm for survival analysis via direct optimization of concordance index. Comput Math Methods Med. 2013;2013:873595. Epub 20131120. doi: 10.1155/2013/873595. PubMed PMID: 24348746; PubMed Central PMCID: PMCPMC3853154.
93. Ebrahimi V, Sharifi M, Mousavi-Roknabadi RS, Sadegh R, Khademian

MH, Moghadami M, et al. Predictive determinants of overall survival among re-infected COVID-19 patients using the elastic-net regularized Cox proportional hazards model: a machine-learning algorithm. BMC Public Health. 2022;22(1):10. Epub 20220105. doi: 10.1186/s12889-021-12383-3. PubMed PMID: 34986818; PubMed Central PMCID: PMCPMC8727465.
94. Jun S, Ha K, Chung S, Joung H. Meat and milk intake in the rice-based Korean diet: impact on cancer and metabolic syndrome. Proc Nutr Soc. 2016;75(3):374-84. Epub 2016/03/16. doi: 10.1017/s0029665116000112. PubMed PMID: 26975473.
95. Cross SH, Mehra MR, Bhatt DL, Nasir K, O'Donnell CJ, Califf RM, et al. Rural-Urban Differences in Cardiovascular Mortality in the US, 1999-2017. Jama. 2020;323(18):1852-4. Epub 2020/05/13. doi: 10.1001/jama.2020.2047. PubMed PMID: 32396176; PubMed Central PMCID: PMCPMC7218488 consulting fees, paid to Brigham and Women's Hospital, from Abbott; fees for serving on a steering committee from Medtronic and Janssen (Johnson & Johnson); fees for serving on a data and safety monitoring board from Mesoblast; consulting fees from Portola, Bayer, and Triple Gene; and fees for serving as a scientific board member from NuPulseCV, Leviticus, and FineHeart. Dr Bhatt disclosed the following relationships—advisory board: Cardax, Cereno Scientific, Elsevier Practice Update Cardiology, Medscape Cardiology, PhaseBio, and Regado Biosciences; board of directors: Boston VA Research Institute, Society of Cardiovascular Patient Care, and TobeSoft; chair: American Heart Association Quality Oversight Committee; data monitoring committees: Baim Institute for Clinical Research, Cleveland Clinic, Duke Clinical Research Institute, Mayo Clinic, Mount Sinai School of Medicine, and Population Health Research Institute; honoraria: American College of Cardiology, Baim Institute for Clinical Research, Belvoir Publications, Duke Clinical Research Institute, HMP Global, Journal of the American College of Cardiology, Medtelligence/ReachMD, Population Health Research Institute, Slack Publications, Society of Cardiovascular Patient Care, and WebMD; other: Clinical Cardiology, NCDR-ACTION Registry Steering Committee, VA CART Research and Publications Committee; research funding: Abbott, Afimmune, Amarin, Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Chiesi, CSL Behring, Eisai, Ethicon, Ferring Pharmaceuticals, Forest Laboratories, Fractyl, Idorsia, Ironwood, Ischemix, Lilly, Medtronic, PhaseBio, Pfizer, PLx Pharma, Regeneron, Roche, Sanofi Aventis, Synaptic, and the Medicines Company; royalties: Elsevier; site co-investigator: Biotronik, Boston Scientific, St Jude Medical, and Svelte; trustee: American College of Cardiology; and unfunded research: FlowCo, Merck, Novo Nordisk, and Takeda. Dr Nasir reported being supported by the Jerold B. Katz Academy of Translational Research. Dr Califf reported serving on the corporate board for Cytokinetics; being the board chair for the People-Centered Research Foundation; and receiving personal fees for consulting from Merck, Amgen, AstraZeneca, Biogen, Genentech, Eli Lilly, and Boehringer Ingelheim; and receiving other funding from Verily Life Sciences and

Google Health. No other disclosures were reported.
96. Lee HY. Socioeconomic Disparities in the Prevalence, Diagnosis, and Control of Hypertension in the Context of a Universal Health Insurance System. J Korean Med Sci. 2017;32(4):561-7. Epub 2017/03/01. doi: 10.3346/jkms.2017.32.4.561. PubMed PMID: 28244279; PubMed Central PMCID: PMCPMC5334151.
97. Chung SJ, Han YS, Lee SI, Kang SH. Urban and rural differences in the prevalence of gender and age specific obesity and related health behaviors in Korea. J Korean Med Sci. 2005;20(5):713-20. Epub 2005/10/15. doi: 10.3346/jkms.2005.20.5.713. PubMed PMID: 16224141; PubMed Central PMCID: PMCPMC2779264.
98. Di Angelantonio E, Kaptoge S, Wormser D, Willeit P, Butterworth AS, Bansal N, et al. Association of Cardiometabolic Multimorbidity With Mortality. Jama. 2015;314(1):52-60. Epub 2015/07/08. doi: 10.1001/jama.2015.7008. PubMed PMID: 26151266; PubMed Central PMCID: PMCPMC4664176.
99. Safar ME, Gnakaméné JB, Bahous SA, Yannoutsos A, Thomas F. Longitudinal Study of Hypertensive Subjects With Type 2 Diabetes Mellitus: Overall and Cardiovascular Risk. Hypertension. 2017;69(6):1029-35. Epub 2017/04/12. doi: 10.1161/hypertensionaha.116.08962. PubMed PMID: 28396537.
100. Valerio L, Peters RJ, Zwinderman AH, Pinto-Sietsma SJ. Association of Family History With Cardiovascular Disease in Hypertensive Individuals in a Multiethnic Population. J Am Heart Assoc. 2016;5(12). Epub 2016/12/23. doi: 10.1161/jaha.116.004260. PubMed PMID: 28003252; PubMed Central PMCID: PMCPMC5210427.
101. Nielsen M, Andersson C, Gerds TA, Andersen PK, Jensen TB, Køber L, et al. Familial clustering of myocardial infarction in first-degree relatives: a nationwide study. Eur Heart J. 2013;34(16):1198-203. Epub 2013/01/09. doi: 10.1093/eurheartj/ehs475. PubMed PMID: 23297314.
102. Mons U, Müezzinler A, Gellert C, Schöttker B, Abnet CC, Bobak M, et al. Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. Bmj. 2015;350:h1551. Epub 2015/04/22. doi: 10.1136/bmj.h1551. PubMed PMID: 25896935; PubMed Central PMCID: PMCPMC4413837 http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.
103. Carter BD, Abnet CC, Feskanich D, Freedman ND, Hartge P, Lewis CE, et al. Smoking and mortality--beyond established causes. N Engl J Med. 2015;372(7):631-40. Epub 2015/02/12. doi: 10.1056/NEJMsa1407211. PubMed PMID: 25671255.

104. Kunzmann AT, Coleman HG, Huang WY, Berndt SI. The association of lifetime alcohol use with mortality and cancer risk in older adults: A cohort study. PLoS Med. 2018;15(6):e1002585. Epub 2018/06/20. doi: 10.1371/journal.pmed.1002585. PubMed PMID: 29920516; PubMed Central PMCID: PMCPMC6007830.
105. Saito E, Inoue M, Sawada N, Charvat H, Shimazu T, Yamaji T, et al. Impact of Alcohol Intake and Drinking Patterns on Mortality From All Causes and Major Causes of Death in a Japanese Population. J Epidemiol. 2018;28(3):140-8. Epub 2017/11/14. doi: 10.2188/jea.JE20160200. PubMed PMID: 29129895; PubMed Central PMCID: PMCPMC5821691.
106. Mukamal KJ, Chen CM, Rao SR, Breslow RA. Alcohol consumption and cardiovascular mortality among U.S. adults, 1987 to 2002. J Am Coll Cardiol. 2010;55(13):1328-35. Epub 2010/03/27. doi: 10.1016/j.jacc.2009.10.056. PubMed PMID: 20338493; PubMed Central PMCID: PMCPMC3865979.
107. Chang JY, Choi S, Park SM. Association of change in alcohol consumption with cardiovascular disease and mortality among initial nondrinkers. Sci Rep. 2020;10(1):13419. Epub 2020/08/10. doi: 10.1038/s41598-020-70304-7. PubMed PMID: 32770048; PubMed Central PMCID: PMCPMC7414908.
108. Burton R, Sheron N. No level of alcohol consumption improves health. Lancet. 2018;392(10152):987-8. Epub 2018/08/28. doi: 10.1016/s0140-6736(18)31571-x. PubMed PMID: 30146328.
109. Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2018;392(10152):1015-35. Epub 2018/08/28. doi: 10.1016/s0140-6736(18)31310-2. PubMed PMID: 30146330; PubMed Central PMCID: PMCPMC6148333.
110. Jee SH, Sull JW, Park J, Lee SY, Ohrr H, Guallar E, et al. Body-mass index and mortality in Korean men and women. N Engl J Med. 2006;355(8):779-87. Epub 2006/08/24. doi: 10.1056/NEJMoa054017. PubMed PMID: 16926276.
111. Primatesta P, Falaschetti E, Gupta S, Marmot MG, Poulter NR. Association between smoking and blood pressure: evidence from the health survey for England. Hypertension. 2001;37(2):187-93. Epub 2001/03/07. doi: 10.1161/01.hyp.37.2.187. PubMed PMID: 11230269.
112. Mattace-Raso FU, van der Cammen TJ, Hofman A, van Popele NM, Bos ML, Schalekamp MA, et al. Arterial stiffness and risk of coronary heart disease and stroke: the Rotterdam Study. Circulation. 2006;113(5):657-63. Epub 2006/02/08. doi: 10.1161/circulationaha.105.555235. PubMed PMID: 16461838.
113. Vlachopoulos C, Alexopoulos N, Panagiotakos D, O'Rourke MF, Stefanadis C. Cigar smoking has an acute detrimental effect on arterial stiffness. Am J Hypertens. 2004;17(4):299-303. Epub 2004/04/06. doi: 10.1016/j.amjhyper.2003.12.014. PubMed PMID: 15062882.
114. Mayl JJ, German CA, Bertoni AG, Upadhya B, Bhave PD, Yeboah J, et al.

Association of Alcohol Intake With Hypertension in Type 2 Diabetes Mellitus: The ACCORD Trial. J Am Heart Assoc. 2020;9(18):e017334. Epub 2020/09/10. doi: 10.1161/jaha.120.017334. PubMed PMID: 32900264; PubMed Central PMCID: PMCPMC7726983.
115. Raal FJ, Alsheikh-Ali AA, Omar MI, Rashed W, Hamoui O, Kane A, et al. Cardiovascular risk factor burden in Africa and the Middle East across country income categories: a post hoc analysis of the cross-sectional Africa Middle East Cardiovascular Epidemiological (ACE) study. Arch Public Health. 2018;76:15. Epub 2018/02/17. doi: 10.1186/s13690-018-0257-5. PubMed PMID: 29449941; PubMed Central PMCID: PMCPMC5812200.
116. Forouzanfar MH, Liu P, Roth GA, Ng M, Biryukov S, Marczak L, et al. Global Burden of Hypertension and Systolic Blood Pressure of at Least 110 to 115 mm Hg, 1990-2015. Jama. 2017;317(2):165-82. Epub 2017/01/18. doi: 10.1001/jama.2016.19043. PubMed PMID: 28097354.
117. Dinesh Shah A, Langenberg C, Rapsomaniki E, Denaxas S, Pujades-Rodriguez M, Gale CP, et al. Type 2 diabetes and incidence of a wide range of cardiovascular diseases: a cohort study in 1.9 million people. Lancet. 2015;385 Suppl 1:S86. Epub 2015/08/28. doi: 10.1016/s0140-6736(15)60401-9. PubMed PMID: 26312908.
118. Read SH, McAllister DA, Colhoun HM, Farran B, Fischbacher C, Kerssens JJ, et al. Incident ischaemic stroke and Type 2 diabetes: trends in incidence and case fatality in Scotland 2004-2013. Diabet Med. 2018;35(1):99-106. Epub 2017/10/19. doi: 10.1111/dme.13528. PubMed PMID: 29044687.
119. Patel N, Chen O, Donahue C, Wang B, Fang Y, Donnino R, et al. Impact of diabetes on heart failure incidence in adults with ischemic heart disease. J Diabetes Complications. 2017;31(11):1597-601. Epub 2017/09/28. doi: 10.1016/j.jdiacomp.2017.07.011. PubMed PMID: 28947278.
120. Tobias DK, Stuart JJ, Li S, Chavarro J, Rimm EB, Rich-Edwards J, et al. Association of History of Gestational Diabetes With Long-term Cardiovascular Disease Risk in a Large Prospective Cohort of US Women. JAMA Intern Med. 2017. Epub 2017/10/20. doi: 10.1001/jamainternmed.2017.2790. PubMed PMID: 29049820.
121. Choi J, Jang J, An Y, Park SK. Blood Pressure and the Risk of Death From Non-cardiovascular Diseases: A Population-based Cohort Study of Korean Adults. J Prev Med Public Health. 2018;51(6):298-309. Epub 2018/12/06. doi: 10.3961/jpmph.18.212. PubMed PMID: 30514060; PubMed Central PMCID: PMCPMC6283742.
122. Sun K, Liu J, Ning G. Active smoking and risk of metabolic syndrome: a meta-analysis of prospective studies. PLoS One. 2012;7(10):e47791. Epub 20121017. doi: 10.1371/journal.pone.0047791. PubMed PMID: 23082217; PubMed Central PMCID: PMCPMC3474781.
123. Sun K, Ren M, Liu D, Wang C, Yang C, Yan L. Alcohol consumption and risk of metabolic syndrome: a meta-analysis of prospective studies. Clin Nutr.

2014;33(4):596-602. Epub 20131014. doi: 10.1016/j.clnu.2013.10.003. PubMed PMID: 24315622.
124. Edwardson CL, Gorely T, Davies MJ, Gray LJ, Khunti K, Wilmot EG, et al. Association of sedentary behaviour with metabolic syndrome: a meta-analysis. PLoS One. 2012;7(4):e34916. Epub 20120413. doi: 10.1371/journal.pone.0034916. PubMed PMID: 22514690; PubMed Central PMCID: PMCPMC3325927.
125. Krishnamoorthy Y, Rajaa S, Murali S, Sahoo J, Kar SS. Association Between Anthropometric Risk Factors and Metabolic Syndrome Among Adults in India: A Systematic Review and Meta-Analysis of Observational Studies. Prev Chronic Dis. 2022;19:E24. Epub 20220505. doi: 10.5888/pcd19.210231. PubMed PMID: 35512304; PubMed Central PMCID: PMCPMC9109643.
126. Putin E, Mamoshina P, Aliper A, Korzinkin M, Moskalev A, Kolosov A, et al. Deep biomarkers of human aging: Application of deep neural networks to biomarker development. Aging. 2016;8(5):1021-33. Epub 2016/05/19. doi: 10.18632/aging.100968. PubMed PMID: 27191382; PubMed Central PMCID: PMCPMC4931851.
127. Emdin CA, Khera AV, Natarajan P, Klarin D, Zekavat SM, Hsiao AJ, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. Jama. 2017;317(6):626-34. Epub 2017/02/15. doi: 10.1001/jama.2016.21042. PubMed PMID: 28196256; PubMed Central PMCID: PMCPMC5571980.
128. Dale CE, Fatemifar G, Palmer TM, White J, Prieto-Merino D, Zabaneh D, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. Circulation. 2017;135(24):2373-88. Epub 2017/05/14. doi: 10.1161/CIRCULATIONAHA.116.026560. PubMed PMID: 28500271; PubMed Central PMCID: PMCPMC5515354.
129. Zhang HS, An S, Ahn C, Park SK, Park B. Obesity measures at baseline, their trajectories over time, and the incidence of chronic kidney disease: A 14 year cohort study among Korean adults. Nutr Metab Cardiovasc Dis. 2021;31(3):782-92. Epub 2021/02/07. doi: 10.1016/j.numecd.2020.10.021. PubMed PMID: 33546946.
130. Kim MK, Ko MJ, Han JT. Alcohol consumption and mortality from all-cause and cancers among 1.34 million Koreans: the results from the Korea national health insurance corporation's health examinee cohort in 2000. Cancer causes & control : CCC. 2010;21(12):2295-302. Epub 2010/10/14. doi: 10.1007/s10552-010-9656-9. PubMed PMID: 20941640.
131. Huzen J, Wong LS, van Veldhuisen DJ, Samani NJ, Zwinderman AH, Codd V, et al. Telomere length loss due to smoking and metabolic traits. Journal of internal medicine. 2014;275(2):155-63. Epub 2013/10/15. doi: 10.1111/joim.12149. PubMed PMID: 24118582.
132. Astuti Y, Wardhana A, Watkins J, Wulaningsih W. Cigarette smoking and telomere length: A systematic review of 84 studies and meta-analysis.

Environmental research. 2017;158:480-9. Epub 2017/07/14. doi: 10.1016/j.envres.2017.06.038. PubMed PMID: 28704792; PubMed Central PMCID: PMCPMC5562268.
133. Wulaningsih W, Serrano FE, Utarini A, Matsuguchi T, Watkins J. Smoking, second-hand smoke exposure and smoking cessation in relation to leukocyte telomere length and mortality. Oncotarget. 2016;7(37):60419-31. Epub 2016/08/11. doi: 10.18632/oncotarget.11051. PubMed PMID: 27509177; PubMed Central PMCID: PMCPMC5312393.
134. Jia L, Zhang W, Chen X. Common methods of biological age estimation. Clin Interv Aging. 2017;12:759-72. Epub 2017/05/27. doi: 10.2147/cia.S134921. PubMed PMID: 28546743; PubMed Central PMCID: PMCPMC5436771.
135. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. BMC Med. 2011;9:103. Epub 20110908. doi: 10.1186/1741-7015-9-103. PubMed PMID: 21902820; PubMed Central PMCID: PMCPMC3180398.
136. Echouffo-Tcheugui JB, Batty GD, Kivimäki M, Kengne AP. Risk models to predict hypertension: a systematic review. PLoS One. 2013;8(7):e67370. Epub 20130705. doi: 10.1371/journal.pone.0067370. PubMed PMID: 23861760; PubMed Central PMCID: PMCPMC3702558.
137. Ramezankhani A, Hadavandi E, Pournik O, Shahrabi J, Azizi F, Hadaegh F. Decision tree-based modelling for identification of potential interactions between type 2 diabetes risk factors: a decade follow-up in a Middle East prospective cohort study. BMJ Open. 2016;6(12):e013336. Epub 20161201. doi: 10.1136/bmjopen-2016-013336. PubMed PMID: 27909038; PubMed Central PMCID: PMCPMC5168628.
138. López-Martínez F, Núñez-Valdez ER, Crespo RG, García-Díaz V. An artificial neural network approach for predicting hypertension using NHANES data. Sci Rep. 2020;10(1):10620. Epub 20200630. doi: 10.1038/s41598-020-67640-z. PubMed PMID: 32606434; PubMed Central PMCID: PMCPMC7327031.
139. Ooka T, Johno H, Nakamoto K, Yoda Y, Yokomichi H, Yamagata Z. Random forest approach for determining risk prediction and predictive factors of type 2 diabetes: large-scale health check-up data in Japan. BMJ Nutr Prev Health. 2021;4(1):140-8. Epub 20210311. doi: 10.1136/bmjnph-2020-000200. PubMed PMID: 34308121; PubMed Central PMCID: PMCPMC8258057.
140. AlKaabi LA, Ahmed LS, Al Attiyah MF, Abdel-Rahman ME. Predicting hypertension using machine learning: Findings from Qatar Biobank Study. PLoS One. 2020;15(10):e0240370. Epub 20201016. doi: 10.1371/journal.pone.0240370. PubMed PMID: 33064740; PubMed Central PMCID: PMCPMC7567367.
141. Silva K, Lee WK, Forbes A, Demmer RT, Barton C, Enticott J. Use and performance of machine learning models for type 2 diabetes prediction in community settings: A systematic review and meta-analysis. Int J Med Inform.

2020;143:104268. Epub 20200907. doi: 10.1016/j.ijmedinf.2020.104268. PubMed PMID: 32950874.
142. Guo CY, Wu MY, Cheng HM. The Comprehensive Machine Learning Analytics for Heart Failure. Int J Environ Res Public Health. 2021;18(9). Epub 20210506. doi: 10.3390/ijerph18094943. PubMed PMID: 34066464; PubMed Central PMCID: PMCPMC8124765.
143. Kraja AT, Hunt SC, Rao DC, Dávila-Román VG, Arnett DK, Province MA. Genetics of hypertension and cardiovascular disease and their interconnected pathways: lessons from large studies. Curr Hypertens Rep. 2011;13(1):46-54. doi: 10.1007/s11906-010-0174-7. PubMed PMID: 21128019; PubMed Central PMCID: PMCPMC3063340.
144. Spooner A, Chen E, Sowmya A, Sachdev P, Kochan NA, Trollor J, et al. A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction. Sci Rep. 2020;10(1):20410. Epub 20201123. doi: 10.1038/s41598-020-77220-w. PubMed PMID: 33230128; PubMed Central PMCID: PMCPMC7683682.
145. Krittanawong C, Virk HUH, Bangalore S, Wang Z, Johnson KW, Pinotti R, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. Sci Rep. 2020;10(1):16057. Epub 20200929. doi: 10.1038/s41598-020-72685-1. PubMed PMID: 32994452; PubMed Central PMCID: PMCPMC7525515.
146. Li Y, Sperrin M, Ashcroft DM, van Staa TP. Consistency of variety of machine learning and statistical models in predicting clinical risks of individual patients: longitudinal cohort study using cardiovascular disease as exemplar. Bmj. 2020;371:m3919. Epub 20201104. doi: 10.1136/bmj.m3919. PubMed PMID: 33148619; PubMed Central PMCID: PMCPMC7610202.
147. Kreitchmann RS, Abad FJ, Ponsoda V, Nieto MD, Morillo D. Controlling for Response Biases in Self-Report Scales: Forced-Choice vs. Psychometric Modeling of Likert Items. Front Psychol. 2019;10:2309. Epub 20191015. doi: 10.3389/fpsyg.2019.02309. PubMed PMID: 31681103; PubMed Central PMCID: PMCPMC6803422.
148. Murabito JM, Nam BH, D'Agostino RB, Sr., Lloyd-Jones DM, O'Donnell CJ, Wilson PW. Accuracy of offspring reports of parental cardiovascular disease history: the Framingham Offspring Study. Ann Intern Med. 2004;140(6):434-40. Epub 2004/03/17. doi: 10.7326/0003-4819-140-6-200403160-00010. PubMed PMID: 15023709.
149. Øygarden H, Fromm A, Sand KM, Eide GE, Thomassen L, Naess H, et al. Can the cardiovascular family history reported by our patients be trusted? The Norwegian Stroke in the Young Study. Eur J Neurol. 2016;23(1):154-9. Epub 2015/08/22. doi: 10.1111/ene.12824. PubMed PMID: 26293608; PubMed Central PMCID: PMCPMC5049640.
150. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ, Jr. Selection Bias Due to Loss to Follow Up in Cohort Studies. Epidemiology. 2016;27(1):91-7. doi: 10.1097/ede.0000000000000409. PubMed PMID: 26484424; PubMed Central

PMCID: PMCPMC5008911.
151. Madden K, Scott T, McKay P, Petrisor BA, Jeray KJ, Tanner SL, et al. Predicting and Preventing Loss to Follow-up of Adult Trauma Patients in Randomized Controlled Trials: An Example from the FLOW Trial. J Bone Joint Surg Am. 2017;99(13):1086-92. doi: 10.2106/jbjs.16.00900. PubMed PMID: 28678121; PubMed Central PMCID: PMCPMC5490332.
152. Larsson SC, Burgess S, Mason AM, Michaëlsson K. Alcohol Consumption and Cardiovascular Disease: A Mendelian Randomization Study. Circ Genom Precis Med. 2020;13(3):e002814. Epub 20200505. doi: 10.1161/circgen.119.002814. PubMed PMID: 32367730; PubMed Central PMCID: PMCPMC7299220.
153. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007;115(7):928-35. doi: 10.1161/circulationaha.106.672402. PubMed PMID: 17309939.
154. Pepe MS, Fan J, Feng Z, Gerds T, Hilden J. The Net Reclassification Index (NRI): a Misleading Measure of Prediction Improvement Even with Independent Test Data Sets. Stat Biosci. 2015;7(2):282-95. Epub 20140823. doi: 10.1007/s12561-014-9118-0. PubMed PMID: 26504496; PubMed Central PMCID: PMCPMC4615606.
155. Pearson TA, LaCroix AZ, Mead LA, Liang KY. The prediction of midlife coronary heart disease and hypertension in young adults: the Johns Hopkins multiple risk equations. Am J Prev Med. 1990;6(2 Suppl):23-8. PubMed PMID: 2383409.
156. Parikh NI, Pencina MJ, Wang TJ, Benjamin EJ, Lanier KJ, Levy D, et al. A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study. Ann Intern Med. 2008;148(2):102-10. doi: 10.7326/0003-4819-148-2-200801150-00005. PubMed PMID: 18195335.
157. Paynter NP, Cook NR, Everett BM, Sesso HD, Buring JE, Ridker PM. Prediction of incident hypertension risk in women with currently normal blood pressure. Am J Med. 2009;122(5):464-71. doi: 10.1016/j.amjmed.2008.10.034. PubMed PMID: 19375556; PubMed Central PMCID: PMCPMC2671636.
158. Kivimäki M, Batty GD, Singh-Manoux A, Ferrie JE, Tabak AG, Jokela M, et al. Validating the Framingham Hypertension Risk Score: results from the Whitehall II study. Hypertension. 2009;54(3):496-501. Epub 20090713. doi: 10.1161/hypertensionaha.109.132373. PubMed PMID: 19597041; PubMed Central PMCID: PMCPMC2828464.
159. Kivimäki M, Tabak AG, Batty GD, Ferrie JE, Nabi H, Marmot MG, et al. Incremental predictive value of adding past blood pressure measurements to the Framingham hypertension risk equation: the Whitehall II Study. Hypertension. 2010;55(4):1058-62. Epub 20100215. doi: 10.1161/hypertensionaha.109.144220. PubMed PMID: 20157053; PubMed Central PMCID: PMCPMC2862166.
160. Kshirsagar AV, Chiu YL, Bomback AS, August PA, Viera AJ, Colindres RE, et al. A hypertension risk score for middle-aged and older adults. J Clin

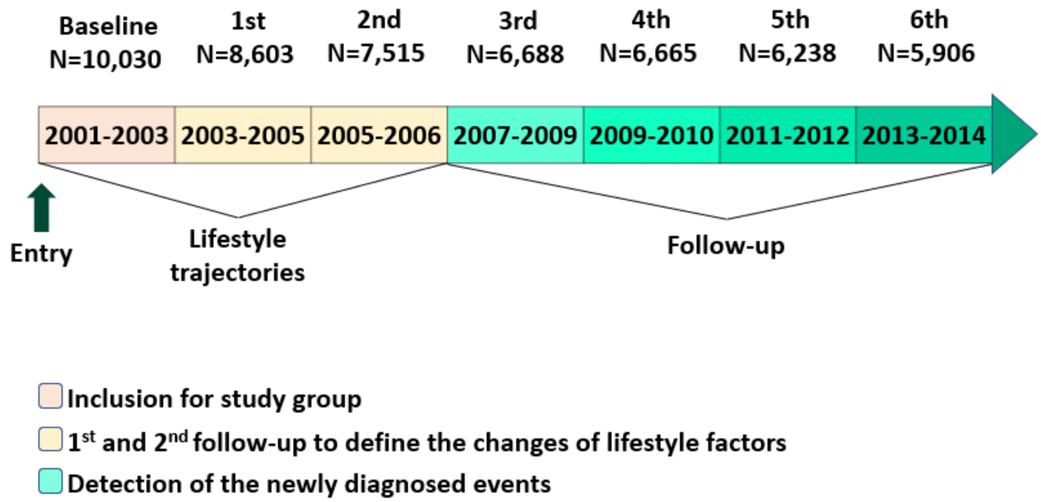
Hypertens (Greenwich). 2010;12(10):800-8. doi: 10.1111/j.1751-7176.2010.00343.x. PubMed PMID: 21029343; PubMed Central PMCID: PMCPMC3683833.
161. Bozorgmanesh M, Hadaegh F, Mehrabi Y, Azizi F. A point-score system superior to blood pressure measures alone for predicting incident hypertension: Tehran Lipid and Glucose Study. J Hypertens. 2011;29(8):1486-93. doi: 10.1097/HJH.0b013e328348fdb2. PubMed PMID: 21720268.
162. Chien KL, Hsu HC, Su TC, Chang WT, Sung FC, Chen MF, et al. Prediction models for the risk of new-onset hypertension in ethnic Chinese in Taiwan. J Hum Hypertens. 2011;25(5):294-303. Epub 20100708. doi: 10.1038/jhh.2010.63. PubMed PMID: 20613783.
163. Lim NK, Son KH, Lee KS, Park HY, Cho MC. Predicting the risk of incident hypertension in a Korean middle-aged population: Korean genome and epidemiology study. J Clin Hypertens (Greenwich). 2013;15(5):344-9. Epub 20130307. doi: 10.1111/jch.12080. PubMed PMID: 23614850; PubMed Central PMCID: PMCPMC8033843.
164. Fava C, Sjögren M, Montagnana M, Danese E, Almgren P, Engström G, et al. Prediction of blood pressure changes over time and incidence of hypertension by a genetic risk score in Swedes. Hypertension. 2013;61(2):319-26. Epub 20121210. doi: 10.1161/hypertensionaha.112.202655. PubMed PMID: 23232644.
165. Choi YH, Chowdhury R, Swaminathan B. Prediction of hypertension based on the genetic analysis of longitudinal phenotypes: a comparison of different modeling approaches for the binary trait of hypertension. BMC Proc. 2014;8(Suppl 1 Genetic Analysis Workshop 18Vanessa Olmo):S78. Epub 20140617. doi: 10.1186/1753-6561-8-s1-s78. PubMed PMID: 25519406; PubMed Central PMCID: PMCPMC4143688.
166. Lim NK, Lee JY, Lee JY, Park HY, Cho MC. The Role of Genetic Risk Score in Predicting the Risk of Hypertension in the Korean population: Korean Genome and Epidemiology Study. PLoS One. 2015;10(6):e0131603. Epub 20150625. doi: 10.1371/journal.pone.0131603. PubMed PMID: 26110887; PubMed Central PMCID: PMCPMC4482533.
167. Otsuka T, Kachi Y, Takada H, Kato K, Kodani E, Ibuki C, et al. Development of a risk prediction model for incident hypertension in a working-age Japanese male population. Hypertens Res. 2015;38(6):419-25. Epub 20141113. doi: 10.1038/hr.2014.159. PubMed PMID: 25391458.
168. Lee JW, Lim NK, Baek TH, Park SH, Park HY. Anthropometric indices as predictors of hypertension among men and women aged 40-69 years in the Korean population: the Korean Genome and Epidemiology Study. BMC Public Health. 2015;15:140. Epub 20150213. doi: 10.1186/s12889-015-1471-5. PubMed PMID: 25886025; PubMed Central PMCID: PMCPMC4332746.
169. Yamakado M, Nagao K, Imaizumi A, Tani M, Toda A, Tanaka T, et al. Plasma Free Amino Acid Profiles Predict Four-Year Risk of Developing Diabetes, Metabolic Syndrome, Dyslipidemia, and Hypertension in Japanese

Population. Sci Rep. 2015;5:11918. Epub 20150709. doi: 10.1038/srep11918. PubMed PMID: 26156880; PubMed Central PMCID: PMCPMC4496670.
170. Lu X, Huang J, Wang L, Chen S, Yang X, Li J, et al. Genetic predisposition to higher blood pressure increases risk of incident hypertension and cardiovascular diseases in Chinese. Hypertension. 2015;66(4):786-92. Epub 20150817. doi: 10.1161/hypertensionaha.115.05961. PubMed PMID: 26283040.
171. Zhang W, Wang L, Chen Y, Tang F, Xue F, Zhang C. Identification of Hypertension Predictors and Application to Hypertension Prediction in an Urban Han Chinese Population: A Longitudinal Study, 2005-2010. Prev Chronic Dis. 2015;12:E184. Epub 20151029. doi: 10.5888/pcd12.150192. PubMed PMID: 26513440; PubMed Central PMCID: PMCPMC4663898.
172. Sathish T, Kannan S, Sarma PS, Razum O, Thrift AG, Thankappan KR. A Risk Score to Predict Hypertension in Primary Care Settings in Rural India. Asia Pac J Public Health. 2016;28(1 Suppl):26s-31s. Epub 20150909. doi: 10.1177/1010539515604701. PubMed PMID: 26354334; PubMed Central PMCID: PMCPMC4724234.
173. Chen Y, Wang C, Liu Y, Yuan Z, Zhang W, Li X, et al. Incident hypertension and its prediction model in a prospective northern urban Han Chinese cohort study. J Hum Hypertens. 2016;30(12):794-800. Epub 20160602. doi: 10.1038/jhh.2016.23. PubMed PMID: 27251078.
174. Niiranen TJ, Havulinna AS, Langén VL, Salomaa V, Jula AM. Prediction of Blood Pressure and Blood Pressure Change With a Genetic Risk Score. J Clin Hypertens (Greenwich). 2016;18(3):181-6. Epub 20151005. doi: 10.1111/jch.12702. PubMed PMID: 26435379; PubMed Central PMCID: PMCPMC8032027.
175. Kanegae H, Oikawa T, Suzuki K, Okawara Y, Kario K. Developing and validating a new precise risk-prediction model for new-onset hypertension: The Jichi Genki hypertension prediction model (JG model). J Clin Hypertens (Greenwich). 2018;20(5):880-90. Epub 20180331. doi: 10.1111/jch.13270. PubMed PMID: 29604170; PubMed Central PMCID: PMCPMC8031110.
176. Wang Y, Ma Z, Xu C, Wang Z, Yang X. Prediction of transfer among multiple states of blood pressure based on Markov model: an 18-year cohort study. J Hypertens. 2018;36(7):1506-13. doi: 10.1097/hjh.0000000000001722. PubMed PMID: 29771738.
177. Xu F, Zhu J, Sun N, Wang L, Xie C, Tang Q, et al. Development and validation of prediction models for hypertension risks in rural Chinese populations. J Glob Health. 2019;9(2):020601. doi: 10.7189/jogh.09.020601. PubMed PMID: 31788232; PubMed Central PMCID: PMCPMC6875679.
178. Syllos DH, Calsavara VF, Bensenor IM, Lotufo PA. Validating the Framingham Hypertension Risk Score: A 4-year follow-up from the Brazilian Longitudinal Study of the Adult Health (ELSA-Brasil). J Clin Hypertens (Greenwich). 2020;22(5):850-6. Epub 20200418. doi: 10.1111/jch.13855. PubMed PMID: 32304277; PubMed Central PMCID: PMCPMC8029849.

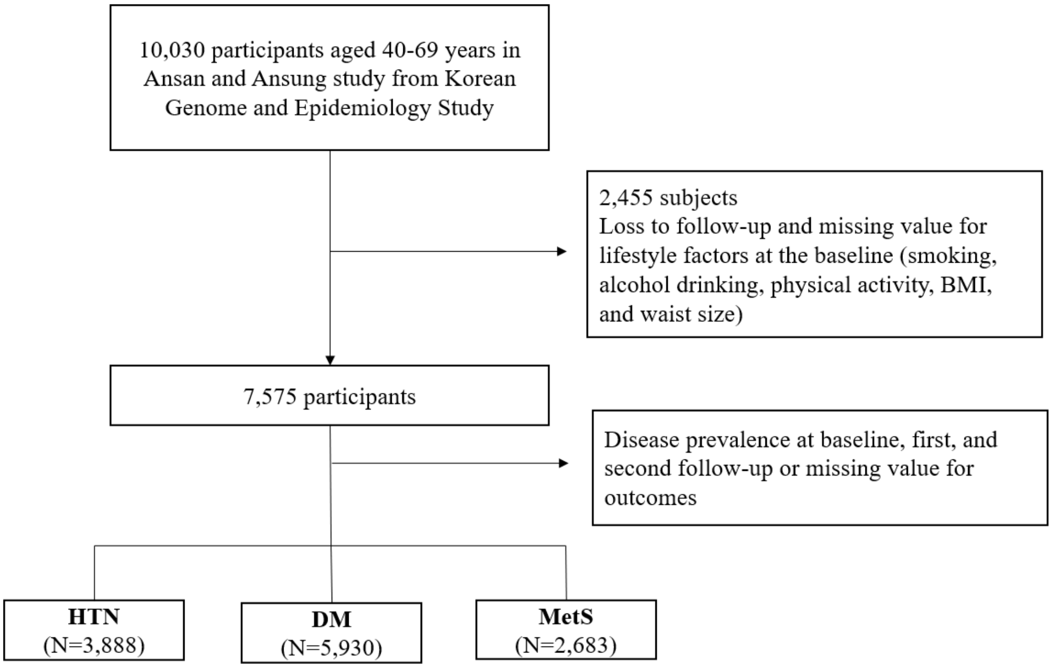
179. Wang B, Liu Y, Sun X, Yin Z, Li H, Ren Y, et al. Prediction model and assessment of probability of incident hypertension: the Rural Chinese Cohort Study. J Hum Hypertens. 2021;35(1):74-84. Epub 20200227. doi: 10.1038/s41371-020-0314-8. PubMed PMID: 32107452.
180. Stern MP, Williams K, Haffner SM. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? Ann Intern Med. 2002;136(8):575-81. doi: 10.7326/0003-4819-136-8-200204160-00006. PubMed PMID: 11955025.
181. Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care. 2003;26(3):725-31. doi: 10.2337/diacare.26.3.725. PubMed PMID: 12610029.
182. Schmidt MI, Duncan BB, Bang H, Pankow JS, Ballantyne CM, Golden SH, et al. Identifying individuals at high risk for diabetes: The Atherosclerosis Risk in Communities study. Diabetes Care. 2005;28(8):2013-8. doi: 10.2337/diacare.28.8.2013. PubMed PMID: 16043747.
183. Aekplakorn W, Bunnag P, Woodward M, Sritara P, Cheepudomwit S, Yamwong S, et al. A risk score for predicting incident diabetes in the Thai population. Diabetes Care. 2006;29(8):1872-7. doi: 10.2337/dc05-2141. PubMed PMID: 16873795.
184. Schulze MB, Hoffmann K, Boeing H, Linseisen J, Rohrmann S, Möhlig M, et al. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. Diabetes Care. 2007;30(3):510-5. doi: 10.2337/dc06-2089. PubMed PMID: 17327313.
185. Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB, Sr. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. Arch Intern Med. 2007;167(10):1068-74. doi: 10.1001/archinte.167.10.1068. PubMed PMID: 17533210.
186. Balkau B, Lange C, Fezeu L, Tichet J, de Lauzon-Guillain B, Czernichow S, et al. Predicting diabetes: clinical, biological, and genetic approaches: data from the Epidemiological Study on the Insulin Resistance Syndrome (DESIR). Diabetes Care. 2008;31(10):2056-61. Epub 20080808. doi: 10.2337/dc08-0368. PubMed PMID: 18689695; PubMed Central PMCID: PMCPMC2551654.
187. Gupta AK, Dahlof B, Dobson J, Sever PS, Wedel H, Poulter NR. Determinants of new-onset diabetes among 19,257 hypertensive patients randomized in the Anglo-Scandinavian Cardiac Outcomes Trial--Blood Pressure Lowering Arm and the relative influence of antihypertensive medication. Diabetes Care. 2008;31(5):982-8. Epub 20080130. doi: 10.2337/dc07-1768. PubMed PMID: 18235048.
188. Chien K, Cai T, Hsu H, Su T, Chang W, Chen M, et al. A prediction model for type 2 diabetes risk among Chinese people. Diabetologia. 2009;52(3):443-50. Epub 20081205. doi: 10.1007/s00125-008-1232-4. PubMed PMID: 19057891.
189. Gao WG, Qiao Q, Pitkäniemi J, Wild S, Magliano D, Shaw J, et al. Risk prediction models for the development of diabetes in Mauritian Indians. Diabet

Med. 2009;26(10):996-1002. doi: 10.1111/j.1464-5491.2009.02810.x. PubMed PMID: 19900231.
190. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. Bmj. 2009;338:b880. Epub 20090317. doi: 10.1136/bmj.b880. PubMed PMID: 19297312; PubMed Central PMCID: PMCPMC2659857.
191. Kahn HS, Cheng YJ, Thompson TJ, Imperatore G, Gregg EW. Two risk-scoring systems for predicting incident diabetes mellitus in U.S. adults age 45 to 64 years. Ann Intern Med. 2009;150(11):741-51. doi: 10.7326/0003-4819-150-11-200906020-00002. PubMed PMID: 19487709.
192. Kolberg JA, Jørgensen T, Gerwien RW, Hamren S, McKenna MP, Moler E, et al. Development of a type 2 diabetes risk model from a panel of serum biomarkers from the Inter99 cohort. Diabetes Care. 2009;32(7):1207-12. doi: 10.2337/dc08-1935. PubMed PMID: 19564473; PubMed Central PMCID: PMCPMC2699726.
193. Chen L, Magliano DJ, Balkau B, Colagiuri S, Zimmet PZ, Tonkin AM, et al. AUSDRISK: an Australian Type 2 Diabetes Risk Assessment Tool based on demographic, lifestyle and simple anthropometric measures. Med J Aust. 2010;192(4):197-202. doi: 10.5694/j.1326-5377.2010.tb03507.x. PubMed PMID: 20170456.
194. Tuomilehto J, Lindström J, Hellmich M, Lehmacher W, Westermeier T, Evers T, et al. Development and validation of a risk-score model for subjects with impaired glucose tolerance for the assessment of the risk of type 2 diabetes mellitus-The STOP-NIDDM risk-score. Diabetes Res Clin Pract. 2010;87(2):267-74. Epub 20091222. doi: 10.1016/j.diabres.2009.11.011. PubMed PMID: 20022651.
195. Liu M, Pan C, Jin M. A Chinese diabetes risk score for screening of undiagnosed diabetes and abnormal glucose tolerance. Diabetes Technol Ther. 2011;13(5):501-7. Epub 20110315. doi: 10.1089/dia.2010.0106. PubMed PMID: 21406016.
196. Alssema M, Vistisen D, Heymans MW, Nijpels G, Glümer C, Zimmet PZ, et al. The Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) update of the Finnish diabetes risk score for prediction of incident type 2 diabetes. Diabetologia. 2011;54(5):1004-12. Epub 20101212. doi: 10.1007/s00125-010-1990-7. PubMed PMID: 21153531.
197. Nanri A, Nakagawa T, Kuwahara K, Yamamoto S, Honda T, Okazaki H, et al. Development of Risk Score for Predicting 3-Year Incidence of Type 2 Diabetes: Japan Epidemiology Collaboration on Occupational Health Study. PLoS One. 2015;10(11):e0142779. Epub 20151111. doi: 10.1371/journal.pone.0142779. PubMed PMID: 26558900; PubMed Central PMCID: PMCPMC4641714.

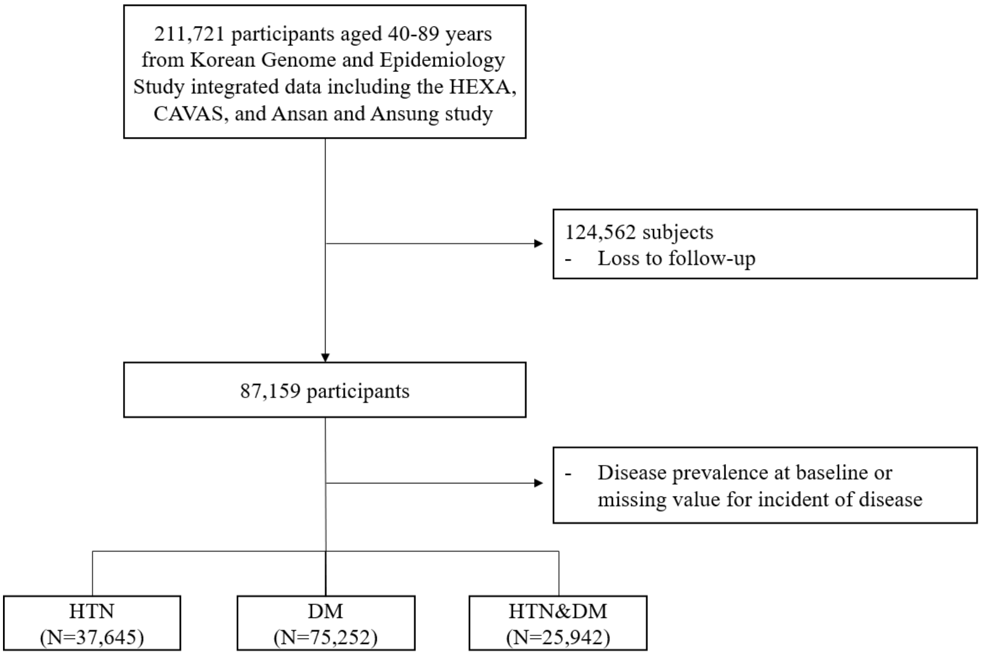
Appendix



Appendix 1. Study design by the timeline of Ansan and Ansung cohort study



Appendix 2. Flow diagram of the study population selection for lifestyle trajectories over time from the Ansan and Ansung follow-up study



Appendix 3. Flow chart of the study population selection for prediction model based on imputation data from the Korean Genome and Epidemiology Study

Appendix 4. Classification of lifestyle trajectories over time

	Category	Definition
Smoking status		
	Constantly never smoker	Constantly never smoker
	Ex-smoker	Ex-smoker
	Constantly decreased	Became non-smoker or decreased dose
	Rise & fall	Rise and fall
	Fall & rise	Fall and rise
	Constantly increased	Constantly smoker with increased dose
Alcohol drinking status		
	Constantly never drinker	Constantly ever drinker
	Ex-drinker	Ex-drinker
	Constantly decreased	Became non-drinker or decreased dose
	Rise & fall	Rise and fall
	Fall & rise	Fall and rise =
	Constantly increased	Constantly drinker with increased dose
Physical activity status		
	Constantly inactive	Constantly inactive
	Decreased	From active to inactive
	Increased	From inactive to active
	Fluctuation	Fluctuation
	Constantly active	Constantly active
BMI status		
	Constantly underweight	Underweight-underweight
	Underweight to normal BMI	Underweight-normal
	Normal BMI to underweight	Normal BMI-under weight
	Constantly normal BMI	Normal BMI-normal BMI
	Became obese	Normal BMI-obesity
	Became non-obese	Obesity-normal BMI
	Constantly obese	Obesity-obesity
Waist size		
	Constantly normal	Normal
	Became abdominal obesity	Became abdominal obesity
	Became non-abdominal obese	Became normal
	Constantly abdominal obese	Abdominal obesity

Appendix 5. The previous risk prediction models for hypertension

Author	Country	Definition of HTN	Risk factors included
Pearson et al, 1990 [155]	USA	Self-reported use of BP medications	Age, SBP, parental history of HTN, and BMI
Parikh et al, 2008 [156]	USA	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, SBP, DBP, parental HTN, and cigarette smoking
Paynter et al, 2009 [157]	USA	Self-reported or SBP≥140 or DBP≥90 mmHg	Age, ethnicity, SBP, DBP, BMI, total grain intake, apolipoprotein B, lipoprotein (a), and C-reactive protein
Paynter et al, 2009 [157]	USA	Self-reported or SBP≥140 or DBP≥90 mmHg	Age, ethnicity, SBP, DBP, BMI, and total to HDL-cholesterol ratio
Paynter et al, 2009 [157]	USA	Self-reported or SBP≥140 or DBP≥90 mmHg	Age, ethnicity, SBP, DBP, total grain intake, apolipoprotein B, lipoprotein (a), and C-reactive protein
Kivimaki et al, 2009 [158]	England	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, parental HTN, BMI, and cigarette smoke
Kivimaki et al, 2010 [159]	England	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, SBP, DBP, parental HTN, BMI, cigarette smoking, and age-DBP interaction
Kshirsagar et al, 2010 [160]	USA	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, SBP, DBP, family history of HTN, DM, BMI, age-DBP interaction, and exercise
Bozorgmanesh et al, 2011 [161]	Iran	SBP≥140 or DBP≥90 mmHg or use of BP medications	Men: SBP, DBP, and cigarette smoking Women: DBP, family history of premature CVD, and waist size
Chien et al, 2011 [162]	Taiwan	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, SBP, DBP, WHC, fasting glucose, uric acid, and BMI
Chien et al, 2011 [162]	Taiwan	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, SBP, DBP, BMI, white blood count, fasting glucose, and uric acid
Lim et al, 2013 [163]	Korea	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, SBP, DBP, parental history of HTN, and BMI

Fava et al, 2013 [164]	Sweden	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, heart rate, BMI, DM, pre-HTN, hypertriglyceridemia, exercise, alcohol consumption, marriage, job, and cigarette smoking,
Choi et al, 2014 [165]	USA	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, age-sex interaction, cigarette smoking, R s10510257 (AA), Rs10510257 (AG), Rs104711 5 (GT)
Lim et al, 2015 [166]	Korean	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, SBP, cigarette smoking, family history of HTN, BMI, and one genetic variable (cGRS or wGRS derived from the 4 SNPs): rs99532 2, rs17249754, rs1378942, rs12945290
Otsuka et al, 2015 [167]	Japan	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, BMI, SBP, DBP, cigarette smoking, alcohol consumption, and parental history of HTN
Lee et al, 2015 [168]	Korea	SBP≥140 or DBP≥90 mmHg or use of BP medications	BMI, waist size, waist-to-hip ratio, and waist-to-height ratio
Yamakado et al, 2015 [169]	Japan	SBP≥140 or DBP≥90 mmHg or use of BP medications	Index I: Leucine, alanine, tyrosine, asparagine, tryptophan, and glycine; Index 2: isoleucine, alanine, tyrosine, phenylalanine, methionine, and histidine
Lu et al, 2015 [170]	China	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, BMI, SBP, DBP, cigarette smoking, alcohol consumption, pulse rate, education level, and genetic risk score
Zhang et al, 2015 [171]	China	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, SBP, DBP, BMI, fasting blood sugar, triglycerides, HDL-cholesterol, hemoglobin, WBC, hematocrit, LC, and NGC
Sathish et al, 2016 [172]	India	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, education level, daily intake of fruits or vegetables, cigarette smoking, alcohol consumption, BP, prehypertension, waist size, and history of high blood glucose

Chen et al, 2016 [173]	China	SBP>140 or DBP>90 mmHg or use of BP medications	Men: age, BMI, SBP, DBP, gamma-GTP, fasting blood glucose, alcohol consumption, age-BMI interaction, and age-DBP interaction Women: age, BMI, SBP, DBP, fasting blood glucose, total cholesterol, neutrophil granulocyte, and alcohol consumption
Niiranen et al, 2016 [174]	Finland	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, history of DM, cigarette smoking, education level, hypercholesterolemia, exercise, and BMI
Kanegae et al, 2018 [175]	Japan	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, BMI, SBP, DBP, LDL-cholesterol, uric acid, proteinuria, cigarette smoking, alcohol consumption, eating rate, DBP by age, and BMI by age
Wang et al, 2018 [176]	China	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, education level, marriage, cigarette smoking, alcohol consumption, BMI, and intake of energy, carbo, fat, and protein
Xu et al, 2019 [177]	China	SBP≥140 or DBP≥90 mmHg or use of BP medications	Men: age, SBP, DBP, parental history of HTN, waist size, and age-DBP interaction Women: age, SBP, DBP, waist size, intake of fruit and vegetable, parental history of HTN, age-waist size interaction, and age-DBP interaction
Syllos et al, 2020 [178]	Brazil	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, sex, SBP, DBP, education level, parental history of HTN, exercise, BMI, neck circumstance, and cigarette smoking
Wang et al, 2021 [179]	China	SBP≥140 or DBP≥90 mmHg or use of BP medications	Age, SBP, DBP, BMI, age by BMI, and parental history of HTN

Appendix 6. The previous risk prediction models for diabetes mellitus

Author	Country	Definition of HTN	Predictors included
Stern et al, 2002 [180]	USA	Fasting plasma glucose≥126 mg/dL, 2-hour fasting glucose ≥200mg/dL or current use of insulin or oral antidiabetic agent	Age, sex, ethnicity, fasting plasma glucose, SBP, HDL cholesterol, BMI, and family history of DM
Lindstrom et al, 2003 [181]	Finland	Fasting plasma glucose≥126 mg/dL or 2-hour fasting glucose ≥200mg/dL	Age, BMI, waist size, BP medication, history of high blood glucose, exercise, and daily consumption of vegetables
Schmidt et al, 2005 [182]	USA	Fasting plasma glucose≥126 mg/dL, 2-hour fasting glucose ≥200mg/dL, current use of insulin or oral antidiabetic agent or report of clinical diagnosis	Age, ethnicity, fasting plasma glucose, parental history of DM, SBP, waist size, height, HDL- cholesterol, and triglycerides
Aekplakorn et al, 2006 [183]	Thailand	Fasting plasma glucose≥126 mg/dL, 2-hour fasting glucose ≥200mg/dL or previous diagnosis of DM	Age, sex, BMI, waist size, HTN, and family history of DM
Schulze et al, 2007 [184]	Germany	Self-reports of DM or use of DM medication or dietary treatment	Age, waist size, HTN, intake of red meat, intake of whole-grain bread, coffee consumption, alcohol consumption, exercise, and cigarette smoking
Wilson et al, 2007 [185]	USA	Fasting plasma glucose≥126 mg/dL or use of DM medication	Fasting plasma glucose, BMI, HDL-cholesterol, parental history of DM, triglyceride level, and blood pressure
Balkau et al, 2008 [186]	France	Fasting plasma glucose≥126 mg/dL or treatment for DM	Men: HTN, waist size, and cigarette smoking Women: HTN, waist size, and family history of DM
Gupta et al, 2008 [187]	Europe	Fasting plasma glucose≥126 mg/dL, 2-hour fasting glucose ≥200mg/dL, self-	Age, sex, fasting plasma glucose, BMI, SBP, triglycerides, total cholesterol, HDL-

		reported history of DM and medication or dietary therapy for DM.	cholesterol, alcohol consumption, and use of non-coronary artery disease medication
Chien et al, 2009 [188]	Taiwan	Fasting plasma glucose≥126 mg/dL, current use of insulin or oral antidiabetic agent	Age, fasting plasma glucose, BMI, WBC, HDL-cholesterol, and triglycerides
Gao et al, 2009 [189]	Mauritius	Fasting plasma glucose≥126 mg/dL, or use of DM medication	Age, sex, BMI, waist size, family history of DM
Hippisley-Cox et al, 2009 [190]	UK	Identified by electronic health records (C10)	Age, BMI, family history of DM, cigarette smoking, HTN, history of CVD, social deprivation, ethnicity, and current treatment with corticosteroids
Kahn et al, 2009 [191]	USA	Self-reported s history of DM or identified by hospital records	Age, Parental history of DM, HTN, ethnicity, cigarette smoking, waist size, height, resting pulse, and weight
Kolberg et al, 2009 [192]	Denmark	Fasting plasma glucose≥126 mg/dL or 2-hour fasting glucose ≥200mg/dL	Adiponectin, C-reactive protein, ferritin, interleukin 2 receptor A, glucose, and insulin
Chen et al, 2010 [193]	Australia	Fasting plasma glucose≥126 mg/dL, 2-hour fasting glucose ≥200mg/dL, or current use of insulin or oral antidiabetic agent	Age, sex, ethnicity, parental history of DM, history of high blood glucose, use of antihypertensive medication, cigarette smoking, exercise, and waist size
Tuomilehto et al, 2010 [194]	Canada, Germany, Austria, Norway, Denmark, Sweden, Finland, Israel, Spain	2-hour fasting glucose ≥200mg/dL	Sex, fasting glucose level, history of HTN, history of CVD, height, acarbose treatment, and serum triglyceride
Liu et al, 2011 [195]	China	Fasting plasma glucose≥126 mg/dL, 2-hour fasting glucose ≥200mg/dL, use of DM medication or self-reported history of DM	Age, history of high blood glucose, HTN, BMI, and high fasting plasma glucose

Alssema et al, 2011 [196]	Netherlands, Denmark, Australia, UK, Sweden, Hungary	2-hour fasting glucose ≥200mg/dL	Age, sex, BMI, waist size, use of anti-hypertensives, history of gestational DM, cigarette smoking, and family history of DM
Nanri et al, 2015 [197]	Japan	Fasting plasma glucose≥126 mg/dL, RBS, or HbA1c≥6.5%, or use of DM medication	Age, sex, BMI, waist size, cigarette smoking, history of HTN, fasting plasma glucose, and HbA1c
Ramezankhani et al, 2016 [137]	Iran	Fasting plasma glucose≥126 mg/dL, 2-hour fasting glucose ≥200mg/dL, or use of DM medication	Men: age, 2-hour plasma glucose, fasting plasma glucose, waist size, waist-to-hip ratio, waist-to-height ratio, cholesterol-to-HDL ratio, BP, history of hospitalization, family history of DM, secondhand smoke, goitre size, use of aspirin, and education level Women: age, 2-hour plasma glucose, fasting plasma glucose, waist size, BMI, waist-to-hip ratio, waist-to-height ratio, cholesterol-to-HDL ratio, triglyceride-to-HDL ratio, BP, pulse rate, glomerular filtration rate, total length of stay in the city, goitre size, family history of DM, use of the ACE inhibitors, current status of pregnancy, use of aspirin, education level, and family history of premature CVD

Appendix 7. Age-standardized prevalence rates [1] of cardiometabolic disorders according to observational year in each age group

		US representative population (NHANES)	**Korean representative population** (KNHANES)	**Korean urban population** (HEXA-KoGES)	**Korean rural population** (HEXA-CAVAS)
Hypertension					
Age 40-49	ASPR in year < 2010	46.0	39.3	40.5	53.6
	ASPR in year ≥ 2010	43.0	41.9	38.7	46.1
	Rate change[2]	**-3.0**	**2.6**	**-1.8**	**-7.5**
Age 50-49	ASPR in year < 2010	60.2	53.3	54.5	65.2
	ASPR in year ≥ 2010	61.2	54.8	54.0	58.4
	Rate change	**1.0**	**1.5**	**-0.5**	**-6.8**
Age 60-49	ASPR in year < 2010	74.1	59.7	66.2	71.7
	ASPR in year ≥ 2010	73.4	60.5	67.7	66.9
	Rate change	**-0.7**	**0.8**	**1.5**	**-4.8**
Diabetes mellitus					
Age 40-49	ASPR in year < 2010	7.4	6.3	3.1	4.6
	ASPR in year ≥ 2010	9.7	7.7	4.6	5.0
	Rate change	**2.3**	**1.4**	**1.5**	**0.4**
Age 50-49	ASPR in year < 2010	16.0	12.2	6.4	7.1
	ASPR in year ≥ 2010	16.7	15.0	10.3	10.3
	Rate change	**0.7**	**2.8**	**3.9**	**3.2**
Age 60-49	ASPR in year < 2010	23.3	21.1	9.8	8.9
	ASPR in year ≥ 2010	22.4	24.9	18.6	15.2
	Rate change	**-0.9**	**3.8**	**8.8**	**6.3**
Hypercholesterolemia					
Age 40-49	ASPR in year < 2010	17.5	9.0	7.4	9.4
	ASPR in year ≥ 2010	14.2	11.2	8.5	6.6

	Rate change	**-3.3**	**2.2**	**1.1**	**-2.8**
Age 50-49	ASPR in year < 2010	21.4	16.4	14.3	16.3
	ASPR in year ≥ 2010	17.9	21.2	15.5	10.3
	Rate change	**-3.5**	**4.8**	**1.2**	**-6.0**
Age 60-49	ASPR in year < 2010	19.2	19.1	12.7	15.7
	ASPR in year ≥ 2010	14.1	27.1	12.7	11.7
	Rate change	**-5.1**	**8.0**	**0**	**-4.0**
Hypertriglyceridemia					
Age 40-49	ASPR in year < 2010	16.6	16.9	11.0	16.7
	ASPR in year ≥ 2010	17.7	17.5	12.0	17.7
	Rate change	**1.1**	**0.6**	**1.0**	**1.0**
Age 50-49	ASPR in year < 2010	18.9	21.0	13.9	20.7
	ASPR in year ≥ 2010	13.6	21.0	13.7	19.1
	Rate change	**-5.3**	**0**	**-0.2**	**-1.6**
Age 60-49	ASPR in year < 2010	19.8	19.8	14.1	21.3
	ASPR in year ≥ 2010	12.6	18.6	13.6	17.4
	Rate change	**-7.2**	**-1.2**	**-0.5**	**-3.9**
Obesity					
Age 40-49	ASPR in year < 2010	36.2	34.1	27.4	39.1
	ASPR in year ≥ 2010	38.9	34.7	27.5	39.7
	Rate change	**2.7**	**0.6**	**0.1**	**0.6**
Age 50-49	ASPR in year < 2010	38.4	39.5	34.0	44.7
	ASPR in year ≥ 2010	41.5	35.6	32.7	43.6
	Rate change	**3.1**	**-3.9**	**-1.3**	**-1.1**
Age 60-49	ASPR in year < 2010	39.8	38.9	38.6	39.9
	ASPR in year ≥ 2010	40.9	38.2	36.5	43.9
	Rate change	**1.1**	**-0.7**	**-2.1**	**4.0**
Metabolic syndrome					
Age 40-49	ASPR in year < 2010	31.4	25.0	13.0	27.2

	ASPR in year ≥ 2010	32.5	21.9	12.6	22.1
	Rate change	**1.1**	**-3.1**	**-0.4**	**-5.1**
Age 50-49	ASPR in year < 2010	37.6	34.5	20.9	37.3
	ASPR in year ≥ 2010	34.5	30.8	19.0	31.9
	Rate change	**-3.1**	**-3.7**	**-1.9**	**-5.4**
Age 60-49	ASPR in year < 2010	47.0	44.3	29.0	42.9
	ASPR in year ≥ 2010	43.2	36.6	25.7	37.0
	Rate change	**-3.8**	**-7.7**	**-3.3**	**-5.9**

1. Age-standardized prevalence rates (ASPRs) (per 100 persons) were calculated using the WHO world standard population.
2. ‘Rate change’ was calculated as [(ASPR in recent year ≥ 2010) – (ASPR in past year < 2010).

Appendix 8. Comorbidity rates (age-standardized prevalence per 100 persons) in each study

	US representative population (NHANES)	**Top 20 rankings**	**Korean representative population** (KNHANES)	**Top 20 rankings**	**Korean urban population** (HEXA-KoGES)	**Top 20 rankings**	**Korean rural population** (HEXA-CAVAS)	**Top 20 rankings**
None	**24.4 (23.0-25.8) ***	1	**29.3 (28.6-30.3)**	1	**30.9 (30.7-31.2) ***	1	**19.5 (19.0-2.00) ***	2
One	**31.5 (30.0-33.0)**		**31.1 (30.2-31.8)**		**35.1 (34.9-34.4) ***		**34.1 (33.5-34.7) ***	
HTN	17.3 (16.1-18.6)	2	17.1 (16.4-17.8)	2	**21.8 (21.6-22.0) ***	2	**22.6 (22.1-23.1) ***	1
Obesity	7.2 (6.4-8.1)	4	7.9 (7.4-8.3)	4	**12.7 (12.5-12.9) ***	4	7.1 (6.8-7.4) *	4
HC	**4.2 (3.5-4.8) ***	6	2.0 (1.8-2.3)	10	**2.8 (2.7-2.9) ***	6	1.9 (1.7-2.0)	9
HTG	1.5 (1.2-1.9) *	13	2.4 (2.1-2.7)	8	1.8 (1.8-1.9) *	10	1.8 (1.6-1.9) *	11
DM	1.3 (0.9-1.6) *	17	1.8 (1.6-2.0)	11	1.1 (1.0-1.1) *	14	0.7 (0.6-0.8) *	18
Two	**26.0 (24.6-27.4) ***		**23.2 (22.6-24.1)**		**23.1 (22.9-23.4)**		**29.7 (29.2-30.3) ***	
HTN, Obesity	12.5 (11.4-13.5)	3	11.6 (11.1-12.2)	3	**12.7 (12.5-12.9) ***	3	**17.2 (16.7-17.7) ***	3
HTN, HTG	2.6 (2.1-3.2) *	9	3.6 (3.2-3.9)	6	2.6 (2.5-2.7) *	8	**4.5 (4.2-4.8) ***	6
HTN, DM	2.4 (2.0-2.9)	10	2.2 (1.9-2.4)	9	1.7 (1.6-1.8) *	12	1.5 (1.4-1.7) *	13
HTN, HC	**3.5 (2.9-4.1) ***	7	1.7 (1.5-1.9)	12	**2.7 (2.6-2.8) ***	7	2.8 (2.5-3.0) *	8
HTG, Obesity	1.0 (0.7-1.2) *	19	1.6 (1.4-1.9)	13	1.1 (1.1-1.2) *	13	1.4 (1.3-1.6)	14
DM, Obesity	1.2 (0.9-1.5)	18	1.0 (0.8-1.2)	17	0.5 (0.5-0.6) *	19	0.5 (0.4-0.6) *	21
HC, Obesity	0.9 (0.6-1.2)	21	0.9 (0.7-1.0)	19	0.9 (0.9-1.0)	15	1.0 (0.9-1.2)	16
HC, HTG	**1.4 (1.0-1.8) ***	15	0.6 (0.4-0.7)	22	0.5 (0.4-0.5)	20	0.5 (0.4-0.6)	22
Three	**13.4 (12.4-14.4) ***		**11.8 (11.5-12.3)**		**8.7 (8.5-8.8) ***		**12.9 (12.5-13.3) ***	
HTN, HTG, Obesity	**3.1 (2.5-3.6) ***	8	4.3 (3.9-4.7)	5	**3.1 (3.0-3.2) ***	5	**5.6 (5.3-5.9) ***	5

HTN, DM, Obesity	**5.2 (4.6-5.9) ***	5	2.8 (2.5-3.1)	7	1.7 (1.7-1.8) *	11	1.7 (1.5-1.8) *	12
HTN, HC, Obesity	**1.8 (1.4-2.2) ***	12	1.3 (1.2-1.5)	15	**1.9 (1.8-2.0) ***	9	**2.9 (2.7-3.2) ***	7
HTN, DM, HTG	0.8 (0.5-1.0)	22	1.0 (0.8-1.1)	18	0.4 (0.4-0.5) *	21	0.6 (0.5-0.7) *	20
HTN, HC, HTG	**1.5 (1.1-1.9) ***	14	0.8 (0.7-1.0)	20	0.7 (0.6-0.7) *	18	**1.2 (1.1-1.3) ***	15
HC, HTG, Obesity	0.4 (0.2-0.6)	23	0.5 (0.4-0.6)	23	0.3 (0.1-0.4) *	23	0.4 (0.3-0.5) *	23
Four	**4.7 (4.0-5.4)**		**4.6 (4.2-5.2)**		**2.2 (2.1-2.2) ***		**3.8 (3.5-4.0) ***	
HTN, HC, HTG, Obesity	1.3 (1.0-1.7)	16	1.1 (0.8-1.3)	16	0.9 (0.8-0.9) *	16	**1.8 (1.7-2.0) ***	10
HTN, DM, HC, Obesity	**1.0 (0.7-1.2) ***	20	0.7 (0.5-0.8)	21	0.4 (0.4-0.5) *	22	0.7 (0.6-0.8)	19
HTN, DM., HTG, Obesity	1.9 (1.5-2.3)	11	1.4 (1.2-1.7)	14	0.7 (0.7-0.8) *	17	0.9 (0.8-1.0) *	17

Abbreviations: HTN, Hypertension; HC, High cholesterolemia; HTG. High triglyceridemia; DM, diabetes mellitus
* p<0.05 for the test for the difference between each group and the KNHANES
Gray colored cells and Bold font: The prevalence rates in each group were higher than those in the KNHANES
Underlined value and normal font: The prevalence rates in each group were lower than those in the KNHANES

Appendix 9. Association between metabolic disease status at the baseline and the risk of cardiovascular disease risk

Characteristics	No. of participants	CVD	
		No. of CVD	Hazard Ratio[2] (95% CI)
Family history of CVD			
No	57,942	1,170	1.00
Yes	14,169	349	1.28 (1.13-1.44)
DM			
No	67,201	1,329	1.00
Yes	4,910	190	1.34 (1.15-1.57)
HTN			
No	57,819	1,028	1.00
Yes	14,289	491	1.33 (1.18-1.49)
LIP			
No	65,407	1,350	1.00
Yes	6,700	169	1.06 (0.90-1.24)
Combination of disease			
None	22,056	225	1.00
DM	20,129	441	1.47 (0.96-2.26)
HTN	989	23	1.51 (1.28-1.78)
LIP	8,875	150	1.32 (1.07-1.63)
DM and HTN	2,066	77	1.95 (1.49-2.54)
DM and LIP	14,211	452	2.24 (1.54-3.24)
HTN and LIP	897	33	2.00 (1.69-2.36)
DM, HTN, and LIP	2,888	118	2.25 (1.79-2.84)
Disease score			
None	22,056	225	1.00
1 disease	29,993	614	1.46 (1.25-1.70)
2 diseases	17,174	562	2.00 (1.70-2.35)
3 diseases	2,888	118	2.25 (1.78-2.84)

Abbreviation, Hypertension (HTN); Diabetes mellitus (DM); Dyslipidemia (LIP)

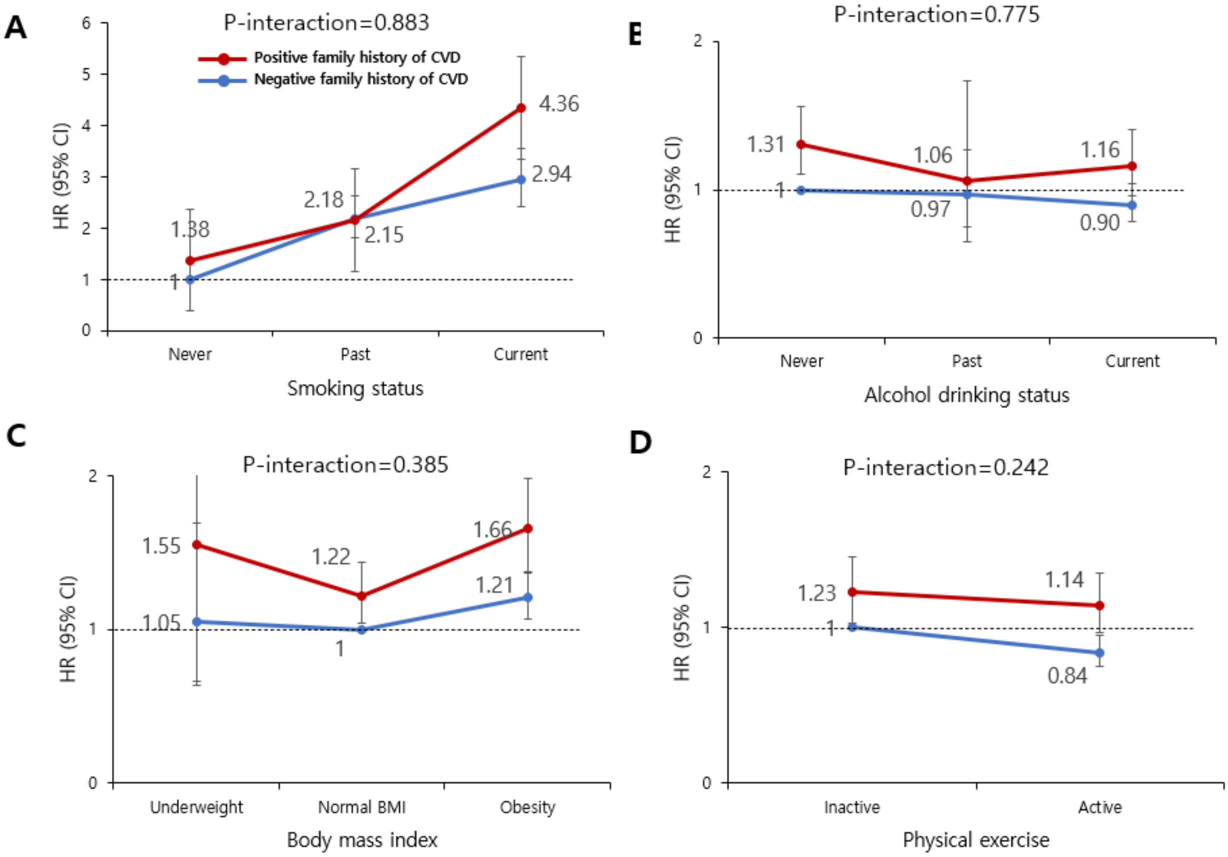
1. Cox proportional hazards regression model were adjusted by sex, age at baseline, body mass index, waist and hip ratio, current smoking status, current alcohol consumption, regular exercise, and family history of cardiovascular disease

Appendix 10. Association between metabolic disease status at the baseline and the risk of myocardial infarction and stroke

Characteristics	**No. of participants**	**Myocardial infarction**		**stroke**	
		No. of MI	**Hazard Ratio[2] (95% CI)**	**No. of stroke**	**Hazard Ratio[1] (95% CI)**
Family history of CVD					
No	57,942	755	1.00	432	1.00
Yes	14,169	228	1.28 (1.11-1.49)	127	1.31 (1.05-1.77)
DM					
No	67,201	861	1.00	488	1.00
Yes	4,910	122	1.36 (1.12-1.65)	71	1.30 (1.01-1.68)
HTN					
No	57,819	668	1.00	369	1.00
Yes	14,289	315	1.29 (1.12-1.49)	190	1.45 (1.20-1.74)
LIP					
No	65,407	849	1.00	522	1.00
Yes	6,700	134	1.32 (1.10-1.58)	37	0.63 (0.45-0.88)
Combination of disease					
None	22,056	142	1.00	84	1.00
DM	20,129	282	1.57 (0.92-2.68)	165	1.30 (0.63-2.68)
HTN	989	15	1.54 (1.25-1.89)	8	1.53 (1.17-2.01)
LIP	8,875	109	1.55 (1.21-2.00)	43	0.99 (0.68-1.43)
DM and HTN	2,066	48	1.97 (1.41-2.75)	32	2.12 (1.39-3.22)
DM and LIP	14,211	288	2.72 (1.75-4.21)	173	1.47 (0.74-2.95)
HTN and LIP	897	24	2.05 (1.66-2.54)	9	1.99 (1.51-2.62)
DM, HTN, and LIP	2,888	75	2.37 (1.77-3.17)	45	2.13 (1.46-3.12)
Disease score					
None	22,056	142	1.00	84	1.00
1 disease	29,993	406	1.55 (1.27-1.88)	216	1.37 (1.05-1.77)
2 diseases	17,174	360	2.08 (1.69-2.55)	214	1.96 (1.50-2.57)
3 diseases	2,888	75	2.37 (1.77-3.17)	45	2.12 (1.45-3.10)

Abbreviation, Hypertension (HTN); Diabetes mellitus (DM); Dyslipidemia (LIP)

1. Cox proportional hazards regression model were adjusted by sex, age at baseline, body mass index, waist and hip ratio, current smoking status, current alcohol consumption, regular exercise, and family history of cardiovascular disease



Appendix 11. Associations of smoking (A), alcohol drinking (B), body mass index (C), and physical activity (D) with cardiovascular disease risk according to family history of cardiovascular disease; Multivariable cox proportional hazards regression model were adjusted by sex, age at baseline, waist and hip ratio, current smoking status, current alcohol consumption, regular exercise, hypertension, diabetes mellitus, and dyslipidemia

Appendix 12. Risk for total and premature cardiovascular death according to lifestyle factors and cardiometabolic diseases (based on the BMI level definition of 18.5-24.9)

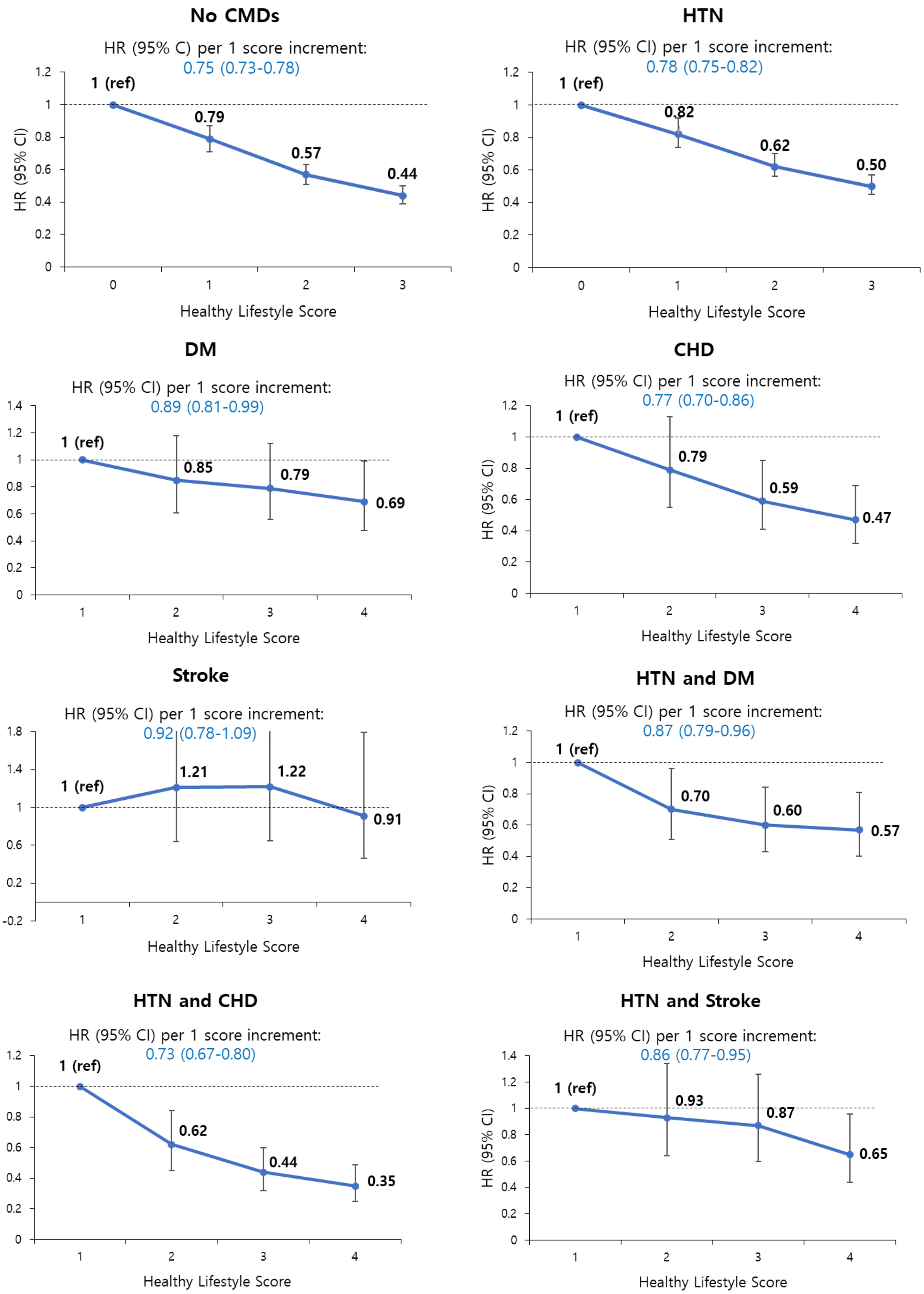
Characteristics	Cohort	CVD death (N=19,442)		Premature CVD death (N=5,774)	
	N	N	HR (95% CI)[2]	N	HR (95% CI)[1]
Healthy lifestyle factors					
Cigarette smoking					
Never	243,481	8,617	1.00	2,491	1.00
Past	40,310	2,682	1.20 (1.14-1.26)	673	1.24 (1.12-1.37)
Current	120,061	5,853	1.77 (1.70-1.85)	2,610	2.01 (1.87-2.17)
[Unhealthy]: Ever	160,371	8,535	1.00	3,283	1.00
[Healthy]: Never	243,481	8,617	0.63 (0.60-0.65)	2,491	0.55 (0.51-0.59)
Alcohol drinking					
Never	233,625	8,829	1.00	2,629	1.00
Past	9,294	1,006	1.57 (1.46-1.68)	261	2.01 (1.76-2.29)
Current	160,933	7,317	1.03 (0.99-1.07)	2,884	1.17 (1.10-1.24)
[Unhealthy]: Ever	170,227	8,323	1.00	3,145	1.00
[Healthy]: Never	233,625	8,829	0.93 (0.90-0.96)	2,629	0.84 (0.79-0.89)
BMI (kg/m^2)					
<18.5	15,681	1,054	1.48 (1.39-1.58)	250	1.45 (1.27-1.66)
18.5-22.9	165,795	6,778	1.00	2,241	1.00
23.0-24.9	101,581	3,947	0.93 (0.89-0.97)	1,381	0.93 (0.87-0.99)
25.0-27.4	77,685	3,081	0.91 (0.87-0.95)	1,063	0.90 (0.84-0.97)
27.5-29.9	30,339	1,438	1.02 (0.96-1.08)	507	1.07 (0.97-1.18)
≥30.0	12,771	854	1.38 (1.28-1.48)	332	1.71 (1.52-1.92)
[Unhealthy]: <18.5 or ≥ 25.0	136,476	6,427	1.00	2,063	1.00
[Healthy]: 18.5-24.9	267,376	10,725	0.93 (0.90-0.96)	3,529	0.92 (0.87-0.97)
Prior cardiometabolic diseases at baseline					
Hypertension					
No	316,412	9,808	1.00	3,702	1.00
Yes	87,440	7,344	1.63 (1.58-1.68)	2,072	1.96 (1.85-2.09)
Diabetes mellitus					
No	383,363	15,068	1.00	5,094	1.00
Yes	20,489	2,084	1.63 (1.55-1.71)	680	2.17 (2.00-2.36)
Chronic heart disease					
No	388,605	15,165	1.00	5,318	1.00
Yes	15,247	1,987	1.67 (1.59-1.75)	456	1.95 (1.77-2.16)
Stroke					
No	397,968	15,847	1.00	5,441	1.00
Yes	5,884	1,305	2.69 (2.54-2.86)	333	3.75 (3.33-4.21)

Appendix 13. Healthy lifestyle score (HLS)[1] for death and premature death[2] from all-cause and CVD according to prior cardiometabolic diseases (CMDs)[3] in the 403,852 ACC participants

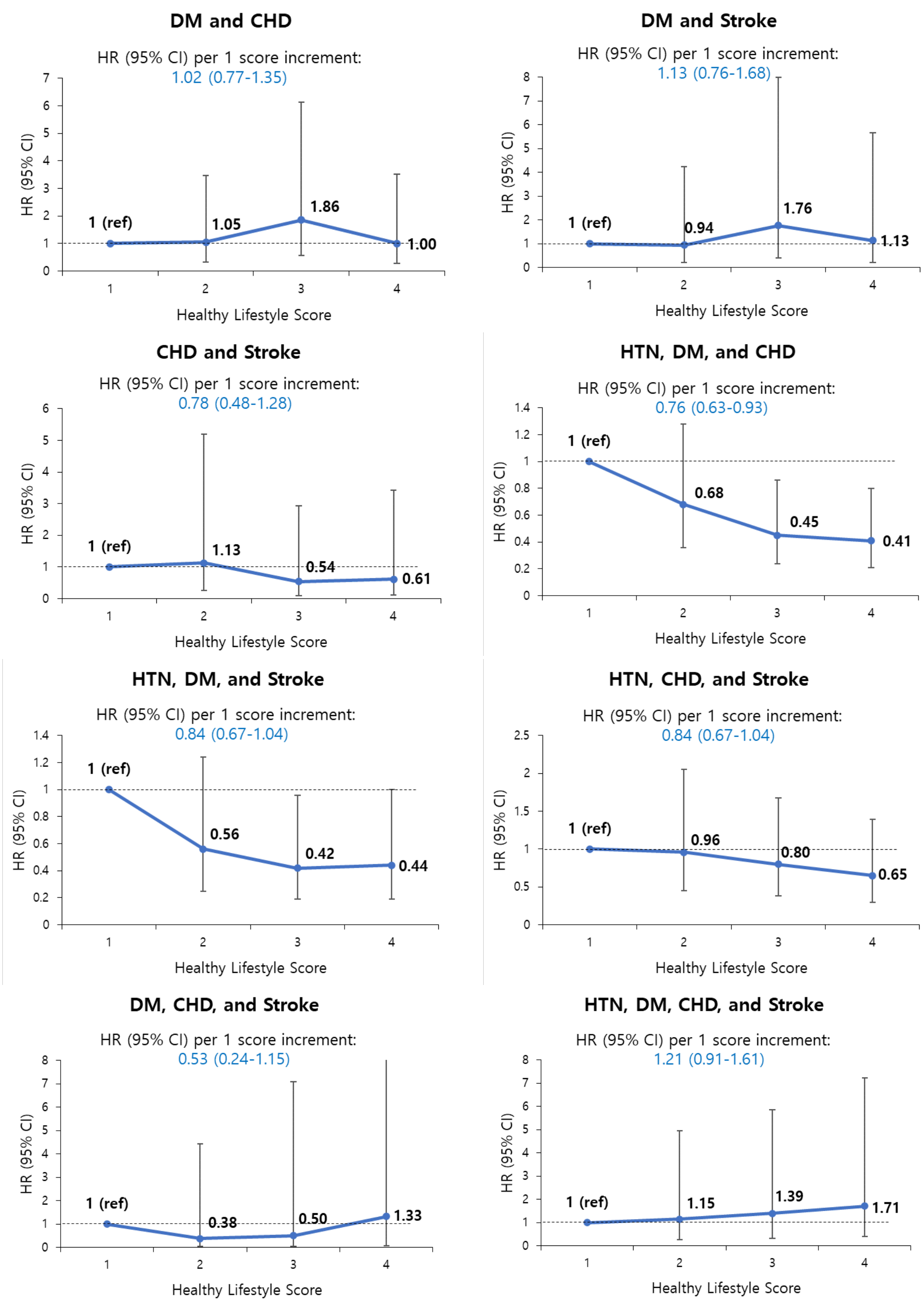
HLS [1]	All-cause death			CVD death		
	No CMD	1 CMD	2-4 CMDs	No CMD	1 CMD	2-4 CMDs
	HR (95% CI)[4]	HR (95% CI)[4]	HR (95% CI)[4]	HR (95% CI)[4]	HR (95% CI)[4]	HR (95% CI)[4]
0	1.00	1.00	1.00	1.00	1.00	1.00
1	0.92 (0.89-0.95)	0.97 (0.93-1.02)	0.92 (0.85-1.01)	0.91 (0.84-0.99)	0.97 (0.89-1.06)	0.86 (0.76-0.97)
2	0.71 (0.68-0.74)	0.74 (0.70-0.78)	0.74 (0.68-0.81)	0.67 (0.61-0.73)	0.76 (0.69-0.83)	0.74 (0.65-0.85)
3	0.61 (0.58-0.64)	0.67 (0.63-0.71)	0.70 (0.63-0.77)	0.54 (0.49-0.60)	0.71 (0.64-0.79)	0.71 (0.61-0.82)
	Premature all-cause death			Premature CVD death		
0	1.00	1.00	1.00	1.00	1.00	1.00
1	0.97 (0.92-1.02)	1.00 (0.92-1.08)	1.02 (0.88-1.19)	0.95 (0.84-1.06)	1.11 (0.97-1.28)	0.92 (0.74-1.15)
2	0.71 (0.67-0.76)	0.74 (0.68-0.81)	0.74 (0.62-0.87)	0.59 (0.52-0.67)	0.76 (0.65-0.90)	0.66 (0.51-0.85)
3	0.60 (0.56-0.65)	0.66 (0.59-0.74)	0.72 (0.59-0.87)	0.44 (0.37-0.51)	0.73 (0.60-0.89)	0.64 (0.47-0.86)

1. Number of healthy lifestyle conditions of cigarette smoking, alcohol drinking, and BMI
2. The ‘Premature death’ defined as ‘death at age < 70 years old’.
3. Number of diseases at baseline including hypertension, diabetes mellitus, ischemic heart disease, and stroke
4. Adjusted for age, gender, alcohol drinking, cigarette smoking, and BMI, excluding each analysis variable.

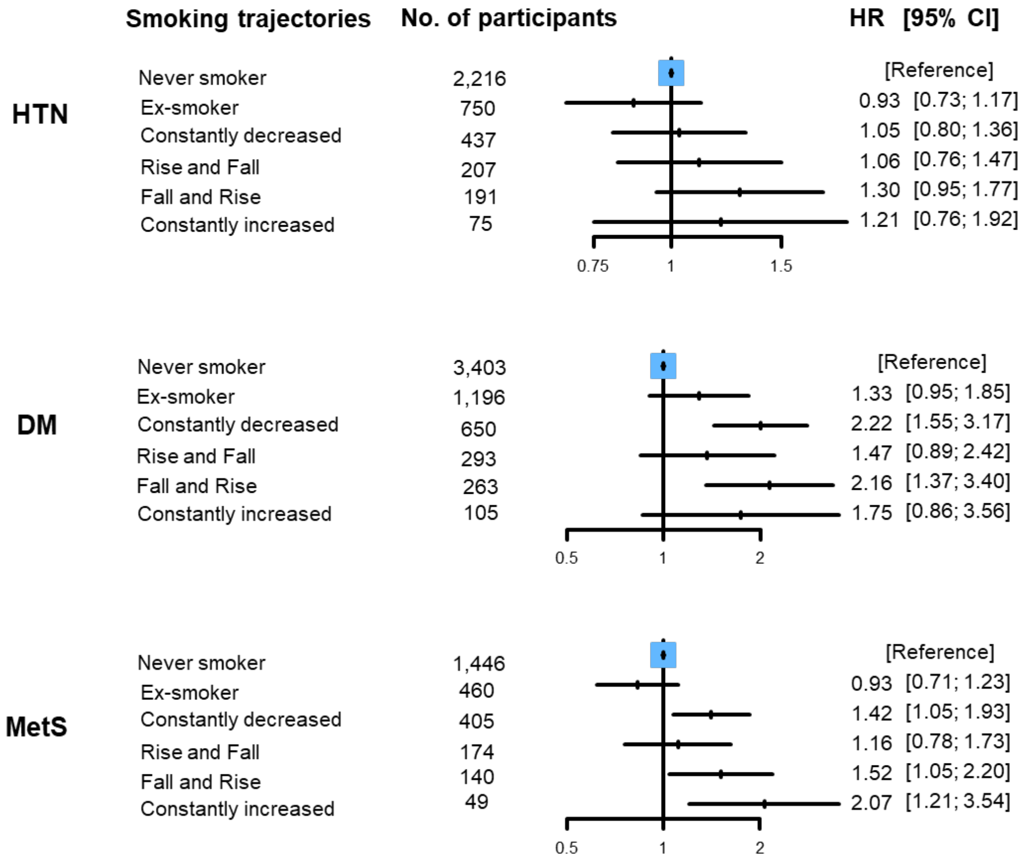
2.



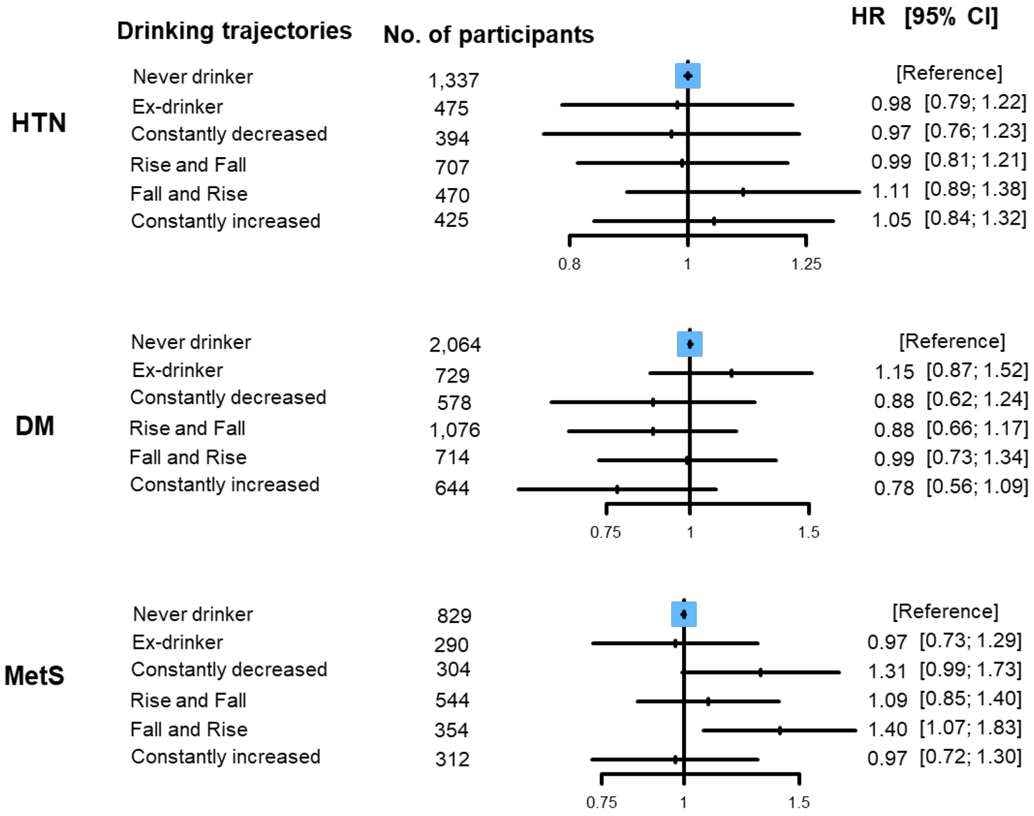
Appendix 14. Association of healthy lifestyle score with cardiovascular death according to disease status



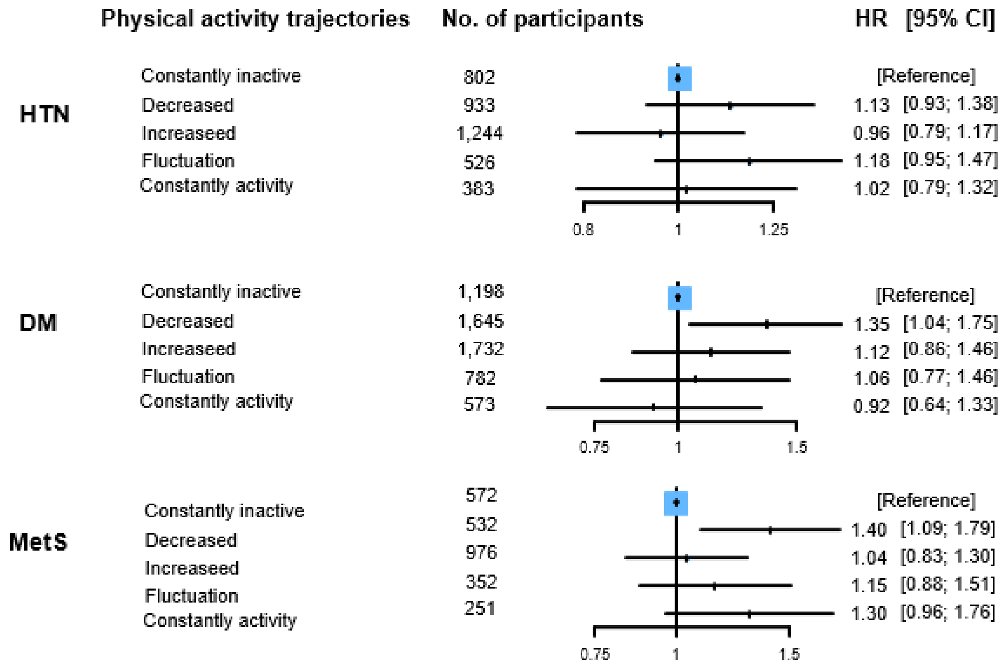
Appendix14(*Continued*). Association of healthy lifestyle score with cardiovascular death according to disease status



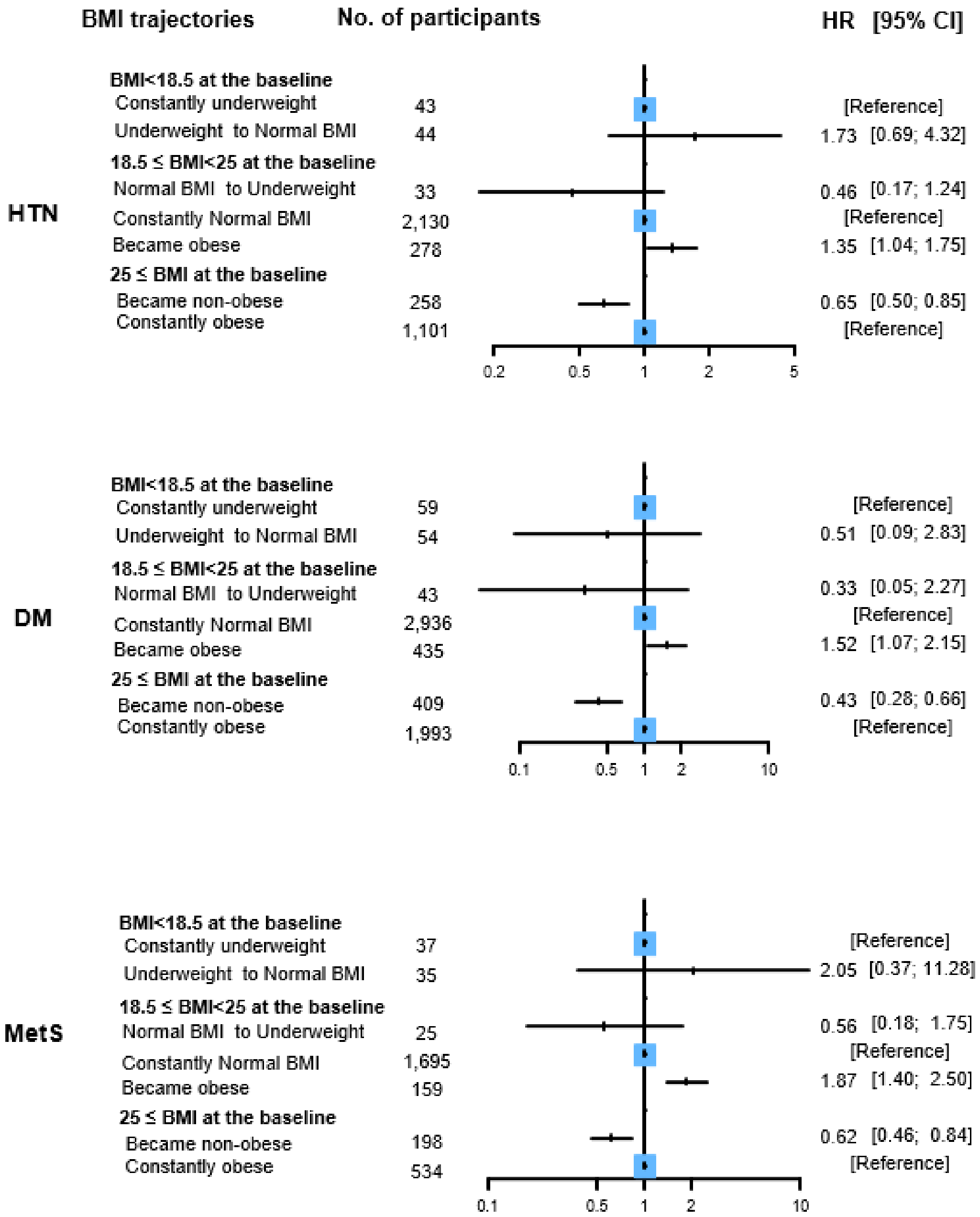
Appendix 15. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to the trajectory of cigarette smoking. (Hazard ratios are adjusted for age, sex, education, income, alcohol intake, physical activity, body mass index, total cholesterol level, and family history of cardiovascular disease)



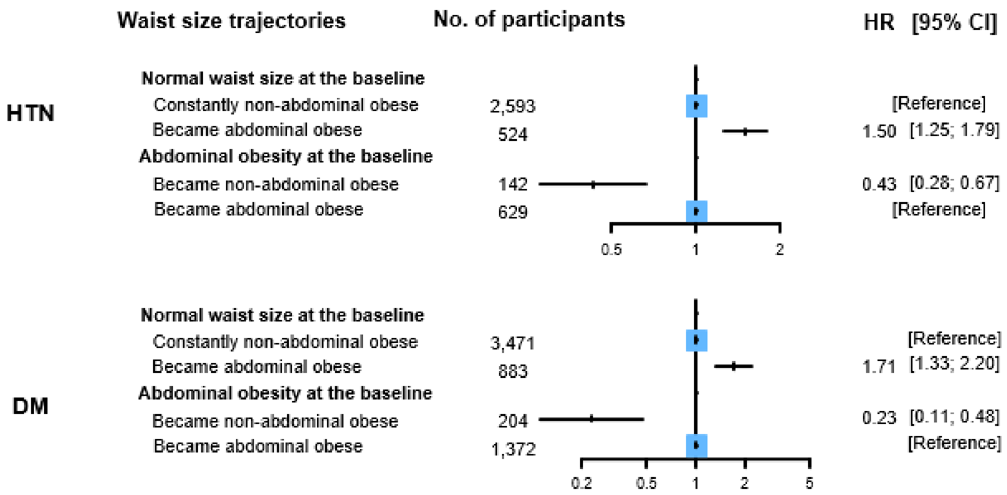
Appendix 16. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to the trajectory of alcohol consumption. (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, physical activity, body mass index, total cholesterol level, and family history of cardiovascular disease)



Appendix 17. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to the trajectory of regular physical activity. (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, alcohol intake, body mass index, total cholesterol level, and family history of cardiovascular disease)



Appendix 18. Adjusted hazard ratios for hypertension, diabetes mellitus, and metabolic syndrome according to the trajectory of body mass index. (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, alcohol intake, physical activity, total cholesterol level, and family history of cardiovascular disease)



Appendix 19. Adjusted hazard ratios for hypertension and diabetes mellitus according to the trajectory of waist size. (Hazard ratios are adjusted for age, sex, education, income, cigarette smoking, alcohol intake, physical activity, total cholesterol level, and family history of cardiovascular disease)

Appendix 20. Equation of biological age in men (Model 1)

$$[\text{Biological age in men}] = 65.1 + 0.6 \times \left(\frac{\text{Year} - 2008.7}{2.30}\right) - 0.4\left(\frac{\text{Height(cm)} - 168.7}{5.83}\right) - 2.6\left(\frac{[\text{Weight(kg)}] - 69.5}{9.32}\right) + 2.1\left(\frac{[\text{Waist size(cm)}] - 85.5}{7.48}\right) - 0.03\left(\frac{[\text{Hip size(cm)}] - 95.7}{5.68}\right) + 0.10\,[(\text{Dyslipidemia}); \text{Yes} = 1; \text{No} = 0\,] - 0.6\,[\,(\text{Allergy}); \text{Yes} = 1; \text{No} = 0\,] + 1.3\,[\,(\text{Tyroid disease}); \text{Yes} = 1; \text{No} = 0\,] + 0.8\,[\,(\text{Asthma}); \text{Yes} = 1; \text{No} = 0\,] - 5.6\,[\,(\text{Smoking status}); \text{None} = 0; \text{Past} = 1; \text{Current} = 0] - 10.1\,[(\text{Smoking status}); \text{None} = 0; \text{Past} = 0; \text{Current} = 1] + 4.7\left(\frac{[\text{Smoking duration(Year)}] - 17.3}{14.14}\right) - 0.6\left(\frac{[\text{Cigarette per day} - 12.3]}{11.21}\right) + 0.2\,[(\text{Drinking status}); \text{None} = 0; \text{Past} = 1; \text{Current} = 0] - 1.0\,[(\text{Drinking status}); \text{None} = 0; \text{Past} = 0; \text{Current} = 1] - 1.3\,[(\text{Secondhand smoking}); \text{Yes} = 1; \text{No} = 0] + 1.1\,[(\text{Regular exercise}); \text{Yes} = 1; \text{No} = 0;] - 1.9\,[(\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 1; \$2{,}000 - 4{,}000 = 0; \geq \$4{,}000 = 0;\,] - 4.0\,[\,(\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 0; \$2{,}000 - 4{,}000 = 1; \geq \$4{,}000 = 0;\,] - 4.2\,[\,(\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 0; \$2{,}000 - 4{,}000 = 0; \geq \$4{,}000 = 1;] - 2.2(\text{Education level}); < \text{Middle school or less} = 0; \text{High school} = 1; \text{College or more} = 0;] - 3.5(\text{Education level}); \text{Middle school or less} = 0; \text{High scholl} = 0; \text{College or more} = 1;] + 3.4\,[(\text{Marital status}); \text{Single} = 0; \text{Married} = 1;] - 4.8\,[(\text{Have occupation}; \text{Yes} = 1; \text{No} = 0;)]$$

Appendix 21. Equation of biological age in women (Model 1)

$$[\text{Biological age in women}] = 57.3+1.1\times(\frac{\text{Year}-20086}{2.62})-0.8(\frac{\text{Height(cm)}-156.3}{5.42})-1.1(\frac{[\text{Weight(kg)}]-57.9}{7.73})+2.0(\frac{[\text{Waist size(cm)}]-78.5}{8.28})-1.0(\frac{[\text{Hip size(cm)}]-93.5}{5.74})+2.6[(\text{Dyslipidemia}); \text{Yes}=1; \text{No}=0\ ]-0.7[\ [\text{Allergy}]; \text{Yes}=1; \text{No}=0\ ]+0.7[\ (\text{Tyroid disease}); \text{Yes}=1; \text{No}=0\ ]+0.7[\ (\text{Asthma}); \text{Yes}=1; \text{No}=0\ ]-2.1[\ (\text{Smoking status}); \text{None}=0; \text{Past}=1; \text{Current}=0]-3.4[(\text{Smoking status}); \text{None}=0; \text{Past}=0; \text{Current}=1]+0.5\left(\frac{[\text{Smoking duration(Year)}]-17.3}{14.14}\right)-0.2\left(\frac{[\text{Cigarette per day}-12.3]}{11.21}\right)-1.7[(\text{Drinking status}); \text{None}=0; \text{Past}=1; \text{Current}=0]-2.2[(\text{Drinking status}); \text{None}=0; \text{Past}=0; \text{Current}=1]-0.8[(\text{Secondhand smoking}); \text{Yes}=1; \text{No}=0]+1.0[(\text{Regular exercise}); \text{Yes}=1; \text{No}=0;]-2.1[(\text{Income level}); <\$1{,}000=0; \$1{,}000-2{,}000=1; \$2{,}000-4{,}000=0; \geq\$4{,}000=0;\ ]-3.7[\ (\text{Income level}); <\$1{,}000=0; \$1{,}000-2{,}000=0; \$2{,}000-4{,}000=1; \geq\$4{,}000=0;\ ]-3.9[\ (\text{Income level}); <\$1{,}000=0; \$1{,}000-2{,}000=0; \$2{,}000-4{,}000=0; \geq\$4{,}000=1;]-3.2[(\text{Education level}); <\text{Middle school or less}=0; \text{High school}=1; \text{College or more}=0;\ ]-4.5[\ (\text{Education level}); \text{Middle school or less}=0; \text{High scholl}=0; \text{College or more}=1;\ ]-1.8[\ (\text{Marital status}); \text{Single}=0; \text{Married}=1\ ;]-1.9[(\text{Have occupation}; \text{Yes}=1; \text{No}=0;)]+1.4[(\frac{[\text{Age at mecharche(Age)}-15.2]}{1.84})]+0.9[\ (\text{Have histroy of oral contraceptive}); \text{No}=0; \text{Yes}=1;\ ]+2.5[\ (\text{Have history of pregnant}); \text{No}=0; \text{Yes}=1;]$$

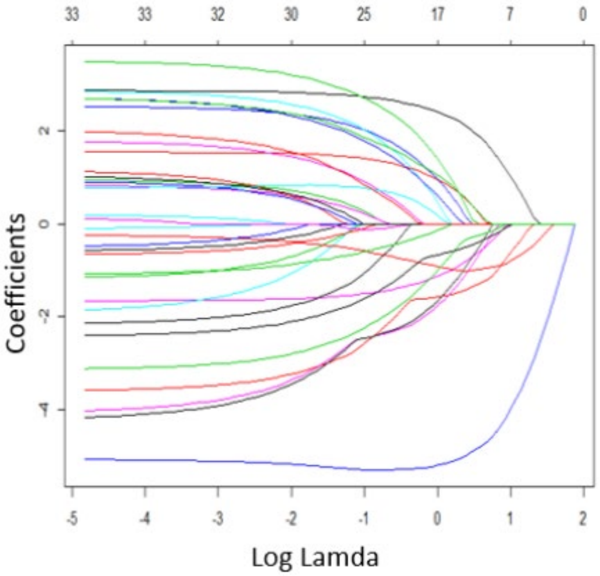
Appendix 22. Example of biological age

Consider a 50-year-old married man with non-smoking, current drinking habit, secondhand smoke, without regular exercise, has an occupation with income level over 200 - 400K/KW, graduated with college, height of 170 cm, weight of 68.5 kg, waist size of 85 cm, hip size of 96 cm and without any of disease history in 2005.

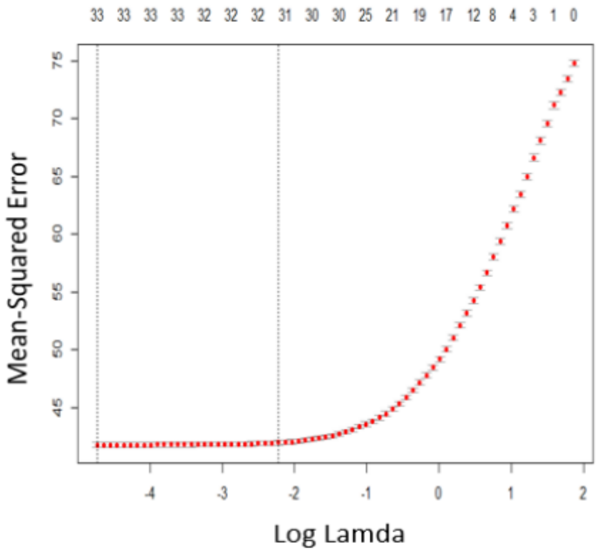
$$[\text{Biological age in men}] = 65.1 + 0.6 \times \left(\frac{\text{Year} - 2008.7}{2.30}\right) - 0.4\left(\frac{\text{Height(cm)} - 168.7}{5.83}\right) - 2.6\left(\frac{[\text{Weight(kg)}] - 69.5}{9.32}\right) + 2.1\left(\frac{[\text{Waist size(cm)}] - 85.5}{7.48}\right) - 0.03\left(\frac{[\text{Hip size(cm)}] - 95.7}{5.68}\right) + 0.10\,[(\text{Dyslipidemia}); \text{Yes} = 1; \text{No} = 0\ ] - 0.6\,[\ (\text{Allergy}); \text{Yes} = 1; \text{No} = 0\ ] + 1.3\,[\ (\text{Tyroid disease}); \text{Yes} = 1; \text{No} = 0\ ] + 0.8\,[\ (\text{Asthma}); \text{Yes} = 1; \text{No} = 0\ ] - 5.6\,[\ (\text{Smoking status}); \text{None} = 0; \text{Past} = 1; \text{Current} = 0] - 10.1\,[(\text{Smoking status}); \text{None} = 0; \text{Past} = 0; \text{Current} = 1] + 4.7\left(\frac{[\text{Smoking duration(Year)}] - 17.3}{14.14}\right) - 0.6\left(\frac{[\text{Cigarette per day} - 12.3]}{11.21}\right) + 0.2\,[(\text{Drinking status}); \text{None} = 0; \text{Past} = 1; \text{Current} = 0] - 1.0\,[(\text{Drinking status}); \text{None} = 0; \text{Past} = 0; \text{Current} = 1] - 1.3\,[(\text{Secondhand smoking}); \text{Yes} = 1; \text{No} = 0] + 1.1\,[(\text{Regular exercise}); \text{Yes} = 1; \text{No} = 0;] - 1.9\,[(\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 1; \$2{,}000 - 4{,}000 = 0; \geq \$4{,}000 = 0;\ ] - 4.0\,[\ (\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 0; \$2{,}000 - 4{,}000 = 1; \geq \$4{,}000 = 0;\ ] - 4.2\,[\ (\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 0; \$2{,}000 - 4{,}000 = 0; \geq \$4{,}000 = 1;] - 2.2(\text{Education level}); < \text{Middle school or less} = 0; \text{High school} = 1; \text{College or more} = 0;] - 3.5(\text{Education level}); \text{Middle school or less} = 0; \text{High scholl} = 0; \text{College or more} = 1;] + 3.4\,[(\text{Marital status}); \text{Single} = 0; \text{Married} = 1;] - 4.8\,[(\text{Have occupation}; \text{Yes} = 1; \text{No} = 0;)]$$

∴ **Biological age = 47.8**

A.



B.



Appendix 23. Coefficient paths and mean-squared error for the Elastic Net model

Appendix 24. Equation of biological age in men (Model 2)

$$[\text{Biological age in men}] = 60.9+0.5\times(\frac{\text{Year}-2008.7}{2.30})-0.5(\frac{\text{Height(cm)}-168.7}{5.83})-3.1(\frac{[\text{Weight(kg)}]-69.5}{9.32})+2.5(\frac{[\text{Waist size(cm)}]-85.5}{7.48})+0.02(\frac{[\text{Hip size(cm)}]-95.7}{5.68})+0.11[(\text{Dyslipidemia}); \text{Yes}=1; \text{No}=0\ ]-0.5[\ (\text{Allergy}); \text{Yes}=1; \text{No}=0\ ]+1.6[\ (\text{Tyroid disease}); \text{Yes}=1; \text{No}=0]+0.8[(\text{Asthma}); \text{Yes}=1; \text{No}=0]+0.8(\frac{[(\text{Pack-year})-15.5]}{16.6})+0.2[(\text{Drinking status}); \text{None}=0; \text{Past}=1; \text{Current}=0]-1.4[(\text{Drinking status}); \text{None}=0; \text{Past}=0; \text{Current}=1]-1.8[(\text{Secondhand smoking}); \text{Yes}=1; \text{No}=0]+1.5[(\text{Regular exercise}); \text{Yes}=1; \text{No}=0;]-2.3[(\text{Income level}); <\$1{,}000=0; \$1{,}000-2{,}000=1; \$2{,}000-4{,}000=0; \geq\$4{,}000=0;]-4.8[(\text{Income level}); <\$1{,}000=0; \$1{,}000-2{,}000=0; \$2{,}000-4{,}000=1; \geq\$4{,}000=0;]-4.9[(\text{Income level}); <\$1{,}000=0; \$1{,}000-2{,}000=0; \$2{,}000-4{,}000=0; \geq\$4{,}000=1;]-2.7(\text{Education level}); <\text{Middle school or less}=0; \text{High school}=1; \text{College or more}=0;]-4.2(\text{Education level}); \text{Middle school or less}=0; \text{High scholl}=0; \text{College or more}=1;]+4.4[(\text{Marital status}); \text{Single}=0; \text{Married}=1;]-5.8[(\text{Have occupation}; \text{Yes}=1; \text{No}=0;)]$$

Appendix 25. Equation of biological age in women (Model 2)

$$[\text{Biological age in women}] = 57.2 + 1.1 \times \left(\frac{\text{Year} - 2008.6}{2.62}\right) - 0.8\left(\frac{\text{Height(cm)} - 156.3}{5.42}\right) - 1.1\left(\frac{[\text{Weight(kg)}] - 57.9}{7.73}\right) + 2.0\left(\frac{[\text{Waist size(cm)}] - 78.5}{8.28}\right) - 0.2\left(\frac{[\text{Hip size(cm)}] - 93.5}{5.74}\right) + 2.6\,[(\text{Dyslipidemia}); \text{Yes} = 1; \text{No} = 0] - 0.7\,[[\text{Allergy}]; \text{Yes} = 1; \text{No} = 0] + 0.7\,[(\text{Tyroid disease}); \text{Yes} = 1; \text{No} = 0] + 0.7\,[(\text{Asthma}); \text{Yes} = 1; \text{No} = 0] - 0.02\left(\frac{[(\text{Pack-year}) - 0.3]}{2.11}\right) - 1.7\,[(\text{Drinking status}); \text{None} = 0; \text{Past} = 1; \text{Current} = 0] - 2.3\,[(\text{Drinking status}); \text{None} = 0; \text{Past} = 0; \text{Current} = 1] - 0.9\,[(\text{Secondhand smoking}); \text{Yes} = 1; \text{No} = 0] + 1.0\,[(\text{Regular exercise}); \text{Yes} = 1; \text{No} = 0;] - 2.2\,[(\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 1; \$2{,}000 - 4{,}000 = 0; \geq \$4{,}000 = 0;] - 3.6\,[(\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 0; \$2{,}000 - 4{,}000 = 1; \geq \$4{,}000 = 0;] - 3.9\,[(\text{Income level}); < \$1{,}000 = 0; \$1{,}000 - 2{,}000 = 0; \$2{,}000 - 4{,}000 = 0; \geq \$4{,}000 = 1;] - 3.2\,[(\text{Education level}); < \text{Middle school or less} = 0; \text{High school} = 1; \text{College or more} = 0;] - 4.5\,[(\text{Education level}); \text{Middle school or less} = 0; \text{High scholl} = 0; \text{College or more} = 1;] - 1.8\,[(\text{Marital status}); \text{Single} = 0; \text{Married} = 1;] - 1.9\,[(\text{Have occupation}; \text{Yes} = 1; \text{No} = 0;)] + 1.4\,\left[\left(\frac{[\text{Age at mecharche(Age)} - 15.2]}{1.84}\right)\right] + 0.9\,[(\text{Have histroy of oral contraceptive}); \text{No} = 0; \text{Yes} = 1;] + 2.4\,[(\text{Have history of pregnant}); \text{No} = 0; \text{Yes} = 1;]$$

Appendix 26. Multicollinearity test for independent variables measured by the variance inflation factor for the variables included in the hypertension prediction model

Variables	**Variance Inflation Factor**
Sex	2.941
Age	1.415
Education level	1.391
Income level	1.345
BMI	2.547
Waist circumstance	3.012
Smoking	1.946
Alcohol consumption	1.223
Physical activity	1.035
Total calorie intake	1.038
SBP, mmHg	1.647
DBP, mmHg	1.608
Diabetes mellitus	1.051
Total cholesterol	1.096
HDL-cholesterol	1.180
Triglyceride	1.133
Albumin/creatinine ratio	1.530
Cardiovascular disease	1.034
Family history of CVD	1.012

Appendix 27. Multicollinearity test for independent variables measured by the variance inflation factor for the variables included in the diabetes mellitus prediction model

Variables	**Variance Inflation Factor**
Sex	2.527
Age	1.405
Education level	1.420
Income level	1.390
BMI	2.766
Waist circumstance	3.294
Smoking	1.913
Alcohol consumption	1.322
Physical activity	1.041
Total calorie intake	1.044
hypertension	1.152
Total cholesterol	1.083
HDL-cholesterol	1.184
Triglyceride	1.159
Cardiovascular disease	1.031
Family history of CVD	1.012

Appendix 28. Multicollinearity test for independent variables measured by the variance inflation factor for the variables included in the comorbidity of hypertension and diabetes mellitus prediction model

Variables	Varian Inflation Factor
Sex	2.805
Age	1.390
Education level	1.368
Income level	1.309
BMI	2.427
Waist circumstance	2.837
Smoking	1.917
Alcohol consumption	1.189
Physical activity	1.033
SBP, mmHg	1.644
DBP, mmHg	1.613
Total calorie intake	1.037
Total cholesterol	1.101
HDL-cholesterol	1.171
Triglyceride	1.121
Albumin/creatinine ratio	1.487
Cardiovascular disease	1.027
Family history of CVD	1.012

요약 (국문초록)

연구 배경: 인구의 고령화와 서구형 생활양식으로 인해 대사 질환 동시 이환 (고혈압, 당뇨병, 및 고지혈증 등을 포함한 두가지 이상의 대사 질환을 가진 것으로 정의)의 유병률이 증가하고 있다. 이러한 대사성 질환은 심혈관계 질환의 위험 증가와 연관된다. 2016년 Global Burden of Disease에 따르면, 심혈관계 질환에 의한 사망은 21세기 주요 사망 원인이며, 우리나라에서는 암에 이어 두번째로 높은 사망원인을 차지한다. 세계보건기구 (The World Health Organization)에서는 음주, 흡연, 비만, 신체 활동, 건강한 식습관을 심혈관계 질환의 예방 가능한 요인으로 지정한 바 있다. 이에 대사 질환 동시 이환에 대한 연구가 필요하다. 따라서, 이 연구의 목적은 1) 한국에서의 대사성 질환과 동시 이환의 유병률을 추정하고; 2) 대사 동시 이환 심혈관계 가족력과 심혈관계 발생 위험을 평가하고, 3)대사 동시 이환에 따른 심혈관계 사망에 대해 생활습관 요인 미치는 영향을 평가하고; 4) 생활 습관 변화와 대사 증후군의 연관성을 확인하고; 5) 대사 동시 이환에 대한 기계학습을 기반으로 한 건강 연령 및 질병 위험 예측 모형을 개발하는 것이다.

연구 방법: 본 연구는 한국인유전체역학조사사업 (KoGES)의 도시기반 (Health examinee-Gem Study, HEXA), 농촌기반 (Cardiovascular disease association study, CAVAS), 지역사회기반 (Ansan and Ansung Study, 2001-2014)를 주로 사용하였고, 추가로 미국 국민건강영양조사 (US National Health and Nutrition Examination Survey, NHANES 2003-2014), 한국국민건강영양조사 (Korea NHANES, KNHANES 2007-2014), 아시아 코호트 연구 (Asia Cohort Consortium)를 사용하였다. 통계방법으로는, 세계보건기구의 세계표준인구를 이용한 직접 표준화 방법을 이용해 대사성 질환의 연령표준화 유병률을 산출하였다. 연구 대상자의 일반적인 특성은 연속형 변수의 경우 Student's t-test, 범주형 변수의 경우 Chi-squared test를 시행하여

비교하였다. 콕스 비례 위험 회귀 분석과 로지스틱 회귀 분석을 수행하여 hazard ratios (HRs), odds ratio (ORs), 95% confidence interval을 추정하였다. 위험 예측 모형의 경우, training set (전체 대상자의 70%)에서 콕스 비례 회귀 분석, random survival forest 기반 모형을 각각 구축하고, test set (전체 대상자의 30%)에서 concordance index (c-index)를 이용해 각 모형의 성능을 평가하였다. 건강 연령 예측 모형의 경우, 10-fold validation을 사용한 elastic net 방법을 이용해 모형을 구축하였다.

연구 결과: 한국과 미국의 대사성 질환과 동시 이환을 비교한 결과, 한국이 미국보다 대사 동시 이환의 유병률이 낮았다. 한국과 미국에서 가장 흔한 대사 질환 조합은 고혈압과 비만이었다. 한국 인구 중 농촌 지역에 거주하는 인구는 도시 지역에 거주하는 인구보다 대사 동시 이환 유병률이 더 높은 것으로 나타났다.

대사 동시 이환, 심혈관계 질환 가족력, 그리고 심근경색과 뇌졸중을 포함한 심혈관계 질환의 위험 연구 결과는 다음과 같다. 고혈압, 당뇨병, 고지혈증이 있고, 심혈관계 가족력이 있는 대상자는 심혈관계 질환 가족력과 질병이 없는 대상자에 비해 유의하게 심혈관계 질환 (HR 2.88, 95% CI: 1.96-4.24), 심근경색 (HR 3.30, 95% CI: 2.06-5.29), 뇌졸중 (HR 2.52, 95% CI: 1.33-4.79) 위험이 증가하는 것을 확인했다

심혈관대사 질환 동시 이환을 가진 대상자에서 생활 습관 요인이 심혈관계 질환 관련 사망에 미치는 영향 연구에서는, '비흡연', '금주', '체질량지수 18.5-27.4kg/m^2' 를 건강 상태로 정의하여 건강한 생활 습관 점수를 산출했다. 생활 습관 요인 중 금연은 심혈관계 질환 사망 위험 감소와 가장 강한 연관성을 보였다. 고혈압, 당뇨병, 관상동맥질환이 있는 대상자에서는 건강한 생활 습관 점수가 1씩 증가할 때마다 심혈관계 사망위험이 24% (HR 0.76, 95% CI: 0.63-0.93)씩 감소했다. 2개 이상의 심혈관계 대사질환이 있는 대상의 경우, 건강한 생활 습관 요인은 3가지 모두 있는 경우 심혈관계 질환 사망 (HR 0.51, 95% CI: 0.42-0.61)과 심혈관계 질환으로 인한 조기 사망위험(HR 0.38, 95% CI: 0.27-0.54)

의 감소에 유의한 영향이 있었다.

지역사회기반 연구자료를 이용한 반복 측정된 생활 습관 요인의 변화에 따른 대사 증후군 위험 연구에서는, 하루 흡연 개피수의 증가 (HR 1.49, 95% CI: 1.09−2.03), 음주량의 light/moderate에서 heavy로 증가는 (HR 1.42, 95% CI: 1.10−1.84) 대사 증후군의 발생 위험의 증가와 유의한 연관성을 보였다. 새롭게 비만 된 대상자는 꾸준히 적정 체중을 유지하는 대상자에 비해 대사성 증후군 (HR 1.88, 95% CI: 1.44−2.45)의 발생 위험의 증가와 유의한 관계를 보였다.

보다 정밀한 개인 맞춤 건강 상태 예측 및 개선을 위해 기계 학습 기반 질병 예측 모형을 개발과 대사 동시 이환에 대한 예측 변수로서의 건강 연령을 개발한 연구에 따르면, 실제 연령에 비해 젊은 건강 연령을 가진 경우, 당뇨병 (HR = 0.63, 95% CI: 0.55–0.72), 고혈압 (HR = 0.74, 95% CI: 0.68–0.81), 당뇨병과 고혈압 동시 이환 (HR = 0.65, 95% CI: 0.47–0.91) 위험도가 낮은 것으로 나타났다. 기계학습기반 예측 모형 연구 결과, 기계 학습 기반의 고혈압과 당뇨병 동시 이환 모형은 높은 통계적 질병 예측력을 보이는 것으로 나타났다.

연구 결론: 본 연구는 한국 인구집단에서 심혈관계 질환 발생 및 사망의 위험을 줄이기 위해 대사 동시 이환에 대한 연구에 대한 필요성을 강조한다. 본 연구에서는 동시 이환을 가진 대상자 중 특히 심혈관계 질환 가족력이 있는 경우에 심혈관계 질환의 발생 위험이 증가하는 것을 확인하였다. 또한 심혈관계 대사 질환 동시 이환을 가진 대상자라도, 금연, 금주, 표준 체질량 지수 유지와 같은 건강한 생활 습관은 심혈관계 질환으로 인한 사망과 조기 사망 위험 감소와 연관성이 있었다. 또한, 건강한 생활습관으로의 변화를 통해 대사 증후군의 위험을 줄이는 데 도움이 되는 것을 확인하였다. 이러한 요인들을 기반으로 기계학습을 이용하여 구축된 질병 예측 모형과 건강연령은 우리나라에서의 대사 질환 동시 이환에 대한 고위험군을 파악하고 이를 미리 예방함으로써, 건강증진을 통해 질병 부담을 줄이는 효과적인 도구로 활용될 수 있을 것으로 기대된다.

Keywords: 대사질환 동시 이환, 생활습관, 심혈관계 질환, 건강 연령, 질병예측모형

Student number: 2016-21993

Acknowledgment

This study was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), and National Genome Research Institute, Korea Center for Disease Control and Prevention, and by a grant from Seoul National University Hospital (2022). This research was funded by the Ministry of Health & Welfare, Republic of Korea (grant No. HI16C1127).