Major risk factors for heart failure and pulmonary heart disease in 500,000 adults in the China Kadoorie Biobank

Valirie Agbor Ndip

Linacre College, University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2024

Abstract

Major risk factors for heart failure and pulmonary heart disease in 500,000 adults in the China Kadoorie Biobank

Valirie Agbor Ndip, Linacre College, University of Oxford

A thesis submitted to Oxford Population Health for the degree of Doctor of Philosophy

Trinity Term 2024

Background: Observational and Mendelian randomisation (MR) studies conducted primarily in Western countries suggest that higher levels of systolic blood pressure (SBP), body mass index (BMI) and waist-to-hip ratio (WHR) are associated with higher risks of HF. However, uncertainties persist about the shape and strength of these associations, especially in non-Western populations. Moreover, studies on the associations of SBP and adiposity with PHD are limited.

Methods: This thesis used data from the China Kadoorie Biobank study, which is a prospective cohort study of 0.5 million middle-aged Chinese adults recruited from 10 geographically diverse areas (5 rural, 5 urban) during 2004–2008 and followed up through record linkage to morbidity and mortality registers. Cox regression yielded adjusted hazard ratios (HRs) for HF and PHD with usual SBP, BMI and WHR. The likelihood ratio test was used to quantify the effect of adjustments for confounding and assess for effect modification and mediation. Mendelian randomisation (MR) provided causal evidence of the most promising associations in the observational analyses. Two-tailed p-values less than 0.05 were considered statistically significant. Data was analysed in R version 4.2.2.

Results: The overall mean (SD) SBP, BMI and WHR at baseline were 131 (21) mmHg, 23.6 (3.4) kg/m^2 and 0.88 (0.07), respectively. In observational analyses, SBP had a

positive and log-linear association with HF (HR per 10 mmHg: 1.13, 95% CI 1.11–1.15; n=8,223) but was not associated with PHD. The MR analysis also showed a positive association of genetically-predicted SBP with HF (HR per 10 mmHg: 1.18, 1.01–1.39; n=1,160). In observational analyses, BMI and WHR had a J-shaped association with HF (HR per 1 SD for BMI >22 kg/m^2: 1.21, 1.18–1.25; for WHR >0.85: 1.17, 1.13–1.21; n=1,160). The associations of BMI and WHR with HF persisted after mutually adjusting for each other. By contrast, BMI and WHR had a reverse J-shaped association with PHD (HR per 1 SD for BMI <26 kg/m^2: 0.61, 0.59–0.63; n=1,104) that attenuated to the null after excluding PHD cases recorded within the first two years of follow-up and prior medical conditions at baseline. In MR, genetically-predicted BMI was positively associated with HF (HR per 1 SD: 1.39, 1.11–1.75) but showed no association with PHD. By contrast, genetically-predicted WHR was not associated with HF or PHD in MR.

Conclusions: In this Chinese population, SBP and BMI were positively associated with HF but not PHD. However, WHR was not associated with HF or PHD. These findings contribute to the limited knowledge of the associations of SBP and adiposity with HF in non-Western populations and provide preliminary evidence of the associations of SBP and adiposity with PHD.

Acknowledgements

I am grateful to the participants of the China Kadoorie Biobank (CKB), without whom this study would not have been possible. I appreciate the Commonwealth Scholarship Commission in the United Kingdom for funding my studies.

I appreciate my DPhil supervisors, Associate Professor Derrick Bennett, Dr Yiping Chen and Professor Robert Clarke, for their invaluable time, patience, guidance and support during this project. I will apply your teachings in my future endeavours.

I am grateful to the wider CKB collaborative group for the encouragement and support. I acknowledge Professor Zhengming Chen, Associate Professor Christiana Kartsonaki, Dr Neil Wright, Dr Andri Iona, Associate Professor Huaidong Du, Dr Iona Millwood, Dr Ahmed Edris, Dr Alfred Pozarickij and Dr Sam Morris and special thanks to the Admin, IT and Data groups.

I acknowledge Bowen Liu for the insightful discussions during the initial phase of my Mendelian randomisation (MR) analysis. I acknowledge Dr Zammy Fairhurst-Hunter for developing the genetic risk scores for adiposity traits used in my MR analysis. I am grateful to Marie-Alice Frippat for her unwavering support with DPhil administrative matters. I thank my friends Besong Micheal, Marwa Elbasheer, Yu-jie Chiou, Abednego Masai, Uchenna Gwacham-Anisiobi, Paul Lokubal and colleagues (Ayodipupo Oguntade, Reneé Pereyra-Elías, Olaa Mohamed-Ahmed, Danial Qureshi and David Shaw) at Oxford Population Health for their support and companionship. Special thanks to my mentors, especially Professor Dora Mbanya, Dr Chobufo Ditah and Dr Jean Jacques Noubiap, for encouraging me to reach for the stars.

To my late parents, I am sure you are proud to see how far I have come. Thank you for teaching me about hard work, passion and dedication. I am indebted to my uncles (Mr. Tezoh Lawrence and Maître Zemenjuoh Etienne), aunts (Mrs. Tezoh Regina and Mrs Zemenjuoh Francisca) and siblings for their guidance and support.

Finally, my deepest appreciation goes to my lovely wife, Larissa, to whom I dedicate this thesis, for her love, patience, and sacrifice throughout this project.

Statement of contribution to work

This thesis is my own work unless stated otherwise. My supervisors conceived the idea to investigate the risk factors for heart failure and pulmonary heart disease. I established the research objectives for this thesis with input from my supervisors. In addition, I developed the data analysis plans, performed all statistical analyses, produced the tables and figures, and interpreted the findings with feedback from my supervisors, Derrick Bennett, Yiping Chen, and Robert Clarke. I wrote up all sections of this thesis. Chapter 4 has been submitted for publication and is under review. As the first author, I wrote the initial draft of the manuscript and revised the manuscript based on the suggestions from all the other co-authors. I have received funding from the Commonwealth Scholarship Commission in the United Kingdom to undertake this DPhil project.

Table of Contents

List of abbreviations

APCSC	Asia-Pacific Cohort Studies Collaboration
ARIC	Atherosclerosis Risk in Communities
BMI	Body mass index
BNP	Brain Natriuretic Peptide
CALIBER	CArdiovascular research using LInked Bespoke studies and Electronic health Records
CFR	Case-fatality rate
CHD	Coronary heart disease
CI	Confidence interval
CKB	China Kadoorie Biobank
CMR	Cumulative mortality risk
COPD	Chronic obstructive pulmonary disease
CVD	Cardiovascular disease
DXA	Dual-energy x-ray absorption
FAR	Floating absolute risk
GIANT	Genetic Investigation of ANthropometric Traits
GWAS	Genome-wide association study
HERMES	HEart failure Molecular Epidemiology for Therapeutic targetS
HF	Heart failure
HR	Hazard ratio
ICBP	International Consortium of Blood Pressure
InSIDE	Instrument Strength Independent of Direct Effect
IPD	Individual participants data
IV	Instrumental variable
IVW	Inverse-variance-weighted
LD	Linkage disequilibrium
LMICs	Low and middle-income countries
LR	Likelihood ratio
MAF	Minor allele frequency
MET	Metabolic equivalent of tasks
MR-cML	Mendelian randomisation using constraint Maximum-likelihood
mre	multiplicative random-effects
MR-RAPS	Mendelian randomisation robust adjusted profile score
OR	Odds ratio
PHD	Pulmonary heart disease
RCT	Randomised controlled trial
RDB	Regression dilution bias
RDR	Regression dilution ratio
RHD	Rheumatic heart disease
RHR	Resting heart rate
SBP	Systolic blood pressure
SD	Standard deviation
SDG	Sustainable development goal
SNP	Single nucleotide polymorphism
TIA	Transient ischaemic attack
TPA	Total physical activity
WHR	Waist-to-hip ratio

List of tables

List of figures

1. Introduction

1.1. Heart failure

1.1.1. Definition and classification

Heart failure (HF) is a complex clinical syndrome resulting from structural and functional impairment of ventricular filling or ejection fraction[1,2]. The European Society of Cardiology and the Chinese Society of Cardiology define HF as a clinical syndrome characterised by symptoms (such as shortness of breath, reduced exercise capacity, orthopnoea, bilateral ankle swelling and fatigue) and signs (such as over 2kg of unintentional weight gain per week, bilateral pedal oedema and lung crepitations), with corresponding elevated blood levels of natriuretic peptides and radiological or imaging evidence of systemic or pulmonary venous congestion[1,2].

HF is a heterogeneous disease that is caused by any disorder affecting the structure or function of the heart, impairing the ability of the heart to pump efficiently[1]. HF chiefly results from conditions causing a dysfunction of the myocardium (heart muscles), such as coronary artery disease, hypertension, and chronic lung disease[1]. It is also caused by abnormal heart rhythms (e.g., atrial fibrillation), systemic diseases like haemochromatosis or disorders of the heart valves (e.g., rheumatic heart disease [RHD]) or pericardium (e.g., chronic pericarditis)[1].

HF diagnosis is based on a combination of characteristic signs and symptoms and relevant clinical evidence of impaired ventricular function[1]. Depending on the affected side of the heart, HF is typically classified into left and right HF[3,4]. Pulmonary heart disease (PHD) is a specific type of right HF resulting from chronic lung disease. In order to compare the epidemiology of left and right HF, PHD will be used as a proxy for right HF in this thesis.

1.1.2. Pathophysiology of left and right heart failure

1.1.2.1. Structure of the heart

Appreciation of the normal structure and function of the heart is essential to understanding how risk factors lead to disease. The heart is a hollow muscular organ about the size of a clenched fist. It is separated into the left and right sides by the septum, which prevents the mixing of venous blood on the right with arterial blood on the left side of the heart (**Figure 1.1**)[5,6]. The two sides of the heart function as separate pumps that contract simultaneously to expel blood out of the heart[6].

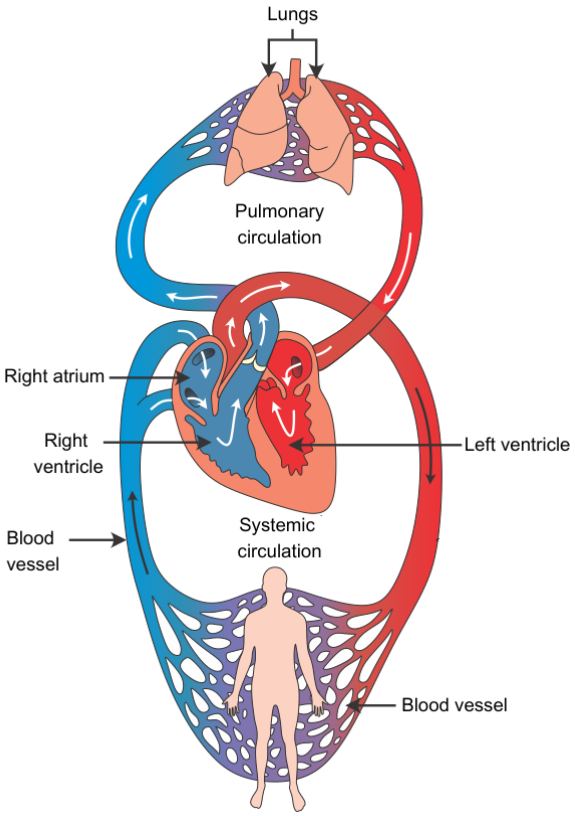
There are two chambers on each side of the heart: the right atrium and ventricle on the right, and the left atrium and ventricle on the left (**Figure 1.1**)[6]. The atria are thin-walled chambers that receive venous blood into the heart[6]. Heart function is chiefly determined by the function of the ventricles, which have thick muscular walls and are responsible for pumping blood out of the heart[5]. The energy needed for cyclic ventricular contraction is obtained from nutrients and oxygen supplied by the coronary arteries[5,6].

1.1.2.2. Functions of the heart

The heart undergoes cycles of contraction and relaxation to perform its function, with filling of the cardiac chambers during relaxation and the expulsion of blood from the heart during contraction. Deoxygenated blood from systemic veins enters the right ventricle, which contracts, generating sufficient force to pump the blood against the pressure and resistance of the pulmonary arteries (known as pulmonary resistance) to the lungs for oxygenation (**Figure 1.1**)[5,6]. The oxygenated blood is then returned to the left ventricle, which contains more myocardial muscle mass than the right ventricle. The left ventricle then contracts to generate sufficient force to pump the arterial blood against systemic resistance to the remainder of the body, where the oxygen in the blood is used

for metabolism[5]. The deoxygenated blood is then returned to the right heart through the systemic veins to repeat the cycle.

Figure 1.1. Anatomy of the cardiovascular system.



The blue colour represents deoxygenated blood, while the red represents oxygenated blood. The figure was annotated with permission from the original author (https://www.toppr.com/ask/question/explain-circulatory-system-through-a-diagram/)

1.1.2.3. Pathophysiology of left and right heart failure

Heart failure has an insidious onset that typically occurs over several years. The pathogenesis of HF begins with an initial insult to the structural and functional integrity of the heart following sufficient exposure to risk factors, causing a structural or functional impairment of the heart, leading to ventricular dysfunction and reduction in cardiac output[5,7]. In most cases, this ventricular dysfunction is asymptomatic for several years as neurohormonal mechanisms are triggered to compensate for the reduction in cardiac

output[3,7,8]. Over time, however, these compensatory mechanisms result in a deterioration of ventricular dysfunction, leading to overt signs and symptoms of HF, which progresses to worsening HF with recurrent hospitalisations and death[7,9]. The mechanisms of left and right HF are very similar and will be described together. Any differences in the mechanisms of left and right HF will be highlighted accordingly.

Ventricular function is chiefly determined by preload, contractility and afterload. These processes will be described individually but can occur either in isolation or simultaneously. Preload refers to the extent to which the myocardial muscles stretch when the ventricles are filled with blood[5,8]. Elevated preload results from an increase in the quantity of blood entering the heart, such as during valvular regurgitation, whereby the heart valves fail to regulate the amount of blood filling the heart ventricles[5,8,10,11]. In the normal state, the myocardial muscle fibres are elastic and contract proportional to the degree of stretching; the ventricles are compliant and accommodate increases in preload[8,10,11]. Progressive ventricular dilation, however, results in worsening ventricular function, leading to heart failure[8,10,11].

Contractility refers to the ability of the ventricular myocardial muscles to contract when filled with blood[5,9]. Factors that affect the structure (e.g., ischaemic heart disease) or function (e.g., atrial fibrillation) of myocardial fibres will cause HF[1,7,9,11].

Afterload is the stress on the myocardial muscles, resulting from resistance to blood flow out of the heart that occurs during ventricular contraction[5,9]. In response to persistent elevations in afterload, the ventricular myocardium hypertrophies by ventricular remodelling[5,9–11]. Ventricular remodelling enables the ventricle to adapt to elevations in the afterload by generating sufficient contractile force to overcome this resistance and maintain the stroke volume (the amount of blood pumped out of the heart after a single contraction).

Ventricular hypertrophy is a thickening of the walls of the left heart chamber and is a key mechanism of HF resulting from non-ischaemic causes. The hypertrophied ventricular walls are eventually replaced by inelastic fibrotic tissues, leading to a deterioration in ventricular function and HF[5,9,10,12]. Ventricular hypertrophy also results in poor oxygen supply from the coronary arteries, leading to ischaemic heart disease, poor ventricular contractility and HF[5,9,10,12]. In the right ventricle, resistance from the pulmonary circulation chiefly determines the afterload. Therefore, diseases that increase pulmonary resistance, including lung diseases, left HF, obesity or pulmonary hypertension, can also result in right HF[4,9,13]. In contrast, the afterload in the left ventricle is primarily determined by the systemic circulation. Sustained elevated resistance in the systemic arteries (e.g., due to systemic hypertension) increases the risk of left HF[5,9,11].

Other mechanisms involved in the development of HF include ischaemia and inflammation[9]. Diabetes and obesity promote inflammation of small blood vessels that supply the myocardium, leading to ischaemia and worsening ventricular function[9,10,12].

1.1.3. Diagnosis of heart failure

The diagnosis of left and right HF is based on the presence of characteristic symptoms and relevant clinical signs, and the preliminary diagnostic step involves the identification of the signs and symptoms of HF[1,2]. However, the most common symptoms include breathlessness, fatigue, ankle oedema and light-headedness. Less common symptoms include persistent cough, wheeze, palpitations and loss of appetite. However, these symptoms for HF and PHD overlap with the presentation of other medical conditions, including obesity and chronic pulmonary disease (COPD), reducing their accuracy of diagnosis[1,2]. Moreover, HF is more common in older adults that suffer from multiple comorbidities that can further confound the clinical presentation. Hence, international and Chinese HF guidelines also recommend screening for objective evidence of cardiac

dysfunction or structural impairment using diagnostic tests, including electrocardiogram, chest X-ray, echocardiogram and plasma levels of specific biomarkers for HF for classification of HF and subtypes of HF and to guide the use of treatment in clinical practice[1,2].

1.1.3.1. Left heart failure

The typical symptoms of left HF include breathlessness and fatigue, and the signs include the presence of a third heart sound and hepatojugular reflex[1,2]. A diagnosis of left HF is more plausible when these signs and symptoms present in patients with relevant medical histories such as hypertension and myocardial infarction[1,2].

Serum pro-brain natriuretic peptide (BNP) is highly specific in differentiating HF from other conditions with a similar presentation[14,15]. Pro-BNP is a hormone synthesised and secreted exclusively by the left ventricular cardiomyocytes when the myocardium is stretched due to increased pre or afterload. Heart failure is unlikely with low or normal serum pro-BNP levels[16]. In addition, electrocardiogram and echocardiography can help identify left ventricular dysfunction and the cause of left HF as well as guide treatment[1,2].

1.1.3.2. Right heart failure

The diagnosis of right HF is based on signs (e.g., enlarged liver and ascites) and symptoms (e.g., early satiety and breathing difficulties) in addition to the presence of relevant medical history (e.g., chronic lung disease, valvular disease and left HF), lifestyle factors (e.g., tobacco use) and family history (e.g., pulmonary arterial hypertension)[17]. In contrast to left HF, pro-BNP is highly sensitive for right HF but not specific, especially in the presence of left HF. Electrocardiogram can identify structural cardiac impairment specific to the right heart, including right ventricular hypertrophy and dilatation of the right atrium[17].

Medical imaging and invasive procedures are usually required to confirm a diagnosis of right HF. The gold standard diagnosis of right HF involves estimation of the right ventricular-pulmonary arterial coupling, a ratio of the right ventricular contractility to the pulmonary arterial elastance—a measure of the stiffness of the pulmonary arteries. In clinical practice, transthoracic echocardiography is the first-line diagnostic tool and enables a rapid assessment of the size and function of the right ventricle[9,17]. Cardiac magnetic resonance imaging can also be used to assess right ventricular function in case of poor image quality from transthoracic echocardiography—the complex geometry and position of the right ventricle can limit the use of transthoracic echocardiography to reliably assess right ventricular function[4]. Cardiac magnetic resonance imaging is the gold standard for evaluating right ventricular function due to its ability to capture images of the right ventricle in multiple planes[9]. Right heart catheterisation is an invasive technique to assess right ventricular function by directly measuring the pressures within the right ventricle and pulmonary arteries. However, access to such tests is limited because of their requirement for specialised equipment and highly-trained staff to obtain accurate pressure measures, and ventricular pressure does not always reflect ventricular volume[18].

1.2. Epidemiology of heart failure and pulmonary heart disease

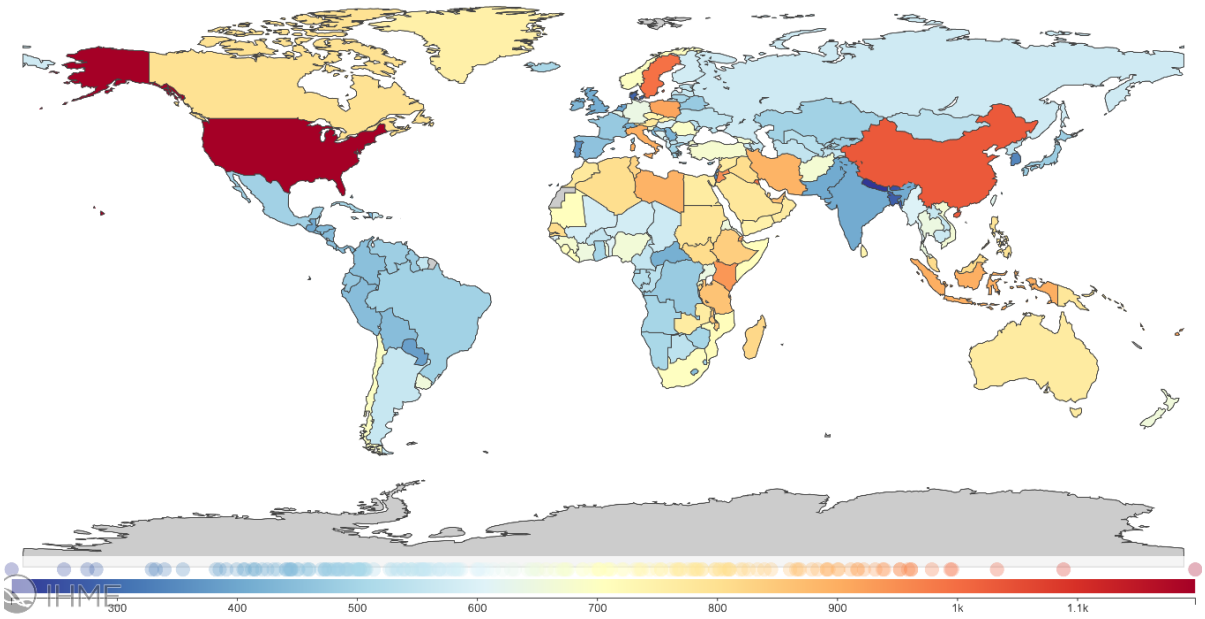
1.2.1. Epidemiology of heart failure

Reliable data on the epidemiology of HF was first reported in 1971 in a seminal publication from the Framingham Study, a population-based study of 5192 individuals without HF at baseline, entitled *"The natural history of congestive heart failure: the Framingham study"*[19]. This landmark study paved the way to understanding the epidemiology of HF, a condition that has evolved to be labelled as a pandemic in the 21st century, affecting about 56 million people worldwide in 2019[12,20].

The prevalence of HF in the general adult population is estimated at 1–3%, but this increases exponentially with age, attaining over 10% among people aged 70 years or older[1,21]. About 1 in 5 adults age 35 or older are expected to develop HF in their lifetime[22]. According to a Global Burden of Disease study, the global age-standardised prevalence of HF was 712 per 100,000 people in 2019 (**Figure 1.2**), but these differed between regions and countries worldwide, reflecting differences in the distribution of risk factors of HF, variable quality of health systems for diagnosis and reporting of HF cases and differential access to treatment for acute cardiac events between populations[20,23–25]. While the global age-standardised prevalence of HF has reduced or remained unchanged between 1990 and 2019, the prevalence estimates differed by income and socioeconomic status, with increases occurring in middle-income and high-income countries[26]. This growing number of HF cases, especially in middle-income countries, may reflect “Westernisation” of lifestyles, rapid improvement in life expectancy or improvements in diagnosis and reporting of HF cases[23,24].

HF has a particularly high mortality rate, which is more extreme than most cancers. In high-income countries, HF is associated with a 1-year mortality risk of 10% after first admission for HF and a 5-year risk of mortality of 42%[1,25]. However, there has been some reduction in HF mortality in high-income countries based on evidence from high-income countries. The Framingham Study reported a reduction in adjusted 5-year mortality rates from 70% in 1950 through 1969 to 59% from 1990 through 1999[27]. A meta-analysis of population-based studies of 1.5 million HF cases demonstrated a reduction in the 5-year mortality of HF from 70% between 1970 and 1979 to 40% between 2000 and 2009[28].

Figure 1.2. Global burden of disease estimates of the age-standardised prevalence of heart failure in 2019.



Prevalence is reported as cases per 100,000. The plot was generated using the open-source dataset and visualisation tool provided by the Global Burden of Disease group (https://www.healthdata.org/data-visualization/gbd-compare).

In 2019, China had the third highest age-standardised prevalence of HF after the United States and Guam[20]. Moreover, the majority (30%) of the global increase in HF cases between 1990 and 2017 occurred in China[21]. The 2012–2015 China Hypertension Survey—a nationally representative study of 31,494 adults aged ≥35 from 14 provinces in China—reported a prevalence of HF of 1.1%, representing about 7.5 million cases[29]. In a cohort of 3335 patients admitted for HF in 14 hospitals in the capital city, Beijing, the 5-year risk of mortality was 55.4%[30]. HF patients living in urban areas in China are estimated to be hospitalised at least three times a year, with each hospital admission costing about US $1,330[31].

1.2.2. Epidemiology of pulmonary heart disease

While the epidemiology of HF has been extensively described, especially in high-income settings, epidemiological evidence on PHD is scarce and rudimentary. In the United States, PHD accounts for 10–30% of all HF admissions. Data on the epidemiology of

PHD has chiefly been extrapolated from patients with chronic lung diseases, with autopsy reports showing evidence of PHD in about 40–50% of chronic lung disease cases[21,32]. COPD cases with right HF have a higher risk of stroke, severe COPD exacerbations requiring hospitalisation and a 1-year cumulative mortality of 44%, which is 4.4-fold greater than the 1-year mortality risk of HF[33–35].

In China, there have been few publications on the epidemiology of PHD. In a recent analysis, Xu et al. estimated the mortality rate of PHD from 2014–2021 in China using nationally representative data from the China National Mortality Surveillance System[36]. The findings demonstrated an age-standardised mortality rate of PHD of 62 per 100,000 people in 2014, which reduced to about 50% in 2021 (29 per 100,000 people)[36].

1.3. Major risk factors for heart failure and pulmonary heart disease

Heart failure occurs following exposure to risk factors that can be multifactorial and have synergistic effects. Following the seminal publication of the Framingham Study, population studies conducted in mainly high-income countries have identified blood pressure (BP), coronary artery disease, adiposity and diabetes as the chief risk factors for HF. This thesis will explore the role of SBP and adiposity as the risk factors of HF and PHD in Chinese adults. These exposures were selected because they are potentially modifiable through pharmacological and non-pharmacological interventions. The burden of these exposures and their associations with HF and PHD worldwide and in China are summarised below.

1.3.1. Blood pressure

1.3.1.1. Measurement of blood pressure

BP is believed to have been first described in India in around 150 BCE, where physicians described the relationship between what they referred to as “pulse quality” and ailments of the brain and heart[37]. Around 300 BCE, the first version of “Yellow

Emperor's Classic of Internal Medicine" described a phenomenon known as "hard pulse disease" and stated that too much salt in food hardens the pulse[38]. In 1628, William Harvey wrote the first description of the circulation of the blood in the body and how blood vessels transport blood pumped from the left side of the heart to the rest of the body and back to the right side of the heart[39]. Harvey's work was pivotal to our understanding of BP but was not fully appreciated for another 300 years.

In 1896, Riva Rocci invented a method of applying uniform pressure around the arm to temporarily restrict arterial blood flow using an inflatable cuff[40,41]. This procedure of restricting blood flow is essential for determining systolic BP (SBP: the pressure at which blood flow is completely stopped when the cuff is inflated) and diastolic BP (DBP: the pressure at which blood flow is restored after relieving the pressure in the inflated cuff)[40,41]. This is the procedure used in most modern BP measurement devices. In 1901, a Russian surgeon, Nikolai Korotkoff, identified a method of using a stethoscope to detect the "click" sounds produced when the blood flow is completely restricted after applying pressure to the arm (i.e., the SBP) and when blood flow is restored after this pressure is relieved (i.e., the DBP)[40,41]. These two methods form the basis of BP measurement that are still used today.

1.3.1.2. Epidemiology of blood pressure

Most guidelines define hypertension or elevated BP as systolic BP (SBP) ≥140 mmHg or diastolic BP (DBP) ≥90 mmHg[42–44], except for the American College of Cardiology/American Heart Association guidelines that recommend lower thresholds for defining elevated BP (SBP ≥130 or DBP ≥80)[45]. Elevated BP is a leading risk factor for cardiovascular diseases such as stroke, ischaemic heart disease and HF, in addition to being the major causes of premature death worldwide, with 8–10 million deaths attributed to hypertension each year[46].

High BP affects about 1.28 billion people aged 30–79 years worldwide, most of whom live in low- and middle-income countries (LMICs). About half (46%) of individuals with high BP are unaware of their status; of those who know their status, only 42% are treated; only 21% of those on treatment have their BP under control. Although the global age-standardised prevalence of hypertension remained stable from 1990–2019, the absolute number of people living with hypertension has doubled during this period, with the majority of this increase occurring in LMICs[47]. The increase in the number of cases with hypertension in LMICs may reflect improvements in life expectancy in these countries and an increase in the proportion of older adults, in addition to the increasing prevalence of risk factors for hypertension, such as obesity and physical inactivity[48].

A nationally representative community-based study of 642,523 adults aged 18–69 established in 2004 to estimate and monitor the prevalence of major risk factors for leading non-communicable diseases showed that the age-standardised prevalence of hypertension in China was 30% in 2018[49]. Between 2004 and 2018, the age-standardised prevalence of hypertension increased steadily from 21% in 2004 to 30% in 2010 and has been decreasing modestly since then. In China, about 274 million adults aged 18–69 had hypertension in 2018, which is greater than the number of cases of hypertension in high-income countries such as the United States (115 million)[50], United Kingdom (11.8 million) and Japan (43 million)[51].

Despite the high number of cases of hypertension compared to high-income countries, China had substantially lower rates of hypertension awareness (38%), treatment (35%) and control (12%) in 2018[49]. A pooled analysis of 1201 population-based studies worldwide, including 104 million individuals, showed that Chinese men and women aged 30–79 years had lower awareness (men: 47.7%, women: 56%), treatment (35%, 47%) and control (14%, 18%) rates compared to men and women in the United States (awareness: men=78%, women=83%; treatment: 66%, 73%; control: 45%, 51%), the

United Kingdom (awareness: 60%, 58%; treatment: 47%, 48%; control: 31%, 29%) and Japan awareness: 66%, 69%; treatment: 46%, 51%; control: 24%, 30%)[47]. The implications for the high prevalence and corresponding low control rates of hypertension in China is a high incidence and prevalence of the complications of hypertension, including stroke, HF and kidney disease.

1.3.1.3. Hypertension and risk of HF and PHD

There is reliable evidence from prospective cohort studies, mostly conducted in high-income countries, that higher levels of SBP and DBP are associated with higher risks of HF. Recent evidence on the risks of BP with cardiovascular disease, including HF, have focused on SBP because it is associated with greater relative risks than diastolic BP, pulse pressure or mean arterial pressure and are modifiable by drug treatment. An individual participant data (IPD) meta-analysis of 48 randomised controlled trials demonstrated a 13% reduction in the risk of HF for every five mmHg reduction in SBP, indicating that elevated SBP causes HF[52]. Higher levels of SBP increase the risk of HF by ventricular remodelling, myocardial fibrosis and atherogenesis[53].

While there is strong evidence for SBP as a risk factor for HF, controversy persists about the shape and strength of these associations, especially in LMICs with high rates of uncontrolled hypertension and where hypertension is the leading cause of hypertension. Population-based observational studies, including disease-specific cohorts such as cohorts of patients with diabetes, have reported linear and J-shaped associations of SBP with HF[54,55]. These uncertainties have implications for population health as reliable estimates of the strength of this association are required to assess the burden of HF attributable to elevated SBP. Furthermore, understanding the shape of the association is necessary to inform policymakers on the range of SBP levels within which individuals remain at increased risk of HF.

In contrast to HF, no study has evaluated the associations between SBP and PHD. Since chronic lung disease is a prerequisite for PHD, one can only extrapolate the associations between BP and PHD from the relationships between BP and chronic lung disease or lung function. Previous studies have reported an inverse association between hypertension and lung function[56] but no association of hypertension with COPD[57]. However, the extent to which the association between BP and lung disease is due to reverse causality—a phenomenon where the exposure-outcome association is biased by pre-existing disease[58]—is uncertain since hypertension is also a common complication of chronic lung disease[59].

1.3.2. Adiposity

1.3.2.1. Definition, measurement and classification of adiposity

After a meal rich in carbohydrates, the body's main and preferred source of energy, the body breaks down the food to produce the energy needed for normal body function[60]. The excess energy from this process is stored as adipose tissue, chiefly located beneath the skin and internal organs[60]. The stored fat is then metabolised during prolonged fasting or exercise as an alternative source of energy. Adiposity is the accumulation of fat in the body and can be broadly classified into central and general adiposity, depending on the location where fat is accumulated[60,61]. Excessive adiposity leads to obesity[60].

Several techniques have been used to quantify adiposity. The gold standard methods include body imaging, such as dual-energy x-ray absorption (DXA) and densitometry[61]. However, these methods have limited use, especially in large-scale epidemiological studies. The tools used for imaging are expensive and cumbersome and require specialised staff to perform such measurements[61]. In addition, densitometry is time-consuming and inconvenient, especially for children and older adults, because it

requires submerging individuals in water[61]. Hence, inexpensive anthropometric measures, including BMI, waist circumference and waist-to-hip ratio (WHR), are simple, reliable and widely used methods to classify adiposity in clinical practice and in epidemiological studies[61].

BMI is an index based on weight and height [weight/height]^2, which are the anthropometric measures with the highest precision (or reproducibility) and accuracy and the least technical error[61]. Even though BMI does not represent body fatness, it correlates strongly with DXA-derived percentage body fat (correlation coefficient (r): women ~ 0.88; men ~ 0.76) and biochemical markers of obesity such as leptin (r ~ 0.73)[61]. Hence, BMI is the most common anthropometric index used in epidemiological studies.

Central adiposity refers to the accumulation of fat around the organs in the lower abdomen, and this includes deposits of fat beneath the skin (subcutaneous fats), within the abdominal cavity or around visceral organs including the stomach and liver (visceral fats) or behind the peritoneum (retroperitoneal fats)[61,62]. Of these, visceral adiposity has been associated with poor cardiometabolic risk profile and elevated risk of disease independent of general adiposity[63,64]. The most common anthropometric methods to indirectly measure central adiposity include waist circumference and WHR[61]. Compared to BMI, these methods have greater variability and, hence, have a more limited reproducibility and are more prone to measurement error[61]. Despite this limitation, studies have demonstrated that central adiposity is a better measure of body fatness than BMI and a stronger predictor of most major diseases in older adults over 65 years[65,66]. These findings highlight the importance of combining central and general adiposity measures when investigating the associations of adiposity with disease risk.

There are substantial differences in fat distribution between ethnic groups and these differences have important implications for health[67,68]. Compared to Caucasians, Africans and Hispanics, East Asians have a higher accumulation of visceral fat for any given BMI[68,69]. Hence, East Asians have a higher risk of cardiometabolic diseases than other populations at lower thresholds of BMI, reflecting their higher proportion of visceral adiposity[70]. In addition, East Asians have less accumulation of subcutaneous fat, which confers less metabolic risk than visceral fat, compared with other populations[69]. Investigating the associations of adiposity with HF and PHD in diverse populations could improve our understanding of how these associations differ between populations and inform ethnic-specific health policies.

1.3.2.2. Epidemiology of obesity

Similar to BP, population levels of adiposity have been increasing globally chiefly due to the urbanisation of lifestyle that accompanies economic development. Over the last four decades, the mean BMI increased by 2.4 kg/m^2, from 21.9 kg/m^2 in 1975 to 24.3 kg/m^2 in 2014[71]. During this time, the age-standardised prevalence of obesity (BMI ≥ 30 kg/m^2) increased by 3-fold in men (from 3.2–10.8%) and 2-fold in women (6.4–14.9%). In 2015, about 640 million adults (266 million men and 375 million women) were obese worldwide: most of these (27%) lived in high-income countries[71]. Obesity is typically associated with four million deaths per year, most of which are cardiovascular disease and cancer-related deaths[72].

Figure 1.3. Changes in the global number of cases of obesity in men and women between 1975 and 2014.

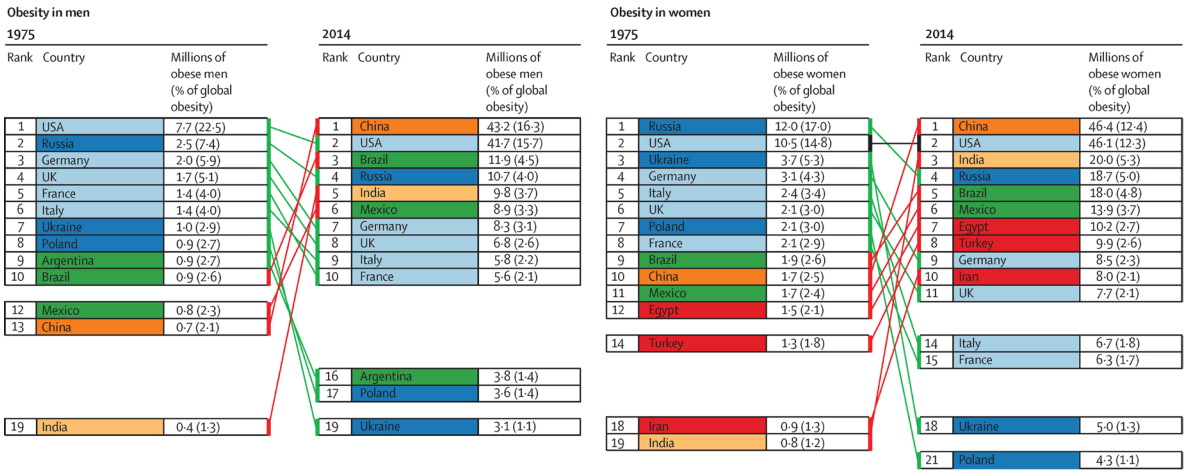


Figure reproduced with permission from the Journal[71].

The epidemiology of obesity in China is comparable to or worse than that observed worldwide. China witnessed a substantial increase in the prevalence of obesity between 1975 and 2014, with the greatest number of cases worldwide in 2014 living in China (**Figure 1.3**)[71]. Consistent with global BMI trends, the mean levels of BMI in China increased by 3 kg/m^2 between 1982 and 2014 from 21 to 24 kg/m^2 [73,74]. In addition, the prevalence of obesity among Chinese adults increased 2-fold between 1992 (3.6%) and 2018 (8.1%)[73,74]. More recent estimates suggest that about 85 million Chinese adults aged 18–69 are obese, which is 3-fold greater than when estimated in 2004[74].

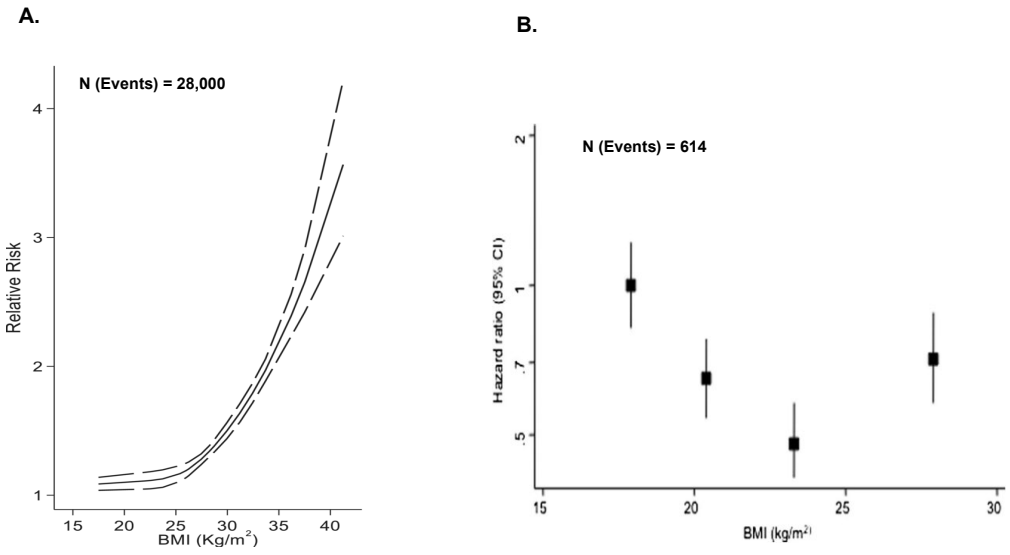
1.3.2.3. Adiposity and risk of HF and PHD

Large meta-analyses of prospective cohort studies, chiefly involving studies from high-income countries, have demonstrated higher risks of HF with higher levels of general (BMI) and central adiposity (waist circumference and WHR)[75,76]. The causal relevance of these associations were supported by findings in a recent Mendelian randomisation study[77], suggesting that higher levels of adiposity cause HF. Adiposity is likely to cause HF by promoting atherogenesis and the development of myocardial infarction through elevated BP, diabetes mellitus and chronic inflammation[53]. Moreover, obesity increases

the risk of obstructive sleep apnoea and, consequently, right HF owing to increased afterload[78]. However, almost all of the studies included in the meta-analysis and Mendelian randomisation were conducted in high-income countries. Uncertainties persist about the shape and strength of the associations of adiposity with HF and PHD in LMICs.

In contrast to the meta-analysis by Oguntade et al., 2023[75] showing a positive log-linear association between BMI and incident HF at BMI levels >25 kg/m^2, an IPD meta-analysis of prospective cohort studies from the Asian-Pacific Region, including 40% of studies from China, demonstrated an inverse association between BMI and HF mortality[79]. Therefore, the extent to which the associations of BMI with HF differ between diverse populations and how much of this variation is due to reverse causality remains unclear.

Figure 1.4. Meta-analyses on the association of body mass index with heart failure.



Summary-level meta-analysis of studies conducted in predominantly Western countries (left panel)[69]. Individual participant data meta-analysis of prospective cohort studies conducted in the Asian-Pacific Region (right panel)[73]. The figures were reproduced with permission from the journals.

There are no studies investigating the association between adiposity and PHD, but studies have shown inverse associations between higher levels of adiposity with the risk of COPD even after extensive adjustment for the major confounders[80,81]. Adiposity is likely to have a similar association with PHD since COPD is the chief cause of PHD. A potential biological mechanism for this association is that low BMI or underweight reduces respiratory muscle mass, leading to COPD[82]. Alternative explanations for the inverse association of BMI with COPD include reverse causality and residual confounding.

1.4. Structure of the thesis

Chapter 2 summarises the findings of the literature review, which uses a systematic approach to identify all eligible published studies on the association of SBP and adiposity (BMI and WHR) with HF and PHD. This chapter aimed to identify the research gaps and inform the objectives of this thesis. The overall and specific objectives of the thesis are presented at the end of this chapter. Chapter 3 describes the design of the CKB study and the characteristics of the participants at baseline and resurvey. Chapter 4 assesses the burden of HF and PHD in the CKB and discusses the value of investigating modifiable risk factors for these conditions. Chapters 5 and 6 investigate the observational associations of SBP with HF and PHD (Chapter 5) and adiposity (BMI and WHR) with HF and PHD (Chapter 6). Chapters 7 and 8 use Mendelian randomisation approaches to evaluate the causal relevance of the associations of SBP and adiposity with HF and PHD. Finally, Chapter 9 discusses the overall findings of this thesis, assesses their strengths and limitations and implications for clinical practice and future research.

2. Literature review

Chapter 1 demonstrates that although previous studies have demonstrated associations of SBP and adiposity (BMI and WHR) with HF, there is limited evidence on the shape and strength of these associations, especially in non-Western countries, including China. In addition, the chapter highlights limitations in the associations of SBP, BMI and WHR with PHD in any population. The present chapter reviews the published prospective cohort studies on the shape and strength of the associations of SBP or adiposity (BMI and WHR) with HF and PHD, respectively. The chapter then identifies gaps in the available evidence to inform the objectives of this thesis. The associations of each exposure with HF and PHD are summarised under separate headings.

2.1. Objectives

The literature review was conducted to (1) summarise the available evidence on the shape and strength of the associations of SBP, BMI and WHR with HF and PHD globally and (2) identify gaps in the available evidence to inform the objectives of this thesis.

2.2. Methods

A systematic approach was used to identify all relevant literature on the topic. Medline, Embase and Global Health Library were searched through Ovid® from database inception to 14 November 2023 with no language restrictions. A combination of search terms was used to design the search algorithm. **Table 2.1** shows the search strategy for Medline, which was adapted to suit other databases.

Database searches were supplemented by a search of Google Scholar and ResearchGate to identify articles that were missed during the search. The reference lists of relevant review articles were also hand-searched to identify relevant articles.

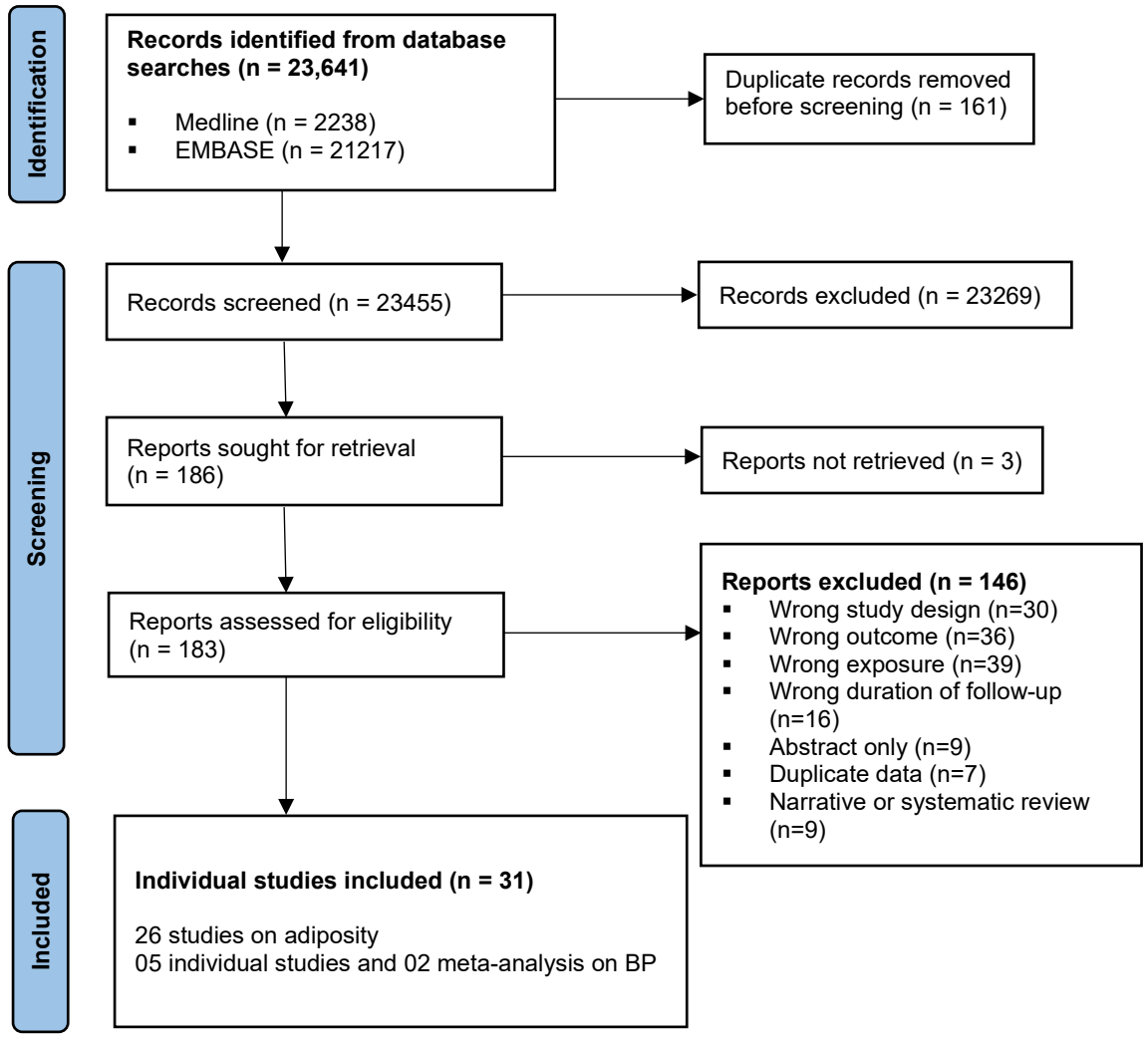
Table 2.1. Search strategy for Medline.

SN	Search terms
1.	exp Adiposity/ or exp Obesity/ or Body Weight/ or (adiposity or obese).mp
2.	exp Blood Pressure/ or exp Hypertension/ or (Blood pressure or hypertension or hypertensive$1).mp.
3.	exp Arterial Pressure/ or (arterial pressure or arter$ pressure$).mp.
4.	or/1-3
5.	exp Heart Failure/ or (heart failure or cardiac failure or cardiac insufficienc*).mp.
6.	exp Pulmonary Heart Disease/ or (pulmonary heart disease or cor pulmonale).mp.
7.	5 or 6
8.	4 and 7
9.	exp animals/ not humans.sh.
10.	8 not 9

Eligible citations identified through the database searches were exported to Endnote X20 for removal of duplicates and then exported to Covidence for screening against the eligibility criteria for this review based on title, abstract and full text. The inclusion criteria for the literature review were: (1) Population-based prospective cohort or case-cohort studies involving at least 500 participants with no prior heart disease or stroke at baseline and at least five years of follow-up. Meta-analyses of these studies were also included; (2) Studies conducted in people ≥15 years old: most studies on HF have included individuals as young as 15 years; (3) the outcome was incident fatal or non-fatal HF or PHD. The exclusion criteria were: (1) studies that reported the association between a risk factor and HF or PHD from a predictive model after adjusting for potential mediators. For example, studies that evaluated SBP as the main exposure and adjusted for left ventricular hypertrophy or studies that investigated adiposity as the main exposure and adjusted for diabetes; (2) Studies that did not adjust for age, which is a strong confounder for most observational associations; (3) For duplicated publications (two or more publications involving the same dataset), only the most comprehensive or recent article was included. The inclusion criteria for the review were carefully selected to include studies that were least influenced by reverse causality, which is a major source of bias in studies investigating the risk factors for chronic diseases with an insidious onset, such as HF and PHD[83–85]. **Figure**

2.1 summarises how citations identified through the literature search were included in this literature review.

Figure 2.1. Flow diagram for study selection.



2.3. Results

2.3.1. Systolic blood pressure and risk of HF and PHD

2.3.1.1. Heart failure

Three meta-analyses (two IPD[79,86] and one summary-level[87] meta-analysis) of prospective cohort studies and nine individual prospective studies reported on the association of SBP with risk of HF[54,55,87–93]. The findings of these studies demonstrated consistent positive associations of higher levels of SBP with higher risks of HF in the general population[79,86–88], individuals registered with a primary care practice[89] and individuals with diabetes[54,55]. The Asia-Pacific Cohort Studies Collaboration (APCSC) and a study conducted by Choi et al., 2018 in Korea were the only studies conducted in non-Western populations[79,91].

Using data from the population-based Framingham Study including 3362 participants with no prior HF in the United States and 518 incident HF cases, Lee et al., 2007 demonstrated a 12% higher risk of HF (HR 1.12, 95% confidence interval [CI]=1.06–1.18) for each 10 mmHg higher usual SBP[87]. Rapsomaniki et al., 2014[89] investigated the association of usual SBP with HF using data from the CALIBER (CArdiovascular research using LInked Bespoke studies and Electronic health Records) study, a large-scale prospective study of 1.25 million individuals with no prior history of heart failure aged ≥30 years that were recruited from 225 primary care practices in the United Kingdom. Similar to the Framingham study, the authors reported a 13% higher risk of HF (1.13, 1.11–1.15) for each 10 mmHg higher usual SBP. The APCSC was an IPD meta-analysis of ~543000 participants and 614 deaths from HF cases from 32 prospective cohort studies conducted in the Asia-Pacific Region including China, Hong Kong, South Korea, Japan, Taiwan and Australia[79]. The APCSC reported similar associations between usual SBP and risk of death from HF (1.13, 1.06–1.21). The estimates from

these studies were supported by a recent summary-level meta-analysis of 24 prospective cohort studies published up to 10 June 2022, indicating a 13% higher risk of HF per 10 mmHg higher levels of SBP (1.13, 1.10–1.16)[94]. Most of the studies included in this meta-analysis were recruited from Western countries: twelve were from North America, 11 were from Europe, and only one was from Asia (Korea).

However, other studies have reported discrepant results about the strength of the association of SBP with risk of HF. Randolph et al., 2014 reported a weaker association of baseline SBP and incident HF using data from the Jackson Heart Study, a population-based cohort study of 5280 Black Americans[88]. The strength of the association was only half as strong as those reported in the CALIBER, APCSC or the meta-analysis by Baffour et al. 2023[94]. The discrepant results may reflect regression dilution bias (RDB) or reverse causality bias as the study did not exclude cases with prevalent heart disease. The strength of the association of usual SBP with risk of fatal HF in the Prospective Studies Collaboration (PSC) was over 2-fold greater than the estimates reported in the Framingham Study, CALIBER and APCSC[86]. Each 10 mmHg higher usual SBP was associated with a 37% higher risk of death from HF (1.37, 1.30–1.44). The PSC was an IPD meta-analysis of data from one million participants with no prior stroke or heart disease and 746 deaths due to HF from 61 prospective cohort studies (Over 80% of the participants from Western countries, with only 1.8% of the participants from China)[86]. The reason for this high estimate is unclear. One possible explanation for the discrepant results may include incomplete adjustment for confounding as the PSC only adjusted for age, sex and study. However, the APCSC, which is a sub-study of the PSC, used similar adjustments as in the PSC but instead reported a similar estimate to that reported in the Framingham and CALIBER studies.

Table 2.2. Population-based prospective cohort studies of the associations of systolic blood pressure with incident heart failure.

Author, Year	Country of study	Study name	Population	Exposure	Comparison	Outcome	FU, Years	Event (N)	HR (95% CI)
Individual studies									
Magnussen, 2019[90]	Multi-country	FINRISK DanMONICA Moli-sani Northern Sweden	Population-based	SBP	10 mmHg higher levels	Non-fatal HF	12.7	5170	**Men:** 1.04 (1.02–1.06) **Women:** 1.09 (1.06–1.11)
Choi, 2019[91]	Korea	Korea National Health Insurance Research Database	Population-based	SBP	**Ref**: **90–99** mmHg		6.7	5248	**<90:** 2.70 (1.36-5.37) **90–99: 1.00 (Reference)** **100–109:** 1.31 (0.89-1.93) **110–119:** 1.43 (0.99-2.06) **120–129:** 1.60 (1.11-2.29) **130–139:** 1.99 (1.38-2.86) **140–149:** 2.47 (1.71-3.57) **150–159:** 2.89 (1.98-4.21) **≥160:** 4.16 (2.86-6.06)
Eryd, 2016[54]	Sweden	NA	Participants with at least one year of diabetes	SBP	**Ref**: 130-139 mmHg	Non-fatal HF	5	6452	**110–119:** 1.20 (1.01–1.42) **120–129:** 1.02 (0.91–1.15) **130–139: 1.00 (Reference)** **140–149:** 1.09 (0.98–1.20) **150–159:** 1.27 (1.13–1.43) **≥160:** 1.40 (1.25–1.56)
Randolph, 2016[88]	USA	Jackson Heart Study	Black Americans	SBP	10 mmHg higher levels	Non-fatal HF	7	340	1.07 (1.00 – 1.14)
Rapsomaniki, 2014[89]	UK	CALIBER	Individuals registered with primary practice	SBP	10 mmHg higher levels	Fatal and non-fatal HF	5.2	10437	1.13 (1.11–1.15)
Zhao, 2014[55]	USA	NA	White and African American participants with diabetes.	SBP	**Ref:** 130-139 mmHg	Fatal and Non-fatal HF	6.5	5089	**<110:** 2.49 (2.02–3.07) **110–119:** 1.59 (1.39–1.81) **120–129:** 1.16 (1.06–1.28) **130–139: 1.00 (Reference)** **140–159:** 1.09 (1.02–1.18) **≥160:** 1.30 (1.18–1.42)
Borne, 2012[92]	Sweden	Malmo Diet and Cancer Study	Population-based	SBP	10 mmHg higher levels	Non-fatal HF	15	764	1.15 (1.11–1.20)

Butler, 2011[93]	USA	Cardiovascular Health Study and Health ABC Study	Population-based	SBP	Ref: <120 mmHg	Non-fatal HF	10	493	**<u>Men</u>** **<120:** 1.00 **120–139:** 1.25 (0.88–1.77) **140–159:** 1.84 (1.28–2.64) **≥160**: 2.11 (1.38-3.23) **<u>Women</u>** **<120:** 1.00 **120–139:** 2.51 (1.55-4.06) **140–159:** 3.09 (1.87-5.11) **≥160**: 3.74 (2.15-6.50)
Lee, 2007[87]	USA	Framingham Heart Study	Population-population	SBP	Per 10 mmHg higher usual levels	Non-Fatal HF	NR	518`	1.12 (1.06-1.18)
Meta-analyses									
Huxley, 2014[95]†	Multiple countries ††	Asia Pacific Cohort Studies Collaboration	Population-based and occupational settings	SBP	Per 10 mmHg higher usual levels	Fatal HF	7.0	614	1.13 (1.06 – 1.21)
PSC, 2002[86]†	Multiple countries §	PSC	Population-based	SBP	Per 10 mmHg higher usual levels*	Fatal HF	13.3	746	1.37 (1.30 – 1.44)
Baffour, 2023[94]	Multi-country	NA	Cohort studies (retrospective, prospective, and case-cohort) and nested case-control	SBP	Per 10 mmHg higher usual levels*	Fatal and non-fatal HF	NA	31,639	1.13 (1.10-1.16)

† Individual participant data meta-analysis of population-based prospective cohort studies.
†† Included studies from Australia, China, Hong Kong, Japan, Korea, and Taiwan
§ Study participants were from Europe (70%), North America or Australia (20%), and Japan and China (10%).
* The risk was reported per 20 mmHg lower usual SBP but has been converted to per 10 mmHg higher usual SBP to allow comparisons across studies.
CALIBER = CArdiovascular research using LInked Bespoke studies and Electronic health Records; CHS = Cardiovascular Health Study; CVD = Cardiovascular diseases; FU = Follow-up; HF = Heart failure; NA = Not applicable; NR = Not reported; PSC = Prospective Studies Collaboration

Four studies reported discrepant results on the shape of the associations of SBP with HF risk[54,55,89,91], with three studies reporting J-shaped associations (**Table 2.2**)[54,55,91] and one reporting a linear association[89]. It is unclear whether the J-shaped associations are real or due to reverse causality, although these studies excluded individuals with CVD at baseline[54,55,91]. Zhao et al., 2014 reported a J-shaped association between SBP and risk of HF among patients newly diagnosed with diabetes and no history of coronary heart disease (CHD) or HF from public hospitals and clinics (**Table 2.2**). The J-shaped association persisted despite additional exclusion of participants who developed HF within the first two years of follow-up[55]. In a cohort of 187,100 patients with diabetes and no history of history of CVD or major diseases at baseline recruited from primary care units and hospital outpatient clinics, Eryd et al., 2016 reported a J-shaped association despite excluding patients with prevalent CVD, in addition to a wide range of prior diseases at baseline, including dementia, cancer and renal disease[54]. Choi et al. 2019 used data of patients with no prior stroke or heart disease from the Korean National Health Insurance Services Health Screening Cohort. After a median duration of 6.7 years of follow-up, they observed a J-shaped association between SBP and risk of HF.

Two studies reported effect modification of the association between SBP and HF by age, with the strongest proportional association in the youngest age group were that attenuated by about 60–80% in the oldest age group but remained significant[89,91]. In contrast, although the APCSC[79] and Zhao et al.[55] also reported an attenuation of the strength of the association with age, there was no evidence of an association in the older age group. The discrepant results may reflect the low power in the latter studies compared to the former study[89]. Two studies reported evidence of effect modification of the association of SBP with HF with stronger effects observed in women than in men, but the Framingham Study, APSC and CALIBER studies reported no evidence of effect modification by sex[79,87,89]. These discrepancies are unlikely to be explained by a lack of

power as the CALIBER study was sufficiently powered to investigate effect modification by sex; an alternative explanation for the significant modification by sex could be chance findings.

2.3.2. Pulmonary heart disease

The literature search, despite being broad an extensive, did not identify any literature on the association of SBP with risk of PHD. No relevant study was found after screening the references of review articles on PHD and searching Google Scholar and ResearchGate for articles using the keywords “blood pressure,” “pulmonary heart disease” and “cor pulmonale.”

2.3.3. Limitations of the available studies and implications for future studies

Despite being an established risk factor for HF and the leading cause of HF in China, uncertainties persist about the shape and strength of the association of SBP with HF. In addition, no study investigated the relevance of confounding and reverse causality bias on the associations of SBP with HF. Large studies are needed to reliably assess the risk of disease across levels of a risk factor and explore how much of the association is mediated through a third variable. Such data are essential to reliably estimate the burden of HF attributed to SBP and the risk of HF associated with the lowest levels of SBP. Furthermore, epidemiological studies require an appropriate duration of follow-up to evaluate and control for reverse causality bias. Moreover, no study has explored the association between SBP and PHD.

2.3.4. Adiposity and risk of HF and PHD

2.3.4.1. Heart failure

A recent global meta-analysis of prospective cohort studies[75] and 18 individual prospective cohort studies reporting on the association between BMI or WHR with HF were included in this synthesis (**Table 2.3**). The studies were mostly conducted in Europe and the USA. Only two studies included participants from non-Western countries[79,96]. All of the individual studies reported on BMI, while only five studies reported on WHR[97–101]. Most of the studies suggested a positive non-linear association of BMI and WHR with HF[75,97,102,103]. In contrast, a few studies from non-Western populations suggested an inverse association of BMI with HF[79,96]. There was no study on the association between WHR and HF in non-Western countries. In addition, no study investigated the associations of BMI and WHR with PHD even in Western countries.

2.3.4.1.1. Body mass index and HF

Most studies reported a positive association between BMI and the risk of HF (**Table 2.3**). In a recent summary meta-analysis of 35 prospective cohort studies mostly conducted in Western countries, including 34,422 incident HF cases, Oguntade et al., 2023 demonstrated higher risks of HF associated with higher levels of BMI. Each 5 kg/m^2 higher BMI was associated with a 42% higher risk of HF (1.42, 1.40–1.44). Studies conducted mostly in Western countries have reported consistent positive associations of BMI with HF[97,98,104–107]. The estimates from previous studies are difficult to compare because most studies used different thresholds for BMI to estimate HF risk. Xing et al., 2023 used data from the UK Biobank, including 0.5 million participants with no history of CVD at baseline and 3134 incident HF cases, and demonstrated similar results to those reported in the meta-analysis by Oguntade et al. (1.40, 1.25–1.56)[104].

Table 2.3. Population-based prospective cohort studies of the associations of adiposity with incident heart failure.

Author, Year	Country of study	Study name	Population	Exposure	Comparison	Outcome	FU, year	Events (N)	HR (95% CI)
Xing, 2023[104]	UK	UK Biobank	Population-based study	BMI	Per 1 SD (4.7 kg/m^2)	Non-fatal HF	12.1	3134	1.40 (1.25–1.56)
Björck, 2020[102]	Sweden	Swedish Medical Birth Register	Population-based study of women; Age range: 18–45 years	BMI	20 – <22.5 kg/m^2	Non-fatal HF	33	2513	<18.5: 1.50 (1.21–1.86) 18.5–<20.0: 1.09 (0.93–1.28) 20–<22.5: 1.00 (Reference) 22.5–<25.0: 1.24 (1.10–1.39) 25–<27.5: 1.56 (1.36–1.78) 27.5–<30: 2.39 (2.05–2.78) 30–<35: 2.82 (2.43–3.28) ≥35: 4.51 (3.63–5.61)
Kokkinos, 2019[105]	USA	ETHOS Veteran cohort	Male veterans, Mean age: 58 years	BMI	Per 1 kg/m^2	Non-fatal HF	13.4	2979	1.02 (1.01–1.03)
Janszky, 2016[103]	Norway	HUNT2	Population-based; Women: Mean age = 61.0 years	BMI	<24.9 kg/m^2	Non-fatal and fatal HF	11.4	946	**Ref :** 1.00 **25.0–27.4:** 1.07 (0.84–1.36) **27.5–29.9:** 1.26 (0.98–1.61) **30.0–32.4:** 1.26 (0.93–1.70) **32.5–34.9:** 1.84 (1.30–2.59) **≥35:** 2.65 (1.86–3.77)
Ndumele, 2016[108]	USA	Atherosclerosis Risk in Communities	Men and women Mean age: 62 years	BMI	Normal weight: 18.5–24.9 kg/m^2	Non-fatal and fatal HF	12.1	869	18.5–24.9: 1.00 (Reference) 25–29.9: 1.12 (0.99–1.26) 30–34.9: 1.50 (1.32–1.72) >35: 2.27 (1.94–2.64)
Björck, 2015[106]	Sweden	Primary prevention study	Men, 47–55 years old	BMI	Per 1 kg/m^2	Non-fatal and fatal HF	35	1855	1.02 (1.01–1.04)
Cui, 2014[96]	Japan	Japan Collaborative Cohort Study	Population-based Men and women Age range: 40–79 years	BMI	23.0–24.9 kg/m^2	Fatal HF	19.3	605	**<19.0:** 1.83 (1.41–2.37) **19.0–20.9:** 1.20 (0.9.–1.55) **21.0–22.9:** 1.22 (0.96–1.56) **23.0–24.9:** 1.00 **25.0–26.9:** 0.89 (0.64–1.25) **≥27.0:** 1.17 (0.81–1.68)
Borné, 2014[97]	Sweden	Malmø Diet and Cancer Cohort	Population-based, Men and women, Age range: 45–73 years	BMI; WHR	BMI: ≤22.5 kg/m^2 WHR: ≤0.81	Non-fatal and fatal HF	15	727	**BMI** Quartile (Q) 1: 1.00 (Reference) Q2 (24.9): 0.98 (0.76–1.25) Q3 (26.9): 1.12 (0.88–1.42) Q4 (30.0): 1.80 (1.45–2.24) **WHR** Q1 (median WHR: 0.81): 1.00 Q2 (0.85): 1.04 (0.82–1.32) Q3 (0.88): 1.13 (0.90–1.42) Q4 (0.93): 1.77 (1.43–2.19)
Djousse, 2012[107]	USA	Cardiovascular Health Study	Population-based, Men and women,	BMI	BMI: Per 4.45 kg/m^2	Non-fatal HF	11.3	1381	BMI: 1.22 (1.15–1.29)

Author, Year	Country of study	Study name	Population	Exposure	Comparison	Outcome	FU, year	Events (N)	HR (95% CI)
			Age range: 65–100 years						
Van Lieshout, 2011[98]	Netherlands	The Rotterdam Study	Population-based, Men and women, age ≥55 years	BMI; WHR	**Men** BMI: Per 3.7 kg/m^2 WHR: Per 0.09 **Women** BMI: Per 3.7 kg/m^2 WHR: Per 0.09	Non-fatal and fatal HF	15	727	**Men** BMI: 1.21 (1.09–1.34) WHR: 1.00 (0.90–1.11) **Women** BMI: 1.19 (1.07–1.31) WHR: 1.22 (1.10–1.35)
Hu, 2010[99]	Finland	FINRISK Studies	Men and women, Age range: 25–74 years	BMI; WHR	**Men** BMI: ≤25 kg/m^2 WHR: 0.88-<0.93 **Women** BMI: ≤25 kg/m^2 WHR: 0.76-<0.80	Non-fatal and fatal HF	18.4	3614	**1. Men** (a) BMI <25: 1.00 25–29.9: 1.25 (1.12–1.39) ≥ 30: 1.99 (1.74–2.27) (b) WHR Q1 (<0.88): 0.88 (0.58–1.31) Q2 (0.88–0.92): 1.00 Q3 (0.93–0.98): 1.06 (0.74–1.50) Q4 (>0.98): 1.71 (1.23–2.37) **2. Women** (a) BMI <25: 1.00 25–29.9: 1.33 (1.16–1.51) ≥ 30: 2.06 (1.80–2.37) (b) WHR Q1 (<0.76): 0.61 (0.31–1.19) Q2 (0.76–0.79): 1.00 Q3 (0.80–0.85): 0.98 (0.59–1.63) Q4 (>0.85): 1.88 (1.17–3.01)
Loehr, 2009[100]	USA	The Atherosclerosis Risk in Communities Study	Population-based, Men and women Age range: 45–64 years	BMI; WHR	**Men** BMI: Per 4.2 kg/m^2 WHR: Per 0.05 **Women** BMI: Per 6 kg/m^2 WHR: Per 0.08	Non-fatal and fatal HF	16	1528	**Men** BMI: 1.47 (1.39–1.57) WHR: 1.50 (1.41–1.60) **Women** BMI: 1.49 (1.39–1.59) WHR: 1.59 (1.46–1.72)
Kenchaiah, 2009[109]	USA	Physicians' Health Study	Physicians, Men, Age range: 40–84 years	BMI	Per 1 kg/m^2	Non-fatal HF	20.5	1109	1.13 (1.11–1.15)
Levitan, 2009[101]	Sweden	Cohort of Swedish Men	Population-based, Men, 45–79 years old	BMI; WHR	BMI: Per 1 kg/m^2 WHR: Per 0.08	Non-fatal and fatal HF	7	718	BMI: 1.07 (1.05–1.08) WHR: 1.10 (1.03–1.18)
Thrainsdottir, 2007[110]	Iceland	The Reykjavik Study	Population-based, Women,	BMI	Per 1 kg/m^2	Non-fatal HF	13	489	1.09 (1.07–1.11)

Author, Year	Country of study	Study name	Population	Exposure	Comparison	Outcome	FU, year	Events (N)	HR (95% CI)
			Age range: 33–84 years						
Murphy, 2006[111]	Scotland	Renfrew–Paisley Study	Women, Mean age 54 years	BMI	Normal weight: 18.5–24.9 kg/m^2	Non-fatal and fatal HF	20	594	**Normal weight:** Ref **Overweight**: 1.26 (1.05–1.50) **Obese**: 2.09 (1.68–2.59)
Kenchaiah, 2002[112]	USA	Framingham Heart Study	Men and women Age range: ≥30 years old	BMI	Per 3.2 kg/m^2	Non-fatal HF	14	496	Men: 1.05 (1.02–1.09) Women: 1.07 (1.04–1.10)
Wilhelmsen, 2001[113]	Sweden	Multifactor Primary Prevention Study	Men, 55–79 years old	BMI	Per 1 kg/m^2	Non-fatal and fatal HF	27	937	1.06 (1.03–1.09)
Meta-analysis									
Ayodipupo, 2023[75]	NA	NA	Population-based prospective cohort studies, Men and women, Age range ≥ 18	BMI; WHR	Per 5 kg/m^2 Per 0.1-unit higher level	Non-fatal and fatal HF	NA	34,422	BMI: 1.42 (1.40–1.44) WHR: 1.33 (1.28–1.37)
Huxley, 2014[79]	Multiple countries	APSC	Population-based and occupational settings	BMI	Normal weight: 18.5–21.9 kg/m^2	Fatal HF	5.1	495	<18.5: 1.72 (1.37-2.16) **18.5–21.9: 1.00 (0.84-1.19)** 22–24.9: 0.83 (0.68-1.01) 25–29.9: 1.02 (0.83-1.27) ≥30: 1.76 (1.22-2.54)

† All included studies were population-based prospective cohort studies.
APCSC = Asia-Pacific Cohort Studies Collaboration, BF = Body fat; BMI = Body mass index; HF = Heart failure; IHD = ischaemic heart disease; CHS = Cardiovascular Health Study; CVD = Cardiovascular diseases; NA = Not applicable; NR = Not reported; WHR = Waist-to-hip ratio

Uncertainties persist about the shape of the association of BMI with risk of HF. The summary level meta-analysis showed a positive log-linear association between BMI and HF risk for elevated BMI levels (≥25 kg/m^2); there was no evidence of an association between BMI and HF risk at lower BMI levels[75]. Similar findings were reported by Xing et al., 2023 and Ndumele et al., 2016 using data from the Atherosclerosis Risk in Communities (ARIC) study, including 13,730 participants with at least normal weight (BMI≥18.5 kg/m^2) and no CVD at baseline. However, studies have also shown a U-shaped association between BMI and HF risk, with the highest and lowest levels of BMI conferring the greatest risks of HF[102]. In a cohort of women aged 18–45 years with no history of HF at baseline, Björck et al., 2020 showed that compared to participants with BMI levels between 20 and 22.4 kg/m^2, those with BMI levels in the lowest range (<18.5 kg/m^2) and in the highest range (≥ 35 kg/m^2) had a 1.5-fold and 4.5-fold higher risk of non-fatal HF[102].

The two studies that evaluated the association between BMI and HF in non-Western populations showed an inverse relationship between BMI and risk of fatal HF[79,96]. Using data from the Japan Collaborative Cohort (JACC) study, a population-based prospective cohort study of 61,571 participants without kidney disease, liver disease, lung disease, CVD or cancer from 45 communities in Japan, Cui et al., 2014 demonstrated an inverse association between BMI and risk of fatal HF[96]. The APSC reported a similar inverse association between BMI and risk of fatal HF[79]. These non-linear associations are possibly due to reverse causality resulting from poor health among those with the lowest levels of BMI. However, the extent to which the inverse association between BMI and HF risk in studies conducted in the Asia and Pacific regions is real or due to reverse causality requires further exploration. Observational studies investigating the association between BMI and risk of cause-specific mortality have also demonstrated J-shaped associations owing to the strong influence of reverse causality[81].

Oguntade et al., 2023 reported a stronger association between BMI and risk of HF in White (1.36, 1.33–1.39) compared to Black individuals (1.10, 1.03–1.19)[75]. However, this comparison did not include Asians due to insufficient data. That meta-analysis also reported stronger associations of BMI with HF in men (1.29, 1.25–1.32) than in women (1.48, 1.45–1.50) and in younger adults (1.45, 1.39–1.51) compared with older adults (1.30, 1.2 7–1.33)[75]. However, multiple individual studies found no evidence of effect modification between BMI and HF risk by sex[98–100,114], despite having sufficient cases to investigate interaction effects. This discrepant findings between the individual studies and the meta-analysis may reflect the ecological fallacy in the meta-analysis estimates[115]. In addition, in meta-analyses, the results from subgroup analyses are likely to be confounded because, contrary to a multivariable meta-regression, subgroup analysis does not account for confounding when investigating heterogeneity between study-specific estimates[116].

2.3.4.1.2. Waist-to-hip ratio and heart failure

Five eligible studies, including a summary-level meta-analysis, reported on the association of WHR with risk of HF and were eligible for this review (**Table 2.3**). All studies were conducted in Western countries. The included studies mostly reported a positive association of WHR with HF risk[75,98,100,101], with evidence of weaker associations in older participants compared to younger participants[75,98] and in men compared to women[75,98]. In contrast, other studies found no evidence of effect modification by age[101] and sex[75,99,101].

The shape of the association between WHR and HF risk is broadly similar to that between BMI and HF. A summary-level meta-analysis of nine cohort studies, including 6,175 HF cases, reported a positive log-linear association of WHR with HF with a threshold of 0.9, below which there was no evidence of an association between WHR and HF risk[75]. Studies using individual participant data have reported similar findings

among participants with no history of cardiovascular disease[97,99]. This shape persisted despite efforts to minimise reverse causality, such as excluding subjects who died within the first few years of follow-up[99]

2.3.4.2. Pulmonary heart disease

Despite the wealth of evidence on BMI and WHR with HF risk, especially in Western countries, the search for this review did not retrieve any study reporting on the association of these exposures with PHD.

2.4. Limitations in the available evidence and implications for future studies

Most of the epidemiology of BMI and WHR with HF have been conducted in Western countries, with very little evidence from studies in non-Western populations. Studies evaluating these associations are needed in China, which has a lower mean BMI than Western populations. More studies are required to understand how the shape of the association of adiposity, especially BMI, with the risk of HF differ between populations and evaluate the extent to which the differences in the shape of associations are influenced by reverse causality bias. Such information is needed to inform public health policies and interventions to prevent HF in these populations.

2.5. Summary of the findings and key research gaps

Overall, there is reliable evidence of positive associations of SBP and adiposity (BMI and WHR) with risks of HF, but most of the evidence has been collected from Western countries. There is a need to investigate these associations in non-Western populations, including China, where the population is leaner and has a higher prevalence of hypertension with low awareness, treatment and control rates for hypertension than in Western populations.

Uncertainties persist about the shape of the association of BMI and WHR with HF. Studies from Western countries have reported a positive log-linear association between

BMI and HF at BMI levels ≥25 kg/m^2, whereas studies from non-Western populations have demonstrated an inverse association[75,79]. Understanding the extent to which the differences in the shapes of these associations reflect true ethnic differences or reverse causality is important to inform policymakers on the levels of these risk factors that put individuals at risk of HF to guide relevant interventions. In addition, reliable evidence on the shapes of these associations will inform randomised trials and clinical practices on the levels at which lowering the levels of a risk factor may reduce or increase HF risk.

In contrast to HF, there is little or no evidence of the risk factors for PHD, even in high-income countries.

2.6. Aims and objectives of the thesis

This thesis aimed to evaluate the incidence and mortality rates of HF and PHD in the 10-year follow-up of 0.5 million adults in the China Kadoorie Biobank (CKB). In addition, the thesis used observational epidemiological and Mendelian randomisation approaches to examine the independence and causal relevance of SBP, BMI and WHR as potentially modifiable risk factors for HF and PHD in order to guide future strategies for prevention of these diseases.

The specific objectives were to:

1. Estimate the incidence and mortality rates of HF and PHD overall and by subgroup.
2. Investigate the shape and strength of the associations of SBP, BMI and WHR with HF and PHD.
3. Assess the extent to which these associations are explained by confounding, intermediate factors or reverse causality bias.
4. Investigate the causal relevance of the most promising associations from the observational analyses using Mendelian randomisation approaches.

3. Study methods and procedures

This thesis used data from the China Kadoorie Biobank (CKB) study (data release version 18.0) to quantify the incidence and mortality rates of HF and PHD and investigate the independence and causal relevance of the associations of SBP, BMI and WHR with HF and PHD using conventional observational epidemiological and Mendelian randomisation approaches. This chapter summarises the (1) design of the CKB study, (2) assessment of the main exposures, covariates and outcomes, (3) statistical methods and (4) characteristics of the study participants.

3.1. Study design, setting and population

The CKB study was approved by the Oxford Tropical Research Ethics Committee at the University of Oxford (approval number: 025-04, 3 February 2005) and the Ethics Review Committee of the Chinese Centre for Disease Control and Prevention (approval number: 005/2004, 9 July 2004).

The CKB is an ongoing open-ended population-based prospective cohort study of 512,722 Chinese adults aged 30–74 years when recruited at baseline. Participants were recruited from ten geographically diverse regions of China. Within each region, study sites were purposely selected based on population stability, social diversity, differences in risk factor exposure, major disease burden and the quality of local morbidity and mortality registers[117]. All permanent residents within each of 100–150 administrative units were identified through official residential records and invited to participate in the study[117]. Health workers and local community leaders invited participants by distributing invitation letters and study information leaflets from door to door, with a response rate of ~30%[117,118].

3.2. Data collection procedure

3.2.1. Baseline survey

The baseline assessment for the study was conducted by medically trained staff with relevant field experience in temporary assessment clinics from 2004–2008[117]. The baseline assessments involved data collection using interview-administered questionnaires and physical measurements. Blood samples were collected for extraction of buffy coats (for deoxyribonucleic acid [DNA] extraction), measurement of plasma glucose concentrations and long-term storage. A typical baseline visit took about 60–75 minutes to complete the questionnaire, clinical measurements and blood collection.

Before data collection, staff at the regional coordination centres received relevant training on the study rationale, methods required to conduct the survey, including use and maintenance of study equipment, and optimum approaches for communication with study participants. The materials and equipment needed for data collection were provided centrally. A pilot study was conducted to test the standard operating procedures. The staff commenced data collection immediately after their training, and data collection was supervised for one week to ensure adherence to the data collection procedures. All staff members had a unique identification number that was used as a login prior to data entry, which permitted evaluation of completeness and quality of data collected by individual staff members to minimise data entry errors.

At the assessment clinics, participants provided informed consent, which permitted the storage of their blood samples and access to their health records for anonymised and unspecified research[117]. The health workers then used electronic questionnaires to collect data on self-reported exposures such as socioeconomic and demographic status (e.g., household income and level of education), lifestyle (e.g., alcohol consumption and smoking) and medical history (e.g., stroke, cancer and COPD). Missing or implausible responses, inconsistencies and logic errors were detected using software and the

interviewer was prompted to rectify these issues. After collecting self-reported exposures, the physical measures of the study participants, including weight, height, waist circumference, BP and RHR, were measured following standard operating procedures[117].

All participants had 10 ml of venous blood collected in the non-fasting state into tubes containing Ethylenediaminetetraacetate (BD Hemogard™, BD, Franklin Lakes, New Jersey, USA), except for participants in Zhejiang who were asked to fast prior to their baseline assessment visit. All participants had duration of time (fasting time) since their last meal recorded. A SureStep Plus desktop device (LifeScan, Milpitas, CA, SA) was used to measure random plasma glucose levels on the day of the examination. Participants with a random plasma glucose between 7.8 and 11.1 mmol/L were invited for measurement for fasting plasma glucose on the following day. The remaining blood samples were preserved in ice packs at 0–4 ºC and transported to the local laboratories within a few hours for centrifugation and aliquoting into four cryovials, one containing a buffy coat with DNA. The plasma aliquots and buffy coat samples were then stored for 3–4 months at -40 ºC before transportation in dry ice for long-term storage at -80 ºC in Beijing. Two cryovials for each participant were also transported to Oxford for long-term storage in liquid nitrogen.

3.2.2. Variables

3.2.2.1. Main exposures

3.2.2.1.1. Blood pressure

BP (measured to the nearest one mmHg) was measured twice for each participant using a regularly calibrated UA-779 digital BP monitor after at least five minutes of rest[118]. A third BP measurement was taken if there was a difference of 10 mmHg between the first two measurements. The last two BP measurements were averaged, recorded and used for analysis.

3.2.2.1.2. BMI and WHR

Anthropometric measurements were taken with participants in light clothing and no shoes[117]. A stadiometer was used to measure standing and sitting heights to the nearest 0.1 cm. A body composition analyser (TANITA-TBF-300GS; Tanita) was used to measure weight. The weight of light clothing was subtracted from measurements of body weight, depending on the season of measurement: 0.5 kg in summer and 2–2.5 kg in winter. Non-stretchable tapes were used to measure waist circumference and hip circumference to the nearest one cm. Waist circumference was measured midway between the lowest rib and iliac crest, whereas hip circumference was measured at the maximum distance around the buttocks. BMI was calculated as weight (in kg) divided by the square of the height (in metres). WHR was computed as the ratio of the waist and hip circumference.

3.2.2.2. Covariates

3.2.2.2.1. Selection of covariates

The following variables were considered as covariates based on the literature review in Chapter 2: demographic and socioeconomic status (age, sex, study area, income, education and occupation), lifestyle factors (alcohol consumption, total physical activity and sedentary time), physical measurements (waist circumference and heart rate), medication use (use of BP-lowering medications) and prevalent conditions (COPD, emphysema or bronchitis, asthma, tuberculosis and self-rated health). Outdoor ambient temperature was included as a covariate to investigate possible confounding on the associations of SBP and adiposity with HF and PHD[119]. An additional variable, age-at-risk, was derived for the prospective cohort analyses. Age-at-risk corresponds to the age at first diagnosis for HF or PHD and accounts for the excess risk of disease attributed to attained age during study follow-up[120]. Marital status was not used as a covariate, as over 90% of CKB participants were married.

3.2.2.2.2. Definition and categorisation of covariates

Age at baseline was categorised into five groups with cut-offs at 39, 49, 59 and 69. Age-at-risk was categorised into 5-year age bands with cut-offs at 35, 40, 45, 50, 55, 60, 65, 70, 75. The highest level of education attained by the participants was categorised as no formal, primary, middle school and high school education or higher (i.e., university or technical college). Annual household income was categorised into six categories: <2,500, 2,500–4,999, 5,000–9,999, 10,000–19,999, 20,000–34,999 and ≥35,000 Yuan. Occupation status was categorised as agriculture and related works and other works, which included administrative or managerial work, professional or technical work, sales and service work, housewife or househusband and participants on retirement.

Regarding lifestyle factors, alcohol consumption was categorised according to the frequency at which participants consumed alcohol in the previous 12 months. Individuals who had never drunk weekly were classified as never regular drinkers; ex-regular drinkers were those who had not drunk in the 12 months preceding baseline assessment but drank alcohol weekly in the past; reduced-intake drinkers included those who used to drink weekly but reported drinking occasionally in the previous 12 months; occasional drinkers had never drunk weekly in the past and reported to have drunk occasionally in the past 12 months; current weekly drinkers were participants who drank alcohol most weeks in the last 12 months. Alcohol intake was categorised into never, ex- (ex-regular drinking and reduced intake), occasional and current drinking.

Participants were classified as never- and ever-regular smokers. Never-regular smokers included participants who smoked less than 100 cigarettes in their lifetime. Ever-regular smokers were participants who smoked at least one cigarette (or ≥1g of tobacco) per day for six successive months. Ever-regular smokers included current and ex-smokers. Ex-smokers included those who had stopped smoking about six months before the baseline assessment.

Total physical activity (TPA) was defined by summing the amount of daily physical activity over the preceding 12 months across four domains: occupational, leisure-time, housework and commuting exercise. Each type of physical activity was quantified using the metabolic equivalent of tasks (MET)[121]: the ratio of energy expended for each kilogram of body weight when performing physical activity to the energy used when the body is at rest. Each domain of physical activity was then expressed in MET-hour/day by multiplying its MET value by the average duration spent doing that activity per day. TPA in MET-hour/day was obtained by summing the value in MET-hour/day for all physical activity domains[122,123]. In addition to TPA, sedentary leisure time was defined by summing the hours spent on leisure-time sedentary activities per day. Sedentary activities included reading, playing mah-jong or cards, watching television and using a computer outside of work.

Doctor-diagnosed diabetes was considered as a "yes" response to the question: *"Has a doctor ever told you that you had diabetes?"* Participants with a history of diabetes were asked further questions including the age at diagnosis and current use of medications to treat diabetes and CVDs. For participants without a previous diagnosis of diabetes, diabetes was diagnosed using their blood glucose level, referred to as "screen-detected" diabetes. Screen-detected diabetes was defined as random plasma glucose ≥7.0 mmol/L with fasting time ≥8 hours, random plasma glucose ≥11.1 mmol/L with fasting time <8 hours or fasting plasma glucose ≥7.0 mmol/L. Prevalent diabetes included diagnosed- and screen-detected diabetes.

3.3. Resurvey sub-study

After baseline assessment, the first, second and third resurvey sub-studies were conducted in 2008, 2013–14 and 2020–21 to assess the validity of baseline measurements, collect additional exposures that were missed in previous surveys and permit the correction of bias[124]. The resurveys included repeat measurements of

questionnaires and clinical measurements (plus some enhancements) in random samples of ~5% of surviving study participants recruited at baseline using identical procedures to those used in the baseline assessment[118]. Resurvey data were limited to those collected at the first and second resurveys for the present thesis. A total of 24,740 and 33,656 participants were invited for the first and second resurvey sub-studies, with response rates of 80% and 75%, respectively.

3.4. Array design and genotyping

Genotyping in the CKB was performed using a custom-designed 800K single nucleotide polymorphism (SNP) variation array on the Affymetrix Axiom platform, with imputation to the 1,000 Genomes Project Phase 3[125]. The CKB array design used the UK Biobank probe list as the starting point. The list was then modified to optimise the selection of SNPs for East Asians. Based on allele frequency and sequencing data for over 12,000 individuals of East Asian ancestry, the array was designed to maximise genome-wide coverage of low-frequency (1–5%) and common (>5%) variants among individual across China by removing variants that are absent or exist at low frequency in East Asians[125]. In addition, the array included specific variants, such as rare missense, loss-of-function and expression quantitative trait loci variants, found in Chinese populations. Finally, an Asian-specific genome-wide grid was developed to maximise the imputation of rare and common variants. Based on the from 100 plates including ~9,000 samples after quality control, the initial CKB array design was revised to: (1) remove poor-quality, failed or uninformative monomorphic probe, (2) add tags for variants that failed quality control, and (3) improve or restore genome-wide imputation coverage by adding variants[125].

A subset of CKB participants were selected for genotyping, including cases with subarachnoid haemorrhage, intracerebral haemorrhage and fatal ischaemic heart disease; randomly selected cases of myocardial infarction, ischaemic stroke and COPD; controls of participants with no CVD at the time of participant selection matched by age,

sex and study area; and a random selection who participated in the second resurvey. Additional genotyped samples came from randomly selected DNA samples or samples from the assessment centres during the second resurvey; these samples were selected to be representative of the CKB population.

Of the 105,408 CKB participants selected for genotyping, samples were excluded based on the following criteria: low call rates (<95%), genotypic and self-reported sex did not match, ancestry outliers, bad linkage, XY aneuploidy, high heterozygosity (3 SD above the mean) and withdrawn or missing consent[125]. These exclusions left 100,706 participants with eligible genotyping data.

Principal component analysis based on 77,000 unrelated individuals was performed to investigate the population structure of the CKB[125]. The first 11 principal components were found to predict participants' study area.

3.5. Follow-up for morbidity and mortality

After data collection, the study participants were followed up for hospital admissions and death through linkage to electronic health records using a unique personal identification number. Hospital admissions were identified by electronic linkage to the national health insurance claims database for all hospitalisations, and study records were updated every six months. Data on participants' vital status were retrieved through linkage to death registries at China's Disease Surveillance Points (DSP) system and supplemented by annual active follow-up[126]. Information on the causes of death was mostly retrieved from official death certificates and supplemented by medical records. In the small percentage (<5%) of deaths that occurred without prior medical consultation, verbal autopsy was used to ascertain the cause of death.

All fatal and non-fatal diseases were coded using the International Classification of Diseases tenth revision (ICD-10) by trained staff blinded to the baseline information of study participants (**Table 3.1**).

Table 3.1. ICD-10 codes used to define heart failure and pulmonary heart disease.

Disease	ICD-10	Code description
Heart failure (I50)	I50.0	Congestive heart failure
	I50.1	Left ventricular failure
	I50.9	Heart failure, unspecified
Pulmonary heart disease (I27)	I27.0	Primary pulmonary hypertension
	I27.20	Pulmonary hypertension, unspecified
	I27.21	Secondary pulmonary arterial hypertension
	I27.81	Cor pulmonale (chronic)
	I27.89	Other specified pulmonary heart diseases
	I27.9	Pulmonary heart disease, unspecified

3.6. Statistical methods

3.6.1. Data exploration, cleaning and manipulation

The distribution of continuous variables, as well as the presence of outliers and implausible values, were assessed using histograms, boxplots and value ranges of these variables. In addition, variables were assessed for missing observations. Missing values for BMI at baseline (n=2) were imputed with the mean BMI of the study population.

3.6.2. Descriptive analysis of baseline data

The baseline characteristics of the study participants were stratified by sex to assess differences in the distribution of baseline characteristics. Categorical and quantitative variables were summarised using percentages and mean (and standard deviation [SD]). Percentages and means were directly standardised to the baseline age and all 10 study areas structure of the CKB population. The mean and percentages for age were

standardised by study area, whereas the percentages for study areas were standardised by age at baseline.

The magnitude of correlations between SBP, BMI, WHR, waist circumference, body fat percentage and RHR was assessed using overall and partial Pearson correlation coefficients. Overall correlation coefficients were estimated for all study participants after adjusting for age at baseline (five-year bands) and sex. Sex-specific partial correlation coefficients were adjusted for age (five-year bands). Correlation coefficients of 0–0.19, 0.2–0.39, 0.4–0.59, 0.6–0.79 and 0.8–1 were interpreted as very weak, weak, moderate, strong and very strong correlations, respectively[127].

3.6.3. Analysis of resurvey data

3.6.3.1. Descriptive analysis

The baseline characteristics of participants in the first and second resurvey data were adjusted by direct standardisation by age at baseline, sex and all 10 study areas. Categorical and quantitative data were summarised using percentage and mean, respectively.

3.6.3.2. Regression dilution ratios for SBP, WHR and BMI

Prospective cohort studies investigating the association between long-term or “usual” levels of a continuous risk factor and an outcome are often based on values of the exposures measured at an initial (or “baseline”) survey. However, it is often impossible to precisely measure the underlying value of such risk factors chiefly because of short-term biological variability, long-term within-person variability and measurement error. Using the baseline values of this risk factor measured with random error in regression analyses systematically underestimates the true risk of disease associated with long-term or usual levels of the risk factor[128]. This systematic underestimation of the long-term effect of a risk factor on an outcome is called regression dilution bias (RDB), and

the magnitude of this attenuation is the regression dilution ratio (RDR). RDB can be corrected by dividing the (biased) regression coefficient relating baseline levels of the risk factor with the outcome by the RDR[128].

Non-parametric and parametric methods can be used to calculate the RDR using measurements of the exposure at baseline and a resurvey sub-study conducted at the mid-point of follow-up[128]. In this thesis, the RDRs for exposures assumed to have an underlying normal distribution, including SBP and BMI, were calculated using Rosner's regression method (a parametric method). In contrast, the assumption-free MacMahon-Peto's ratio of ranges method was used to calculate the RDR for non-normally distributed exposures, i.e., WHR. The first resurvey (conducted in 2008) and second resurvey (conducted in 2013–2014) sub-studies were not conducted at the mid-point of follow-up: the mid-point of follow-up should be 2011 since the end of follow-up for the version of the CKB data used for this thesis was 2018. However, the values of the risk factors at the mid-point of follow-up were estimated by averaging values for the first and second resurveys.

The MacMahon-Peto method to quantify the regression dilution ratio (RDR) uses mean values at resurvey in individuals classified at baseline and provides an unbiased estimate of usual values[128–130]. To calculate the RDR using MacMahon-Peto's method in this thesis, participants recruited at baseline and resurvey were divided into quintile groups (or fifths) based on their baseline measurements. The quintile-specific mean values of the risk factor were then calculated at baseline and mid-point of follow-up (estimated by averaging values of the risk factor at the first and second resurveys). The RDR was estimated as the ratio of the range of the means in the top and bottom fifths in the baseline survey to the range of the means in the top and bottom fifths at resurvey at the mid-point of the duration of follow-up.

Rosner's regression method is a parametric method and assumes that the risk factor has an underlying normal distribution[128]. The method is also based on the fact that the resurvey measure is an unbiased estimate of the baseline measurement, conditionally on the baseline measurement. Rosner's method is preferred over the MacMachon Peto method for parametric data because it uses all values of a variable to estimate the RDR (rather than grouping the data) and can be extended to adjust for other covariates[128]. In this thesis, the RDR was calculated using this method by first regressing the estimated measurement of the risk factor at the mid-point of follow-up on the baseline measurement. The RDR was then the reciprocal of the regression coefficient. The variance of the RDR was calculated as follows[128]:

$$var\ (RDR) = \frac{RDR^2\ (RDR^2 - 1)}{n}$$

Where *n* is the sample size and the *RDR* = 1/(regression coefficient).

The methods to correction for RDB in prospective associations are described in **Chapters 5 and 6**.

3.7. Results

3.7.1. Baseline characteristics of all CKB participants

Table 3.2 shows the characteristics of all CKB participants recruited at baseline and by sex. The mean (SD) age of the 517,724 participants recruited at baseline was 52.0 (10.7) years and 59% were women. Most participants resided in a rural area (55.9%), were married (90.6%) and were agricultural or factory workers (55.8%). In addition, 67.8% and 28.2% of the participants had a middle school education or higher and a household monthly income of <10,000 Yuan. On average, men (52.9 years) were about a year older than women (51.4 years). Compared to women, men were more likely to be married (93.3 vs. 88.4%) and agricultural or factory workers (62.3 vs. 50.7%). However, men were less likely to have household monthly incomes of <10,000 Yuan compared to women (**Table 3.2**).

About 38.1% and 14.9% of all CKB participants were ever-regular smokers and current drinkers. Compared to men, women almost did not smoke or drink; 85.6% and 33.4% of the men were ever-regular smokers and current drinkers compared to 5.2% and 2.1% of the women, respectively (**Table 3.2**). Men had higher levels of physical activity than women (22.4 vs. 20.1 MET-hour/day).

The participants had normal mean (SD) levels of BMI [23.7 (3.4) kg/m^2], waist circumference [80.3 (9.8) cm] and DBP [77.8 (11.2) mmHg] but high-normal SBP [131.1 (21.3) mmHg]. Compared to women, men had higher mean levels of SBP (132.3 vs. 130.5 mmHg) and waist circumference (82.1 vs. 79.1 cm) but a lower body fat percentage (22.1 vs. 32.1%) (**Table 3.2**).

About 35% of all the participants had hypertension and 10.4% reported having poor health. Less than 10% of the participants had a history of diabetes, COPD, emphysema, asthma or bronchitis. Compared to women, men had a higher prevalence of

hypertension (men vs. women: 36.3 vs. 34.7%) and COPD (8.2 vs. 6.4%). However, men were less likely to have a history of diabetes (5.5 vs. 6.3%), report shortness of breath on walking (4.5 vs. 6.9%) and rate their health poorly (8.8 vs. 11.5%). About 12% of CKB were on BP-lowering medications, and intake of BP-lowering medications was similar for men and women, **Table 3.2**.

Table 3.2. Characteristics of study participants by sex at baseline (2004–2008).

Characteristics	Sex		All
	Men	Women	
No. of participants	210,205	302,519	512,724
Demographic and SES factors, %			
Age, mean (SD), years	52.9 (10.9)	51.4 (10.4)	52.0 (10.7)
Age group			
30–39	14.1	15.9	15.1
40–49	28.0	31.0	29.8
50–59	30.3	31.0	30.7
60–69	19.8	16.6	17.9
70–79	7.8	5.5	6.4
Living in rural area	56.7	55.1	55.9
Ambient outdoor temperature, mean (SD), °C	15.9 (9.2)	16.3 (9.0)	16.1 (10.1)
Married	93.3	88.4	90.6
Middle school education or higher	67.7	68.4	67.8
Household monthly income <10,000 Yuan	26.0	29.8	28.2
Agriculture or factory worker	62.3	50.7	55.8
Lifestyle factors, %			
Ever regular smoker	85.6	5.2	38.1
Current drinkers	33.4	2.1	14.9
Total physical activity, mean (SD), MET-hr/d	22.4 (13.4)	20.1 (10.4)	21.1 (13.9)
Sedentary time, mean (SD), hours/day	3.1 (1.4)	2.9 (1.4)	3.0 (1.5)
Physical measurements, mean (SD)			
BMI, kg/m²	23.5 (3.1)	23.8 (3.3)	23.7 (3.4)
SBP, mean (SD), mmHg	132.3 (18.8)	130.5 (20.2)	131.1 (21.3)
DBP mean (SD), mmHg	79.2 (11.1)	76.8 (10.7)	77.8 (11.2)
Resting heart rate, bpm	77.8 (12.1)	79.7 (11.5)	78.9 (11.8)
Waist circumference, cm	82.1 (9.2)	79.1 (9.1)	80.3 (9.8)
Body fat percentage, %	22.1 (5.9)	32.1 (6.9)	27.9 (8.4)
Prevalent conditions, %			
Hypertension	36.3	34.7	35.2
Taking BP-lowering drugs	10.2	12.4	11.5
Diabetes	5.5	6.3	5.9
COPD	8.2	6.4	7.2
Emphysema or bronchitis	2.9	2.3	2.6
Asthma	0.6	0.5	0.5
Shortness of breath on walking	4.5	6.9	5.9
Self-reported poor health	8.8	11.5	10.4

Means and proportions were directly standardised by age, sex, and all 10 CKB areas.
bpm = beats per minute; COPD = Chronic obstructive pulmonary disease; DBP = Diastolic blood pressure; hr/day = hour/day; MET = Metabolic Equivalent Task; SD = Standard deviation; SOB = Shortness of breath

3.7.2. Baseline characteristics by study area

Table 3.3 summarises the selected baseline characteristics of the participants by the 10 study areas. There were substantial variations in the participant characteristics across the 10 study areas. The mean (SD) age of the participants ranged from 49.4 (10.7) years in rural Gansu to 54.3 (10.5) years in urban Liuzhou. Mean levels of education were lower in rural than urban areas. Participants from urban study areas were twice as likely to have completed middle school education than those from rural areas. Overall, the proportion of individuals with completed middle school education varied from 25.1% in rural Gansu to 88.3% in urban Harbin. Likewise, there were substantial differences in mean ambient outdoor temperatures across the study areas that varied from 7.3ºC in Harbin to 24.5ºC in Haikou (**Table 3.3**).

On average, participants from rural study areas had about 5 mmHg higher mean SBP levels than those in urban areas, and the mean SBP levels varied from 123.7 (19.2) mmHg in urban Haikou to 135.3 (22.2) mmHg in rural Zheijang. Similar variations across study areas were observed among participants' BMI, which ranged from 22.4 (3.2) kg/m^2 in rural Hunan to 25.7 (3.5) kg/m^2. The prevalence of diabetes was about 2-fold greater in the urban areas compared to the rural areas, with the area-specific prevalence ranging from 3.5% in rural Hunan to 10.4% in urban Qingdao. The percentage of individuals taking BP-lowering medications varied considerably across the study areas, with no obvious rural-urban difference. The intake of BP-lowering medication was below average in 60% (3/5) of the rural areas compared to 20% (1/5) of the urban areas.

Table 3.3. Selected characteristics of study participants by study area at baseline (2004–2008).

Study area	N	Age, years†	Women††	Middle school or higher††	SBP, mmHg†	BMI, kg/m²†	WHR†	Temperature, ºC†	Diabetes††	Taking BP-lowering drugs††
Rural										
Gansu	49887	49.4 (10.7)	59.7	74.4	133.5 (23.2)	22.7 (3.2)	0.9 (0.1)	12.1 (9.0)	3.9	6.9
Sichuan	55686	51.6 (10.5)	61.5	50.8	129.6 (18.2)	23.3 (3.2)	0.9 (0.1)	17.2 (7.3)	4.1	3.2
Zhejiang	57704	52.8 (9.9)	58.4	63.6	135.3 (22.2)	22.9 (3.6)	0.9 (0.1)	17.7 (9.7)	4.9	15.8
Hunan	59900	52.0 (10.5)	56.0	41.4	131.6 (20.6)	22.4 (3.2)	0.9 (0.1)	18.0 (8.6)	3.5	10.3
Henan	63356	50.9 (10.4)	55.7	63.4	134.9 (20.7)	24.3 (3.5)	0.9 (0.1)	15.4 (10.1)	5.1	14.1
Urban										
Haikou	29685	53.2 (11.7)	64.0	81.2	123.7 (19.2)	23.4 (3.3)	0.9 (0.1)	24.5 (4.5)	6.6	7.6
Qingdao	35508	50.8 (10.2)	56.7	80.2	132.5 (20.8)	25.7 (3.5)	0.9 (0.1)	13.7 (8.2)	10.4	13.7
Liuzhou	50174	54.3 (10.5)	61.9	82.0	126.7 (18.9)	23.7 (3.4)	0.9 (0.1)	21.6 (7.4)	7.5	12.4
Suzhou	53269	52.1 (10.3)	58.1	68.0	132.7 (19.1)	24.0 (3.2)	0.9 (0.1)	17.4 (8.6)	5.3	16.5
Harbin	57555	53.4 (11.4)	60.2	91.4	127.0 (19.2)	24.5 (3.4)	0.9 (0.1)	7.3 (14.5)	9.7	11.6
All	512,724	52.0 (10.7)	59.0	67.8	131.1 (21.3)	23.7 (3.4)	0.9 (0.1)	16.1 (10.1)	5.9	11.5

†Summarised as mean (standard deviation [SD]);

††Summarised as percentage;

Means and proportions were directly standardised by age and sex;

BMI=Body mass index; SBP=Systolic blood pressure; WHR=Waist-to-hip ratio

3.7.3. Correlation of baseline SBP, BMI and WHR with selected characteristics

The partial correlations of baseline SBP, BMI and WHR with selected baseline characteristics for all participants and by sex are summarised in **Table 3.4**. The correlations were similar for both men and women. SBP was weakly correlated with BMI, WHR and waist circumference and very weakly correlated with fat percentage and heart rate. There were moderate to strong correlations between BMI and other adiposity indices. BMI was strongly correlated with waist circumference and body fat percentage but moderately correlated with WHR. However, BMI had almost no correlation with heart rate. WHR was strongly correlated with waist circumference but only moderately correlated with BMI and weakly correlated with body fat percentage.

Table 3.4. Correlation coefficients between systolic blood pressure, adiposity traits and resting heart rate.

Variables		BMI	Waist circumference	Waist-to-hip ratio	Fat percentage	Heart rate
SBP	**All:**	0.25	0.23	0.19	0.26	0.14
	Men:	0.25	0.23	0.20	0.18	0.14
	Women:	0.25	0.23	0.20	0.18	0.14
BMI		**All:**	0.85	0.56	0.86	0.03
		Men:	0.83	0.52	0.73	0.03
		Women:	0.83	0.52	0.73	0.03
Waist circumference			**All:**	0.78	0.77	0.06
			Men:	0.78	0.53	0.05
			Women:	0.78	0.53	0.05
Waist-to-hip ratio				**All:**	0.54	0.10
				Men:	0.27	0.08
				Women:	0.27	0.08
Fat percentage					**All:**	0.05
					Men:	0.08
					Women:	0.08

Partial correlation coefficients were adjusted for age and sex.

3.7.4. Baseline characteristics of the resurveyed participants

Of the 512,724 participants recruited at baseline, 19,786 (3.9%) and 25,239 (4.9%) participated in the first and second resurveys. The characteristics of the participants recruited in the first and second resurvey sub-studies were broadly similar to the CKB population at baseline, except for the proportion of ever-regular smokers that were lower in the first (36.6%) and second (35.6%) resurveys than at baseline (38.1%), **Table 3.5**.

Table 3.5. Baseline characteristics of resurveyed participants.

Characteristics	Baseline	First resurvey	Second resurvey
No. of participants	512,724	19,786	25,239
Demographic and SES factors, %			
Age, mean (SD), years	52.0 (10.7)	52.0 (10.6)	51.5 (10.2)
Living in a rural area	55.9	58.8	56.7
Ambient outdoor temperature, mean (SD), °C	16.1 (10.1)	15.0 (11.5)	16.5 (10.4)
Married	90.6	91.3	92.1
No formal education	18.6	19.9	18.8
Household monthly income <10,000 Yuan	28.2	32.5	31.6
Agriculture or factory worker	55.8	55.3	54.6
Lifestyle factors, %			
Ever regular smoker	38.1	36.6	35.6
Current drinkers	14.9	14.5	14.6
Total physical activity, mean (SD), MET-hour/day	21.1 (13.9)	20.7 (14.2)	21.4 (14.0)
Sedentary time, mean (SD), hours/day	3.0 (1.5)	3.0 (1.5)	3.0 (1.5)
Physical measurements, mean (SD)			
BMI, kg/m²	23.7 (3.4)	23.7 (3.4)	23.7 (3.4)
SBP, mmHg	131.1 (21.3)	132.2 (21.8)	131.0 (20.8)
DBP, mmHg	77.8 (11.2)	77.9 (11.3)	77.6 (11.0)
Resting heart rate, bpm	78.9 (11.8)	79.1 (11.8)	78.9 (11.7)
Waist circumference, cm	80.3 (9.8)	79.8 (9.7)	79.5 (9.5)
Waist-to-hip ratio	0.9 (0.1)	0.9 (0.1)	0.9 (0.1)
Body fat percentage, %	27.9 (8.4)	28.4 (8.5)	28.3 (8.3)
Prevalent conditions, %			
Hypertension	35.2	36.4	34.8
Taking anti-hypertensive drugs	11.5	10.6	10.1
Diabetes	5.9	5.5	4.7
COPD	7.2	7.3	7.1
Asthma	0.5	0.7	0.5
Self-reported poor health	10.4	10.5	9.9

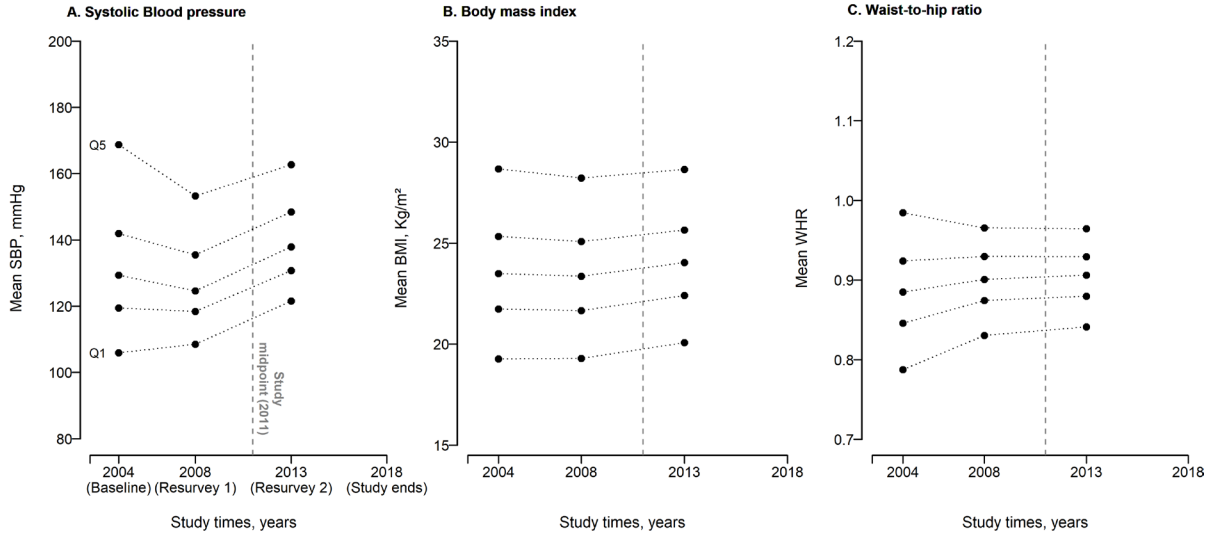
Means and proportions where directly standardised by the age, sex, and all 10 CKB areas.
bpm = beats per minute; COPD = Chronic obstructive pulmonary disease; DBP = Diastolic blood pressure; hr/day = hour/day; MET = Metabolic Equivalent Task; SD = Standard deviation; SOB = Shortness of breath

3.7.5. Regression dilution ratios for SBP, BMI and WHR

Table 3.6 shows the baseline-defined quintile-specific means of SBP, BMI and WHR of the participants at baseline and first and second resurveys. Overall, the range of quintile-specific SBP and WHR mean values reduced between the baseline and first and second resurveys, consistent with substantial regression to the mean (**Table 3.6**, **Figure 3.1**).

In contrast, the ranges of the quintile-specific BMI means remained fairly constant with time, indicating that the baseline BMI average in this population is an excellent estimate of long exposure. The mean times from baseline to the first and second resurveys were 2.6 years and 8.0 years, and these were similar for all baseline-defined groups. The RDRs calculated using Rosner's and Peto's methods were similar. The RDR for SBP, WHR and BMI were 0.61 (0.60–0.62), 0.61 (0.60–0.61) and 0.88 (0.88–0.89), respectively.

Figure 3.1. Regression dilution ratios for SBP, BMI and WHR.



The black dots are the quintile-specific means. Q=Quintile; Resurveys 1 & 2 = First and second resurveys.

Table 3.6. Quintile-specific means of SBP, BMI and WHR at baseline and first and second resurvey.

	Mean exposure values				Mean interval from baseline (years)				
Baseline quintiles	Baseline	First resurvey (RS 1)	Second resurvey (RS 2)	Average of RS 1 and 2 means	To RS 1 (T1)	To RS 2 (T2)	Average of mean T1 and T2	Rosner's RDR (SE)	MacMahon-Peto's RDR
Systolic blood pressure, mmHg									
Q1 (<113.5)	105.66	108.08	121.05	117.55	2.58	7.92	5.25	0.61 (0.014)	0.61
Q2 (114.0 to <123.5)	118.85	117.60	129.84	126.47	2.56	7.94	5.25		
Q3 (124.0 to <133.0)	128.24	123.38	136.59	132.84	2.61	7.95	5.28		
Q4 (133.5 to <149.0)	140.38	134.07	147.57	143.97	2.63	8.03	5.33		
Q5 (149.5+)	168.18	153.97	163.17	160.30	2.66	8.09	5.38		
Body mass index, kg/m^2									
Q1 (<20.9)	19.31	19.38	20.12	19.81	2.59	8.02	5.31	0.88 (0.004)	0.92
Q2 (20.6 to <22.7)	21.71	21.64	22.41	22.10	2.64	8.01	5.32		
Q3 (22.4 to <24.4)	23.45	23.32	24.01	23.73	2.64	8.01	5.32		
Q4 (24.2 to <26.5)	25.29	25.04	25.65	25.41	2.61	7.96	5.29		
Q5 (26.2+)	28.59	28.15	28.56	28.39	2.57	7.93	5.25		
Waist-to-hip ratio									
Q1 (<0.85)	0.79	0.83	0.84	0.84	2.62	8.05	5.34	0.61 (0.013)	0.65
Q2 (0.82 to <0.89)	0.85	0.88	0.88	0.88	2.61	8.00	5.30		
Q3 (0.86 to <0.92)	0.89	0.90	0.91	0.90	2.60	7.99	5.29		
Q4 (0.89 to <0.96)	0.92	0.93	0.93	0.93	2.58	7.94	5.26		
Q5 (0.93+)	0.98	0.96	0.96	0.96	2.64	7.95	5.29		

RDR=Regression dilution ratio; SE=Standard error; T1 and 2=Mean time intervals from baseline to first and second resurveys.

Table 3.7. Age-specific regression dilution ratios for men and women.

	SBP			BMI			WHR		
Subgroup	Rosner's RDR (95% CI)	MacMahon-Peto's RDR	n	Rosner's RDR (95% CI)	MacMahon-Peto's RDR	N	RDR (95% CI)	MacMahon-Peto's RDR	n
Men									
30–59	0.60 (0.55–0.65)	0.59	7135	0.88 (0.87–0.90)	0.94	8276	0.67 (0.63–0.70)	0.65	8275
60–69	0.62 (0.53–0.71)	0.51	1904	0.87 (0.84–0.90)	0.97	2030	0.69 (0.63–0.76)	0.68	2030
70+	0.56 (0.35–0.77)	0.53	596	0.86 (0.80–0.91)	0.96	603	0.71 (0.60–0.82)	0.68	603
Women									
30–59	0.66 (0.62–0.69)	0.62	10559	0.91 (0.90–0.92)	0.90	13771	0.57 (0.52–0.61)	0.65	13770
60–69	0.63 (0.56–0.71)	0.56	2525	0.90 (0.88–0.92)	0.94	2727	0.68 (0.62–0.74)	0.62	2727
70+	0.56 (0.35–0.76)	0.49	671	0.91 (0.87–0.95)	0.93	704	0.55 (0.35–0.75)	0.62	704
All	0.61 (0.58–0.64)	0.61	23,390	0.88 (0.87–0.89)	0.92	28,111	0.61 (0.58–0.64)	0.65	28,109

BMI: Body mass index; CI: Confidence interval; n: number of participants; RDR: Regression dilution ratio; SBP: Systolic blood pressure; WHR: Waist-to-hip ratio.

Table 3.7 shows the age-specific RDRs for men and women. There were some variations in the age-specific Rosner's RDR. For SBP, the ratios were smaller in men and at older ages, suggesting greater variability in baseline BMI values for men than in women and for older participants. The smaller ratios among older participants could reflect random error due to the relatively small sample size of this group. For BMI, the ratios were smaller in men, but there was no age gradient. Contrary to the findings for BMI, the ratios for WHR were smallest in women.

3.7.6. Heart failure and pulmonary heart disease cases by ICD-10 codes

Table 3.8 shows the number of HF and PHD cases by ICD-10 subcategories. A total of 8223 HF and 6404 PHD cases were reported by December 2018 (the study censoring date). Unspecified HF (n=4636, 56.4%) was the most predominant HF category, followed by left ventricular HF (2756, 33.5%) and congestive HF (831, 10.1%). Cases of unspecified PHD (5649, 88.2%), chronic cor pulmonale (430, 6.7%), and other specified PHD (182, 2.8%) were the most reported PHD subcategories (**Table 3.8**).

Table 3.8. Heart failure and pulmonary heart disease events by ICD-10 code.

Disease	Code description (ICD-10)	No. of cases
Heart failure (I50)	Congestive heart failure (I50.0)	831
	Left ventricular failure (I50.1)	2756
	Heart failure, unspecified (I50.9)	4636
	All cases	**8223**
Pulmonary heart disease (I27)	Primary pulmonary hypertension (I27.0)	24
	Pulmonary hypertension, unspecified (I27.20)	39
	Secondary pulmonary arterial hypertension (I27.21)	80
	Cor pulmonale (chronic) (I27.81)	430
	Other specified pulmonary heart diseases (I27.89)	182
	Pulmonary heart disease, unspecified (I27.9)	5649
	All cases	**6404**

3.8. Summary

The CKB study is an open-ended study that permits the collection of a broad range of disease outcomes by electronic linkage of participants to electronic health records. The open-ended design of this study, including the collection of a broad range of exposures and DNA samples, now permits the generation and testing of a wide range of hypotheses. Being a large-scale prospective cohort study with careful measurement of exposure at baseline and resurvey makes CKB a unique resource for investigating epidemiological associations and minimising known sources of bias. The availability of genotyping data also helps to investigate the causal relevance of any observational associations. In addition to other large-scale population-based biobanks, this study should provide insights into the shape and strength and causal relevance of the associations of SBP, BMI and WHR with HF and PHD in this population to guide prevention of these diseases.

4. Incidence and mortality of heart failure and pulmonary heart disease

4.1. Background

Heart failure (HF) and pulmonary heart disease (PHD) are major causes of death and disability. About 30% of the increase in the global burden of HF cases over the last three decades has originated in China[21]. While the burden and long-term prognosis of HF have been extensively studied in high-income countries, there is limited available evidence on the incidence and prognosis of PHD[21]. Autopsy studies suggest that PHD occurs in about half of COPD cases[32], and cases in high-income countries with both COPD and PHD have higher risks of stroke and all-cause mortality[33,35]. Previous single centre studies in urban areas of China suggested that HF cases were hospitalised about three times each year, with each admission costing about US $1,330[31]. However, no study has evaluated the incidence, prevalence or prognosis of PHD in China, but since COPD is common in China, the prevalence of PHD is likely to be substantial[131].

Before investigating the major risk factors of HF and PHD in this thesis, it is worthwhile to understand the burden of these conditions on population health. Understanding the burden of HF and PHD in the CKB could have important implications for policymakers to prioritise resources and implement targeted policies to reduce disease burden through prevention, early diagnosis and treatment.

This chapter aims to compare the first hospitalisation rates and short-term (28-day case-fatality ratios [CFRs]) and long-term (5-year cumulative mortality risk [CMR]) mortality risks for HF and PHD, overall and by sociodemographic factors and comorbidities in Chinese adults.

4.2. Statistical analysis

4.2.1. Outcome definitions

All first HF and PHD events diagnosed during outpatient consultations were excluded from this analysis as outpatient HF cases were only reported by certain CKB study centres. Because HF and PHD are insidious diseases, population-based estimates of the incidence rates of these diseases using such data will be seriously underestimated; hospitalisation rates, therefore, are a more reliable estimate of disease burden.

The primary outcomes were (i) First fatal or non-fatal HF and PHD hospitalisation rates and (ii) 28-day CFR and 5-year CMR of death from any cause following the first hospitalisation for HF or PHD.

4.2.2. Hospitalisation rates

HF or PHD hospitalisation rates were defined as the number of first hospitalisations per 100,000 person-years of follow-up. Hospitalisation rates across subgroups were adjusted by direct standardisation to the age-at-risk (five groups: 30–49, 50–59, 60–69, 70–79, 80+), sex and area (urban and rural) structure of the general CKB population. HF or PHD hospitalisation rates were standardised by age-at-risk (age at disease diagnosis) rather than baseline age to account for the excess conferred by additional age acquired during follow-up[120]. The Normal approximation to Poisson rate sums was used to estimate 95% confidence intervals (CI) of directly standardised hospitalisation rates[132].

4.2.3. Mortality risk

The 28-day CFR and five-year CMR were estimated for participants with first hospitalisation for HF or PHD. The 28-day CFR was the proportion of deaths from any cause among participants with a first HF or PHD hospitalisation occurring within 28 days of diagnosis.

In order to estimate the CMR, participants were followed from the date of first HF or PHD hospitalisation to death from any cause. The CMRs (and corresponding 95% CI) were calculated as one minus the estimate of the Kaplan-Meier survival probability (and corresponding 95% CI)[133].

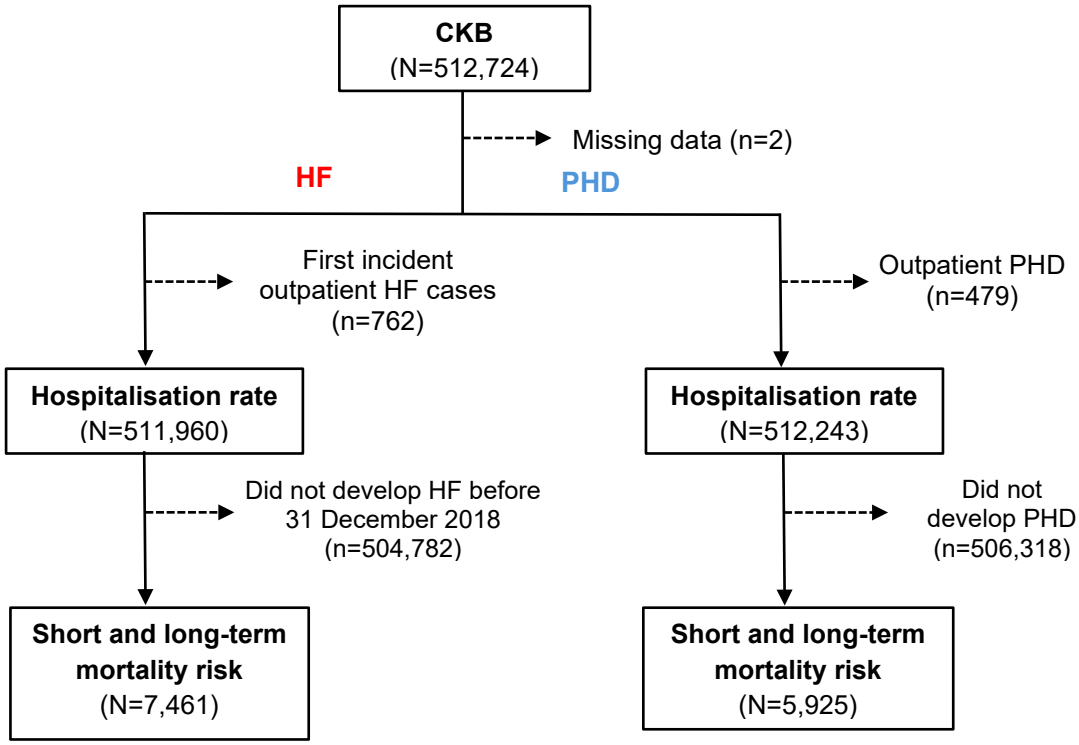
The CFRs and CMRs across subgroups, such as education and number of comorbidities, were adjusted for age at first hospitalisation for HF or PHD, sex and area (urban and rural) by direct standardisation to HF or PHD cases. All time-to-event analyses were censored on 31 December 2018 (the global study censoring date). The standardised estimates of hospitalisation rates and CFR by subgroups reflect the age-sex-area structure of the subgroups[132]. The heterogeneity and trend in hospitalisation rates, 28-day CFRs, and 5-year CMRs across subgroups were evaluated using the Likelihood ratio (LR) test for heterogeneity (for nominal subgroups, e.g., sex) and trend (for ordinal subgroups, e.g., age groups), respectively.

4.3. Results

4.3.1. Characteristics of the study population

Among the 512,724 included participants, 56,549 (11%) died and 4,028 (<1%) were lost to follow-up. **Figure 4.1** depicts the selection of participants for this analysis. A total of 8,233 and 6,404 incident cases of HF and PHD were recorded during a median of 10.1 years of follow-up, respectively. After excluding outpatient cases of incident HF (n=762) and PHD (n=479), 7,461 and 5,925 first hospitalisations for HF and PHD were included in the analysis. The mean age at first hospitalisation was 63.7 (SD 10.3) for HF and 63.7 (SD 10.3) for PHD (**Table 4.1**). About 13% (942/7461) of the HF cases also had PHD, and 16% (942/5925) of the PHD cases had HF.

Figure 4.1. Selection of study participants for various types of analysis.



The analyses were conducted separately for heart failure (HF) and pulmonary heart disease (PHD). Therefore, the number of participants excluded at each stage depended on the trait analysed.

Compared to the overall CKB population, participants with HF or PHD were about a decade older, mostly men, had lower levels of education and income, and had a higher prevalence of comorbidities (**Table 4.1**). PHD cases were mostly rural residents (85.5%) and agricultural workers (63.6%) and had a higher prevalence of COPD (50.9%) than the general CKB population. In contrast, cases with HF had a higher prevalence of CVD and CVD risk factors, including coronary heart disease (12.5%), hypertension (49.5%), and diabetes (14.5%) than the overall CKB population.

Table 4.1. Baseline characteristics of study participants with first hospitalisation for first heart failure and pulmonary heart disease.

Characteristics	Participants with HF* (n=7461)	Participants with PHD* (n=5925)	All CKB Participants (n=512,724)
Age, mean (SD), year	62.6 (9.3)	64.0 (8.4)	52.0 (10.7)
Women, n (%)	55.3	45.3	59.0
Demographic and SES, n (%)			
Rural area	54.6	85.5	55.9
No formal education	31.1	39.8	18.6
Annual household income<10,000 Yuan	38.5	59.1	28.2
Agriculture and related work	38.4	63.6	41.7
Lifestyle factors, n (%)			
Ever-regular smokers			
Men	76.5	85.3	74.4
Women	10.7	21.3	3.2
Regular alcohol drinkers			
Men	26.2	23.4	33.3
Women	2.3	4.1	2.1
Physical measurements, mean (SD)			
SBP, mmHg	142 (23.9)	137 (23.2)	131 (21.3)
DBP, mmHg	79.5 (12.3)	77.9 (12.0)	77.8 (11.2)
Prevalent conditions, n (%)			
Hypertension	59.5	45.8	35.2
Diabetes	14.5	5.86	5.91
IHD	12.5	4.57	3.02
Stroke and TIA	4.93	2.08	1.73
COPD	19.3	50.9	7.23

Abbreviations: HF=heart failure; PHD=pulmonary heart diseases; BMI=body mass index; SBP=systolic blood pressure; DBP=diastolic blood pressure; IQR=Interquartile range; SD=standard deviation; IHD=ischaemic heart disease; TIA=Transient Ischaemic Attack; SES=Socioeconomic status; COPD=chronic obstructive pulmonary disease.

* Participants with HF and PHD are not mutually exclusive

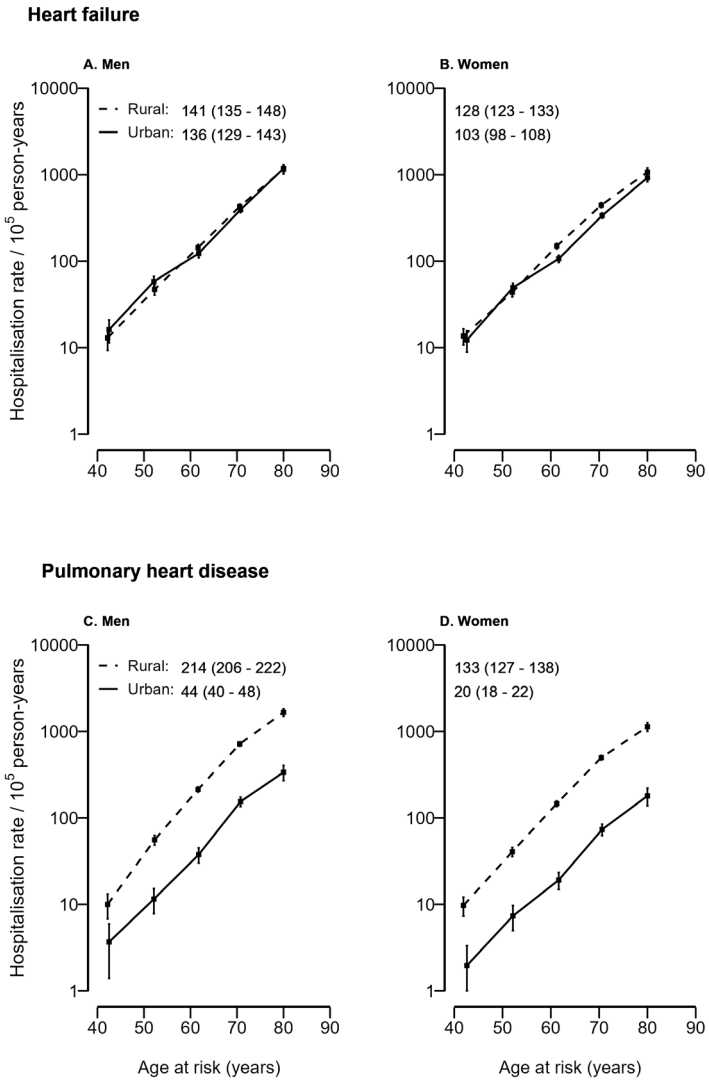
4.3.2. Hospitalisation rates for HF and PHD

The overall crude hospitalisation rates of HF and PHD were 125 (95% CI 122–128) and 99 (96–101) per 10^5 person-years, and the rates increased with age (**Figure 4.2**). The standardised hospitalisation rates of HF were greater in men (132, 128–137) than in women (120, 116–123) and in rural (130, 126–134) than in urban areas (118, 114–122). The standardised hospitalisation rates of PHD were greater in men (123, 119–127) than in women (80, 77–83) but, importantly, were about 6-fold greater in the rural (164, 160–169) than in urban areas (30, 28–32).

When stratified by sex, the hospitalisation rates of HF were similar in rural and urban men but were slightly higher in rural than in urban women older than 55 years old (**Figure 4.2**). By contrast, the hospitalisation rates of PHD were higher in rural than in

urban men (214, [206–222] vs 133[127–138]) and greater in rural than in urban women (133 [127–138] vs 20 [18–22]) (**Figure 4.2**).

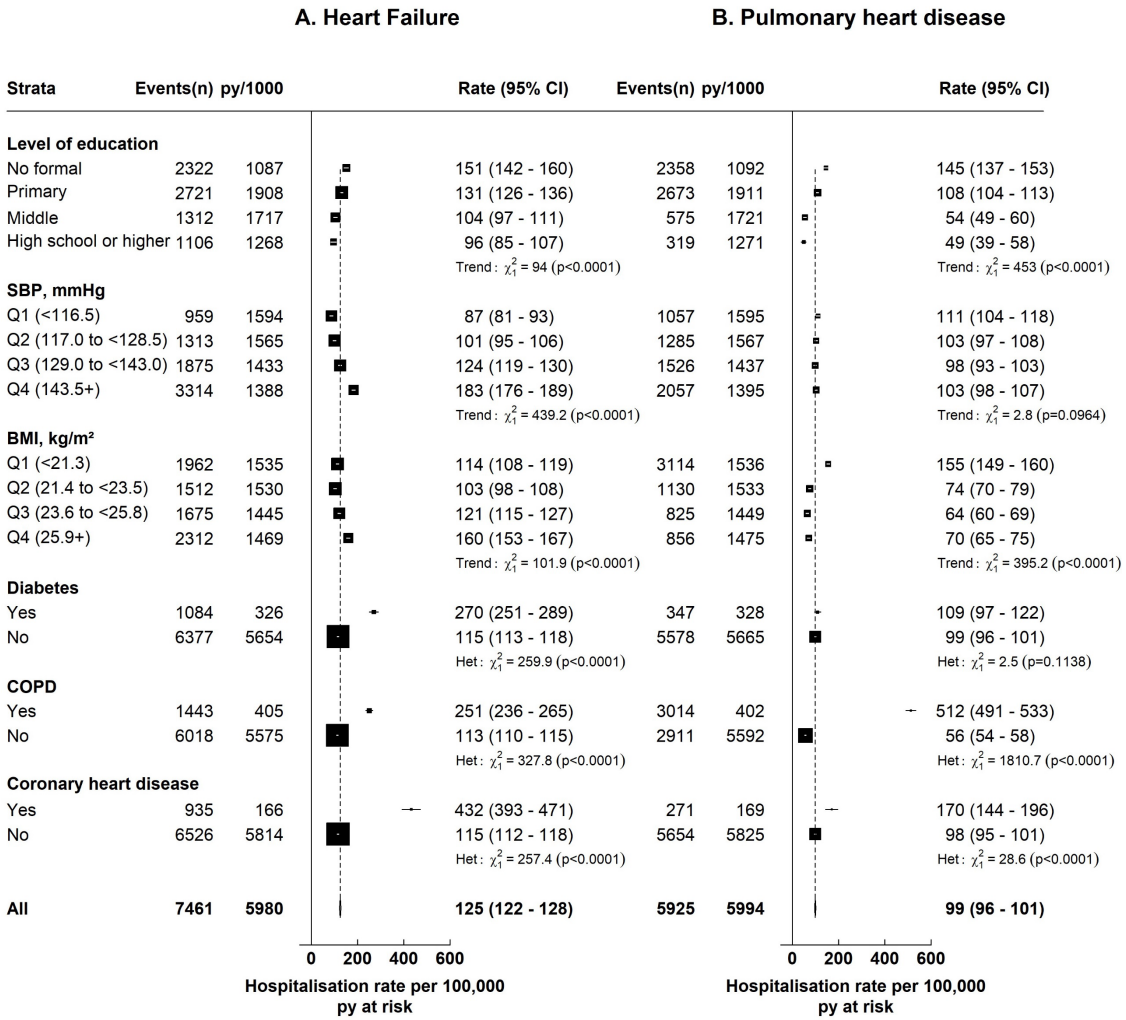
Figure 4.2. Age-specific hospitalisation rates for heart failure and pulmonary heart disease by sex and area.



Age-specific standardised hospitalisation rates were plotted against the mean age-at-risk in each age-at-risk group. The vertical lines through the black squares are the 95% confidence intervals (CI). The y-axis was log-transformed to visualise log-linear trends in hospitalisation rates across age-at-risk groups.

The standardised hospitalisation rates of HF and PHD were both higher in individuals with lower levels of education and prevalent COPD and coronary heart disease (**Figure 4.3**). Individuals with diabetes and those with higher SBP and BMI levels had higher standardised hospitalisation rates of HF (**Figure 4.3**).

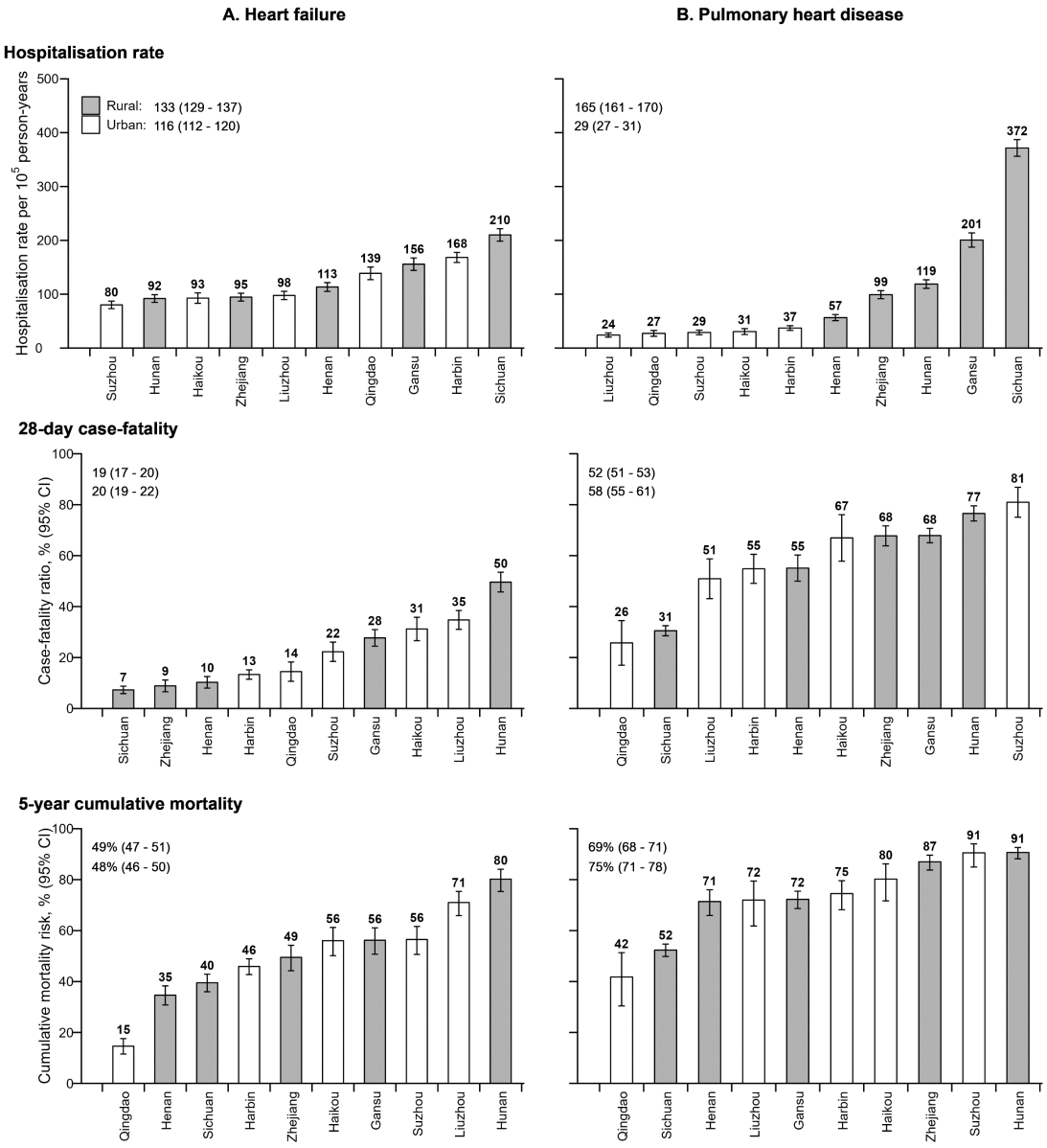
Figure 4.3. Standardised hospitalisation rates for heart failure and pulmonary heart disease by subgroup.



Hospitalisation rates were directly standardised by age-at-risk, sex, and area (urban/rural). BMI = Body mass index, COPD = chronic obstructive pulmonary disease, Het = Heterogeneity, py = person-years, SBP = systolic blood pressure.

The hospitalisation rates of HF and PHD varied substantially by study area (**Figure 4.4**) and were highest in Sichuan (210 per 10^5 person-years) and lowest in Suzhou (82 per 10^5 person-years) (**Figure 4.4**). The standardised hospitalisation rates of PHD were highest in Sichuan (377 per 10^5 person-years) and lowest in Liuzhou (24 per 10^5 person-years) (**Figure 4.4**).

Figure 4.4. First hospitalisation, 28-day case-fatality and 5-year cumulative mortality risk by study area.



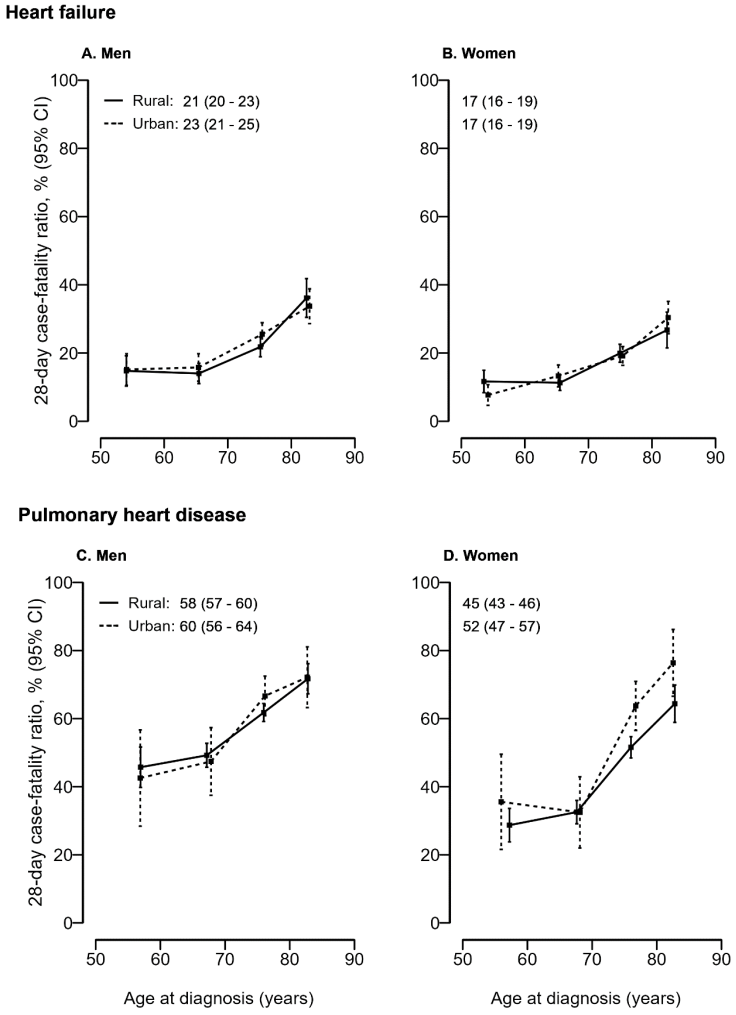
Hospitalisation rates were directly standardised by age-at-risk and sex; 28-day case-fatality ratio (CFR) and cumulative mortality risks (CMR) were directly standardised by age at diagnosis and sex.

4.3.3. 28-day case-fatality ratio of HF and PHD

Overall, 1,451 and 3,128 deaths occurred within 28 days of the first HF (n=7,461) and PHD (n=5,925) hospitalisation. The overall crude 28-day CFRs of HF and PHD were 18% (95% CI 17–19) and 53% (52–54). The standardised 28-day CFRs of HF and PHD were higher in men than in women (HF: 22 [20–23] vs 18 [16–19]; PHD: 58 [56–60] vs

46 [45–49]) with no rural-urban differences (HF: 19 [18–20] vs 20 [19–21]; 52 [51–53] vs 56 [53–60]). Similarly, sex-stratified analyses showed no rural-urban differences in the 28-day CFRs of HF and PHD in men and in women (**Figure 4.5**).

Figure 4.5. Age-specific 28-day case-fatality ratios after first hospitalisation for heart failure and pulmonary heart disease by sex and area.

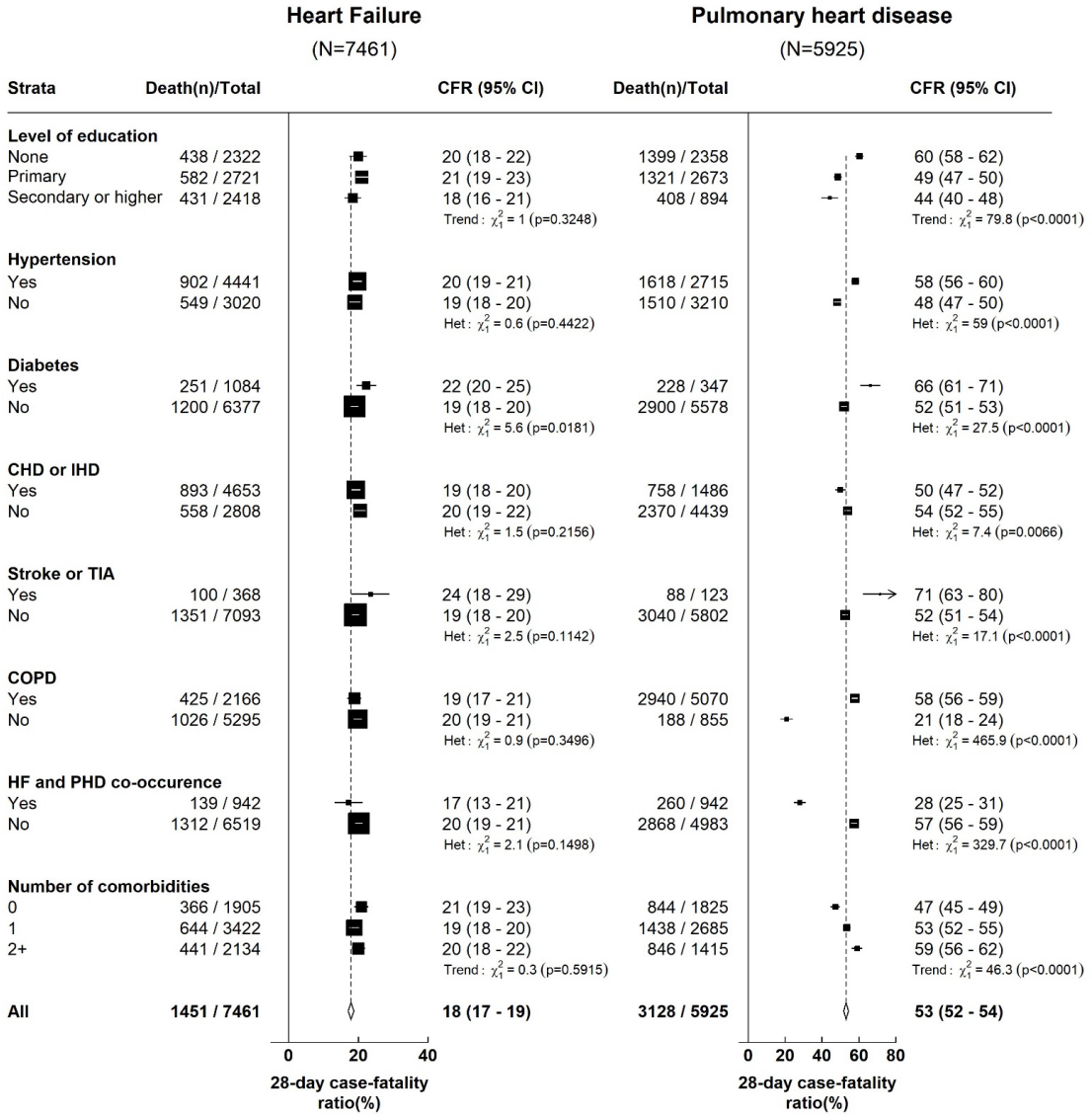


Age-specific case-fatality ratios (CFR: black squares) for heart failure (HF) and pulmonary heart disease (PHD) were directly standardised by the sex and area structure of CKB participants with first incident hospitalisation for HF and PHD, respectively. The vertical error bars through the squares are the 95% confidence intervals (CI).

The standardised 28-day CFR of HF and PHD were similar by levels of education, prevalent hypertension, diabetes, stroke or TIA, COPD, number of comorbidities, and among those with comorbid incident PHD (**Figure 4.6**). By contrast, the 28-day CFR of PHD was higher among those with lower levels of education and those with prevalent

hypertension, diabetes, stroke or TIA, COPD, and higher numbers of comorbidities at baseline, but was lower in individuals with incident ischemic heart disease and comorbid incident HF (**Figure 4.6**).

Figure 4.6. Standardised 28-day case-fatality ratio for heart failure and pulmonary heart disease by subgroup.



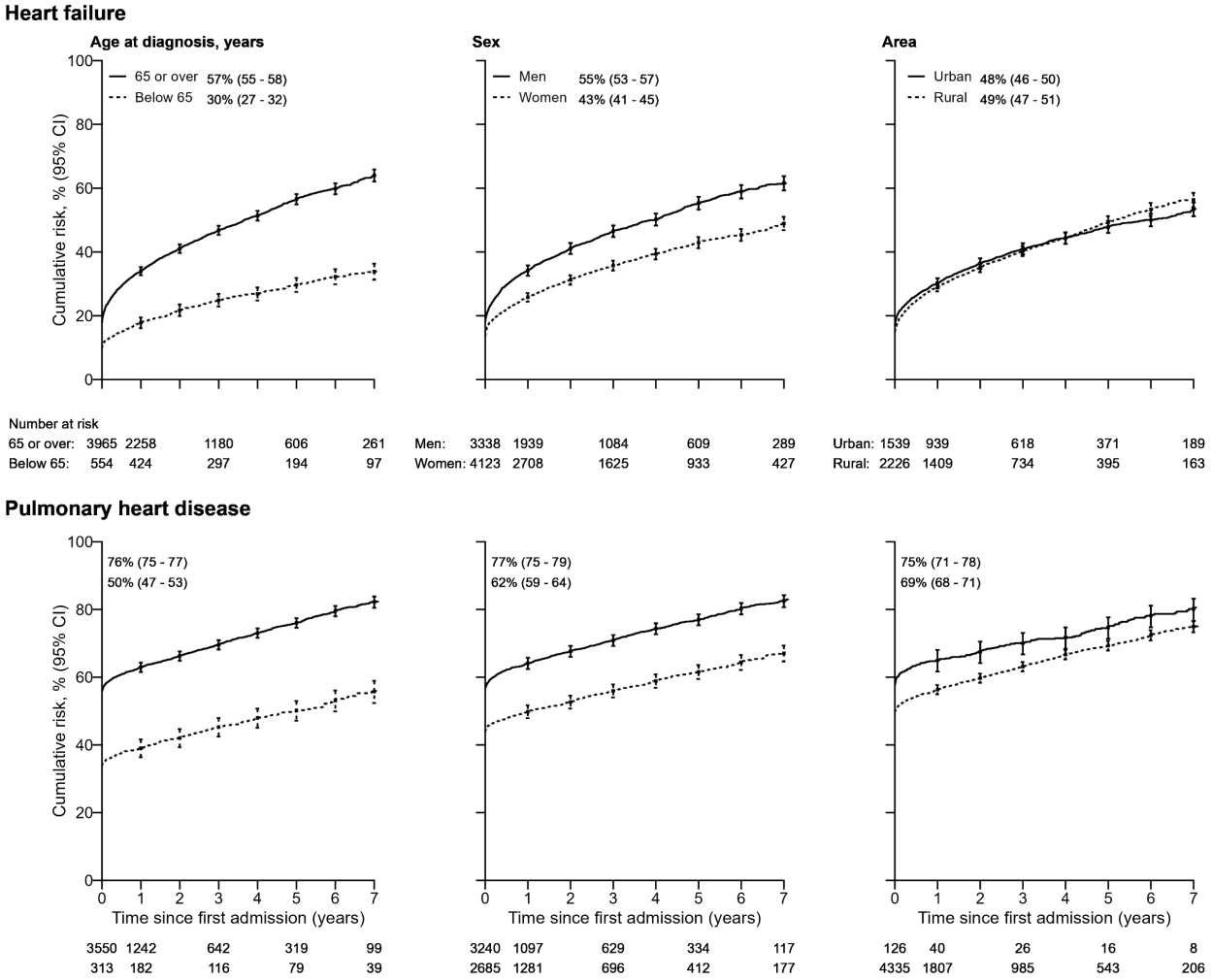
Case-fatality ratios (CFR) were directly standardised by age at diagnosis, sex and area (urban and rural). COPD = Chronic obstructive pulmonary disease, CHD = Prevalent coronary heart disease, IHD incidence ischaemic heart disease, TIA = Transient ischaemic attack, Het = Heterogeneity

The standardised 28-day CFRs of HF varied substantially across CKB areas, with no obvious rural-urban differences (**Figure 4.4**).

4.3.4. Cumulative mortality risk of HF and PHD

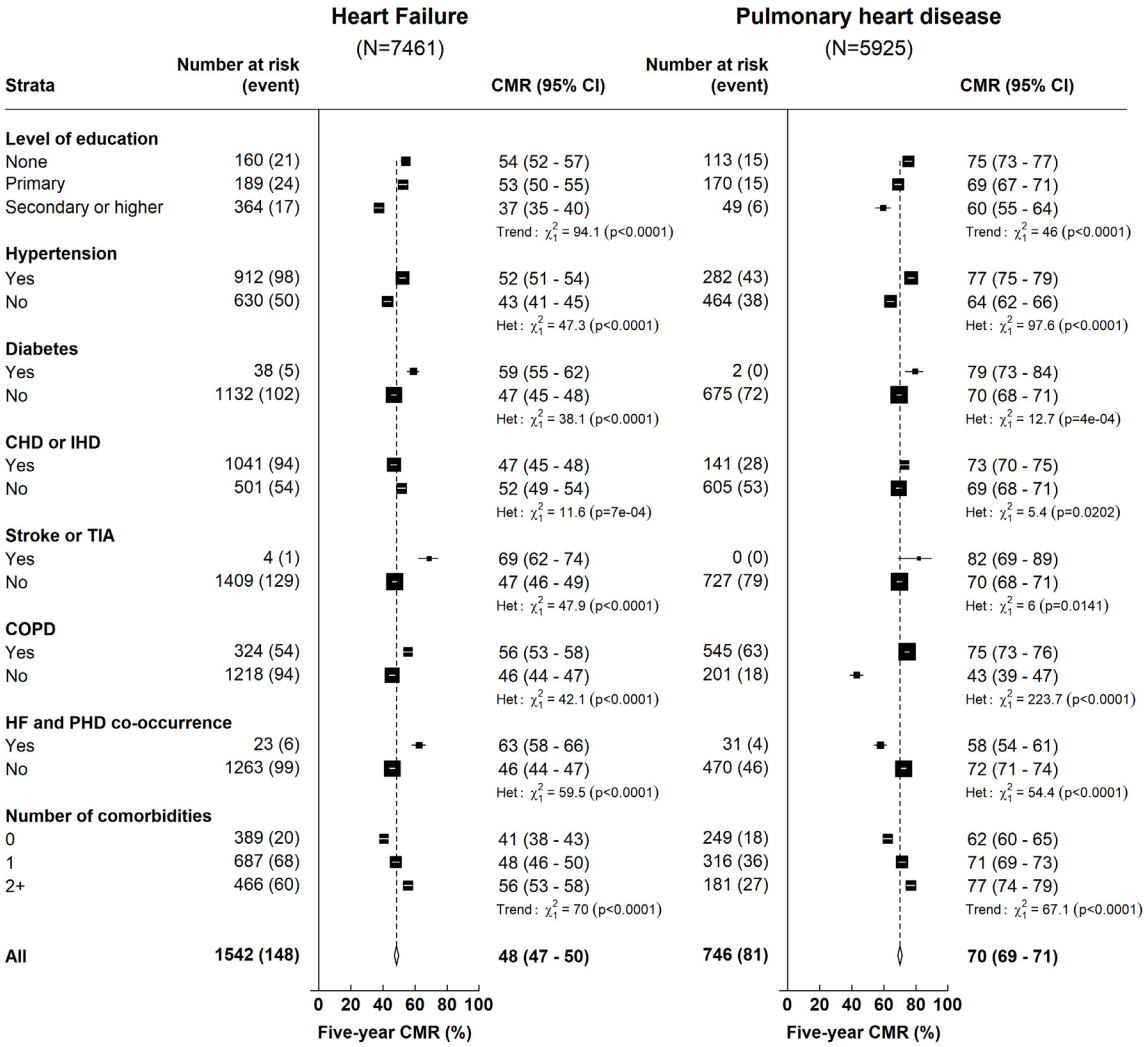
HF and PHD were associated with high 5-year CMRs of 47% (95% CI 35–58) and 70% (59–82), respectively. The standardised 5-year CMR of HF and PHD were higher in men than in women (HF: 55% [53–57] vs 43% [41–45]; PHD: 77% [75–79] vs 62% [59–64]) and in those ≥ 65 years at diagnosis compared to younger individuals (HF: 57% [55–58] vs 30% [27–32]; PHD: 76% [75–77] vs 50% [47–53]) (**Figure 4.7**). There were no rural-urban differences in the 5-year CMR of HF and PHD (**Figure 4.7**). The standardised 5-year CMR for both HF and PHD were higher in individuals with lower levels of education and a higher number of comorbidities (**Figure 4.8**). Participants with hypertension, diabetes, stroke or TIA, and COPD had higher 5-year CMRs (**Figure 4.8**). Moreover, the 5-year CMRs of HF and PHD varied substantially by study area (**Figure 4.8**).

Figure 4.7. Five-year cumulative mortality risks after a first hospitalisation for heart failure and pulmonary heart disease by age, sex and area.



The black squares and vertical error bars represent the point estimates of the cumulative mortality risks and their respective 95% CI. Cumulative mortality curves were directly standardised by age at diagnosis, sex, and area as appropriate.

Figure 4.8. Standardised 5-year cumulative mortality risk for heart failure and pulmonary heart disease by subgroup.



Cumulative mortality risks (CMRs) were directly standardised by age at diagnosis, sex and area (urban and rural). All conventions are as in **Figure 4.6**.

4.4. Discussion

4.4.1. Summary of the main findings

The present analyses of hospitalisation rate and prognosis of HF and PHD in Chinese adults demonstrated high hospitalisation rates of HF and PHD. HF and PHD hospitalisation rates increased with age, lower levels of education and comorbidities and were higher in rural than in urban areas. HF cases had high short-term and long-term mortality risks. The 28-day (short-term) and 5-year (long-term) mortality risks of PHD were 2.5 and 1.5-fold greater than those for HF. The short-term mortality risk for HF (18%) increased with age and was higher in men than in women; the short-term mortality risk for PHD (53%) was greater in men than in women and increased with age, number of comorbidities and among those with lower levels of education. Over half of all HF cases died within 5 years of a first hospitalisation; 70% of the PHD cases died within 5 years of a first PHD hospitalisation. The 5-year mortality risk of HF and PHD was higher in men than in women, older adults, those with lower levels of education and participants with a higher number of comorbidities.

4.4.2. HF and PHD hospitalisation rates

While the present study estimated hospitalisation rates of HF, most previous studies reported incidence of HF and included both inpatient and outpatient HF cases[27,31,134–136], which limits comparisons with the present study. Consistent with previous studies, we found higher hospitalisation rates of HF in China than in Western countries[27,31,134,136]. The hospitalisation rates of HF were higher among older individuals and those with lower levels of education. Individuals with hypertension, diabetes, prior CVD, and those with a higher number of comorbidities had higher hospitalisation rates of HF. Age was the chief determinant of HF hospitalisation, and the rates in those aged 80 years or older were at least 70-fold greater than those aged 40 years. The hospitalisation rates of PHD were higher among individuals with prior COPD and, to a lesser extent, in those with coronary

heart disease. In addition, the rates did not vary by the presence of hypertension or diabetes, indicating that the hospitalisation rate of PHD was chiefly determined by COPD or other chronic lung diseases, while the hospitalisation rate of HF reflected differences in the prevalence of established CVD risk factors.

No previous study has examined geographic differences in the hospitalisation rates of HF and PHD in China, but the present study demonstrated substantial differences in the hospitalisation rates of PHD and HF within China. These geographic differences in the hospitalisation rates of HF and PHD chiefly reflect differences in the prevalence of risk factors, including socioeconomic, demographic, behavioural, and environmental factors[137].

4.4.3. HF and PHD mortality rates

Data on the 28-day CFR (short-term mortality) of HF and PHD are limited even in high-income countries[27,138], and prospective studies of HF cases in China and other Asian countries have mostly reported in-hospital mortality risks for HF[135]. The estimated short-term mortality risk for HF in the present study (18%) was comparable with those reported in a prospective multicentre registry of 4153 acute HF cases in the Czech Republic (20%)[138]. However, the short-term mortality risk for HF in this study was about 2-fold greater than those reported in the USA in the Framingham Heart Study of 1075 HF cases (11%)[27], a recent study of 105,399 incident HF hospitalisation and death records between 2007 and 2016 in Malaysia (11%)[139], and a pooled estimate from a meta-analysis of 1.5 million chronic and stable HF cases (4.3%)[28]. Twenty-eight-day CFR is an important measure of hospital quality of care and predicts long-term disease outcomes[140]. The CFR is likely to vary substantially by calendar year, country, and geographic region partly due to differences in clinical practice and health policies for HF care, including policies to improve accessibility and affordability of outpatient follow-up and evidence-based medicines.

The 5-year CMRs (long-term mortality) for HF in CKB (47%) were comparable to those of a previous prospective hospital-based multicentre cohort study of 3335 urban Chinese adults with acute HF living in Beijing, China (55.4%)[30], but were lower than the pooled estimate from a meta-analysis of 1.5 million chronic and stable HF cases in high-income countries (43.3%)[28]. These differences could reflect variations in the age of HF patients in high-income (mean age ~70 years) compared to those in developing countries (mean age ~60 years), including China and other Asian countries[135,141]. Older participants are more likely to have a higher prevalence of comorbidities, including ischaemic heart disease and atrial fibrillation, which are associated with higher risks of all-cause mortality. Furthermore, differences in the long-term mortality risks could be explained by geographical differences in the causes of HF. In contrast with Europe, high-income Asia Pacific, or North America, ischaemic heart disease is a less common cause of HF in China[21]. HF populations with a higher prevalence of ischaemic heart disease are more likely to have a higher mortality risk[30].

The long-term mortality risk of PHD in this study was about 2.5-fold greater than that in a prospective study of 4862 UK patients with COPD and incident HF[142]. These differences could reflect disparities in the available treatments for PHD or access to care in both settings. Importantly, PHD cases in the CKB were from a lower socioeconomic background and, hence, may have reduced access to high-quality care. The high 5-year CMR of about 50% for HF and 75% for PHD were much higher than 26% for ischaemic stroke and subtypes or 23% for acute myocardial infarction in the present study population[143,144]. In addition, the 5-year CMR of PHD was considerably higher than for intracerebral haemorrhage (60%) in this population, which is the stroke subtype with the highest mortality rate[143].

The short-term and long-term mortality risks of PHD were about 1.5 to 3-fold greater than for HF, indicating that PHD is a more severe disease and has a more adverse

short-term and long-term prognosis. PHD results from severe, chronic respiratory diseases such as COPD, which are associated with substantial morbidity and mortality. In a cohort of Chinese adults admitted for COPD exacerbations and a meta-analysis of COPD patients, patients with PHD had more comorbidities, frequent recurrent hospital admissions, longer duration of hospital stay, and higher mortality than those without PHD[33,145].

Participants with PHD and HF had lower short-term and long-term mortality risks compared to those without HF. The reasons for these unexpected findings are unclear. However, possible explanation could be survivor bias[146–148]. Individuals with PHD who survive long enough to develop HF are more likely to have milder forms of PHD compared to those without HF, leading to better survival among those with both PHD and HF.

4.4.4. Implications of the findings for analyses of risk factors for HF and PHD in this thesis

The present analyses identified age as the strongest determinant of HF and PHD hospitalisation, perhaps reflecting suboptimal disease prevention programs in early middle age. Previous reports have also reported strong positive associations of age with incident HF[27,31,136]. In addition, there were substantial variations in HF and PHD hospitalisation rates across CKB study areas, which were more striking for PHD, suggesting regional differences in risk exposure that may require targeted interventions to prevent HF and PHD. Taken together with the results in Chapter 3, which demonstrated differences in participants' mean age and levels of factors such as SBP and BMI across CKB study areas, these findings highlight the importance of investigating the risk factors for HF and PHD in the CKB after careful control for confounding by age and study area.

Moreover, HF hospitalisation rates were higher among participants with traditional cardiovascular disease risk factors or comorbidities (BMI, hypertension, diabetes and coronary heart disease), whereas PHD hospitalisation was considerably higher in participants with prevalent COPD. These findings suggest that COPD and other chronic respiratory diseases could be the primary drivers of the epidemiology of PHD in the CKB, and less so by traditional cardiovascular disease risk factors.

Furthermore, the current analysis indicated a J-shaped relationship of BMI with HF hospitalisation and an inverse relationship with PHD, suggesting that the prospective associations of BMI with HF, especially PHD, are susceptible to confounding and reverse causality.

4.4.5. Strengths and limitations

This study had several strengths, including being the first study conducted in China to compare the long-term mortality risks for HF and PHD. In addition, the large number of events, long-term follow-up and the time when the events occurred enabled reliable estimation of the short- and long-term mortality risks after HF or PHD. Furthermore, the breadth of exposure collected at baseline in the CKB study allowed the investigation of HF and PHD hospitalisation rates by demographic groups and comorbidities.

However, the study also had several limitations including the lack of generalisability of the findings to the overall population of China because the CKB is not representative of the Chinese population. HF and PHD cases in CKB have not currently been adjudicated, and hence, we cannot exclude the possibility of some misclassification of these cases.

4.5. Summary

The present analysis demonstrates high hospitalisation and mortality rates for HF and PHD in Chinese adults. The findings demonstrated that HF and PHD are important public health problems that warrant prioritising to improve population health.

Understanding the causes of HF and PHD could guide the development of more effective strategies for preventing these conditions.

The findings from the present analysis highlight potential sources of bias. The analysis identified age and study areas as strong confounders and highlighted the importance of confounding and reverse causality bias in any analyses of BMI as risk factors for HF or PHD. These findings will inform analyses in subsequent chapters investigating the associations of SBP and adiposity as risk factors for HF and PHD.

5. Associations of systolic blood pressure with risk of heart failure and pulmonary heart disease

5.1. Background

High BP is a major cause of CVD and premature mortality worldwide, causing over 25 million deaths each year. As part of a global target, the WHO seeks to reduce the global prevalence of hypertension by about a third. The Chinese government implemented nationwide programmes to improve healthy lifestyles and reduce the risk factors of disease, including hypertension. These programmes led to a 5% reduction in the prevalence of hypertension between 2010 (29.6%) and 2018 (24.7%), but awareness (38%), treatment (35%), and control rates (12%) remained poor[149].

Higher levels of SBP are associated with higher risks of HF in prospective cohort studies. The majority of these studies, however, were conducted in high-income countries[54,55,86–89]; only the Asian Pacific Cohort Studies Collaboration (APCSC) reported on the association of SBP with HF in non-Western countries. The APCSC found that the relative risks of HF per 10 mmHg higher levels of SBP in different studies varied from 1.07 to 1.37, respectively. The few studies that reported on the shape of the associations between SBP and HF reported conflicting results, with some studies reporting a J-shaped association[54,55] and others reporting a linear association[89].

Despite SBP being an established risk factor for HF and the leading cause of HF in China, uncertainties still persist about the shape and strength of this association. In addition, no study has investigated the relevance of bias from confounding and reverse causality in the associations of SBP with HF. Large studies are needed to reliably assess the risk of disease across the full range of risk factors and assess the extent to which these associations are mediated through a third variable. Such data are essential to reliably estimate the burden of HF attributed to SBP and risk of HF associated with the lowest levels of SBP. Furthermore, epidemiological studies need a long follow-up period

to evaluate and address reverse causality bias. Moreover, no study has explored the association between SBP and PHD.

This chapter investigates the association of usual SBP with HF and PHD in the CKB. Specifically, it aims to (i) evaluate the shape and strength of these associations, (ii) identify effect modifiers, and (iii) assess the extent to which the associations can be explained by confounding and reverse causality bias.

5.2. Methods

5.2.1. Study population

This analysis included participants at baseline but excluded those with histories of rheumatic heart disease (RHD, n=937), coronary heart disease (CHD, n=15,472), or stroke or transient ischaemic attack (TIA, n=8,884) at baseline to prevent reverse causality[58].

5.2.2. Outcome definition

The outcomes for this analysis were fatal and non-fatal incident HF (I50) for the analysis of HF and fatal and non-fatal incident PHD (I27) for the analysis of PHD. Observations were censored at the first HF or PHD event, loss to follow-up, death from any cause or the global censoring date (31 December 2018).

5.2.3. Statistical analysis

5.2.3.1. Data exploration, cleaning and manipulation

The distribution of continuous variables and the presence of outliers and implausible values were assessed using histograms, boxplots and value ranges of these variables. In addition, variables were assessed for missing observations. Missing values for BMI at baseline (n=2) were imputed using the mean BMI values for the study population.

5.2.3.2. Descriptive analyses

The baseline characteristics of the study participants were stratified by quintiles of SBP at baseline (SBP cut-offs: 114, 122.9, 131.9, 145.9) to assess the distribution of baseline characteristics across levels of SBP. Categorical and quantitative variables were summarised using percentages and mean (and standard deviation [SD]), which were directly standardised to the age, sex and area (all 10 study areas) structure of the CKB population.

5.2.3.3. Selection of confounders

All analyses were adjusted for potential confounders. The following confounders were selected based on prior knowledge and published evidence from prospective cohort studies investigating the association of SBP with HF: age, sex, study area, education, income, occupation, smoking, alcohol consumption, sedentary time, TPA, BMI and heart rate. In addition, other potential confounding variables were included in the multivariable analysis if they led to a ≥10% change in the logHR of the main exposure-outcome association[150,151] before and after adjustment for the potential confounder[151]. This approach further identified ambient outdoor temperature as a potential confounder, which was included in the final multivariable analysis.

5.2.3.4. Association of SBP with HF and PHD

Stratified multivariable Cox proportional hazard regression models were used to estimate hazard ratios (HRs) for the associations between sex-specific quintiles of SBP and first incident HF and PHD using time since study onset as the timescale. Minimally adjusted models were stratified by age-at-risk (5-year age bands, nine categories), sex, and all 10 study areas and adjusted for baseline age, baseline age^2, monthly household income (four categories: <10,000, 10,000–29,999, 30,000–34,999 and ≥35,000 Yuan), education (four categories: no formal, primary, middle and high school or higher education), occupation (two categories: agricultural and non-agricultural workers),

alcohol consumption (four categories: never, ex-, occasional and current drinkers), smoking (three categories: never, ex-smokers/occasional and current smokers), sedentary time, (sedentary time)^2, TPA, TPA^2, heart rate, (heart rate)^2, BMI, BMI^2 and use of BP-lowering drugs. Failure to account for medication use can bias epidemiological associations between SBP and disease[152]. This analysis adjusted for medication use to improve the validity and reliability of the findings[152,153].

For continuous independent variables, the LR test for non-linearity was used to assess deviations from the linearity assumption[154]. Departure from linearity was assessed by comparing the change in χ^2 statistic between a nested model with the covariate in its linear form to a more complex model containing an additional quadratic term of the covariate. In case of a significant change in the χ^2 statistic, indicating a significant departure from linearity, a quadratic term of the covariate was added to the model to account for non-linearity; Otherwise, the covariate was modelled in its continuous form[154].

5.2.3.5. Shape of the association

Risk estimates for categorical risk factors with multiple levels are usually presented relative to a “reference group.” If this reference group contains too few cases, the CIs in the non-reference group can be inflated, obscuring any risk difference between these groups. Easton et al., 1991 proposed floating absolute risks (FARs) as an alternative approach to present risk estimates for categorical exposures with multiple categories[155,156]. In 2004, Plummer proposed a method for improving estimation of the FAR variances initially suggested by Easton et al[157]. The FAR method permits pairwise comparisons of the risk between any two categories rather than restricting comparisons of group-specific HRs to an arbitrarily-defined reference category[155,156].

The FAR method was used to calculate the variance of the log risk in each baseline-defined fifth of SBP, including the "reference group"[155–157]. The variance of the log risk, which reflects the amount of information in each SBP group, was then used to calculate the 95% CI in all SBP groups. The FARs were plotted against the mean levels of SBP at resurvey (usual SBP levels) for each baseline-defined SBP group to inspect the shape of the association and correct it for regression dilution bias. In the presence of linear relationships, the FARs were connected using regression lines. The intercept and slope of the lines were estimated by regressing the quintile-specific means of usual SBP on the log FARs, and the regression estimates were weighted by the inverse of the variance of the log FAR in each SBP group[158]. Unless stated otherwise, all logarithms are to the base e (natural logs).

The group-specific CIs of the FARs are narrower than conventional CIs because they do not reflect the CI for the relative risk compared to the reference category[156]. Hence, FAR CIs were only used when plotting; conventional CIs were used when directly comparing two groups (e.g., top versus bottom quintiles).

5.2.3.6. Strength of the association

In the case of a linear association between baseline SBP and HF or PHD, log HRs (and their standard errors) per 1 mmHg higher baseline SBP for disease outcomes were multiplied by 10 to obtain the log HR per 10 mmHg higher baseline SBP. The scaled log HRs (and their standard errors) were then divided by the regression dilution ratio (RDR) before exponentiating to obtain HRs per 10 mmHg higher usual SBP; the 95% confidence intervals were calculated as exp([log HR/RDR] ± 1.96×[SE/RDR]), where "exp" is the exponential function, SE is the standard error of the log HR and 1.96 is the Z value for a 95% CI for a two-tailed hypothesis test[132].

RDRs were estimated using Rosner's regression method as described in Chapter 3 with adjustments for age at baseline, sex and urbanicity (urban/rural)[128]. In age-sex-specific analyses, the log HRs for each age-sex group were corrected for regression dilution using age-sex-specific RDRs; the sex-specific log HRs and their standard error were combined using inverse variance-weighted fixed-effect meta-analysis to calculate the age-specific HRs.

5.2.3.7. Quantifying the effects of confounding

The confounding effects of covariates included in the final multivariable Cox regression model was assessed by estimating changes in the logHR of the association between usual SBP and disease outcomes before and after adjustment for the covariates[159,160]. Although there is no formal test for confounding, the percentage change in logHR permits a semi-quantitative assessment of the potential confounding effect of covariates included in this analysis[159,160]. Percentage changes in the logHR can also indicate whether the observed association is likely to be explained by residual confounding[159,160]. A considerable reduction in the logHR—a two-third reduction would be arbitrarily considered a substantial reduction—suggests any observed association between the exposure and outcome is likely due to residual confounding rather than being causal[159,160].

5.2.3.8. Subgroup analyses

The LR test was used to investigate effect modification by age-at-risk, sex, urban or rural residence, smoking, alcohol consumption and diabetes by fitting interaction terms into the multivariable Cox regression models. The LR tests for heterogeneity and trend were used to assess effect modification across levels of nominal and ordinal categorical variables, respectively.

5.2.3.9. Sensitivity analysis

The effect of reverse causality was investigated by sequentially excluding: (1) the first two years of follow-up; (2) poor self-reported health at baseline; (3) prevalent chronic lung diseases, including COPD, tuberculosis, emphysema, bronchitis and asthma; and (4) Co-occurring HF and PHD events.

The Cox proportionality assumption was assessed by visual inspection of the correlation of scaled Schoenfield residuals with the log of survival time, and it was satisfied[161,162]. All hypothesis tests were two-tailed, and two-tailed p-values less than 0.05 were considered statistically significant. Data were analysed using R version 4.2.2[163]. Meta-analysis was conducted using the *"metafor"* package[164]. Plots were created using the *"ggplot2"*[165] and *"Jasper"*[166] packages in R.

5.3. Results

5.3.1. Baseline characteristics of the study population

A total of 488,755 of the 512,724 participants recruited at baseline were included in the present analyses. In total, 23,969 participants with prevalent RHD, CHD and stroke or TIA were excluded from these analyses. **Table 5.1** shows the baseline characteristics of the participants included in the analysis. The mean (SD) age of the participants was 51.6 (10.6) years, 59% were female, 57% lived in rural areas and 57% were either agricultural or factory workers (**Table 5.1**). About 86% of men (compared to 4.9% of women) were ever-regular smokers, whereas 34% of men (compared to 2.1% of women) were currently alcohol drinkers. About 34%, 5.4% and 7% of the participants had hypertension, diabetes and COPD.

Concerning the trend of baseline characteristics with quintiles of SBP (**Table 5.1**), participants in the higher quintiles of SBP were older, more likely to be men, live in rural

areas, be current drinkers, and have diabetes and higher BMI and resting heart rate (RHR), respectively.

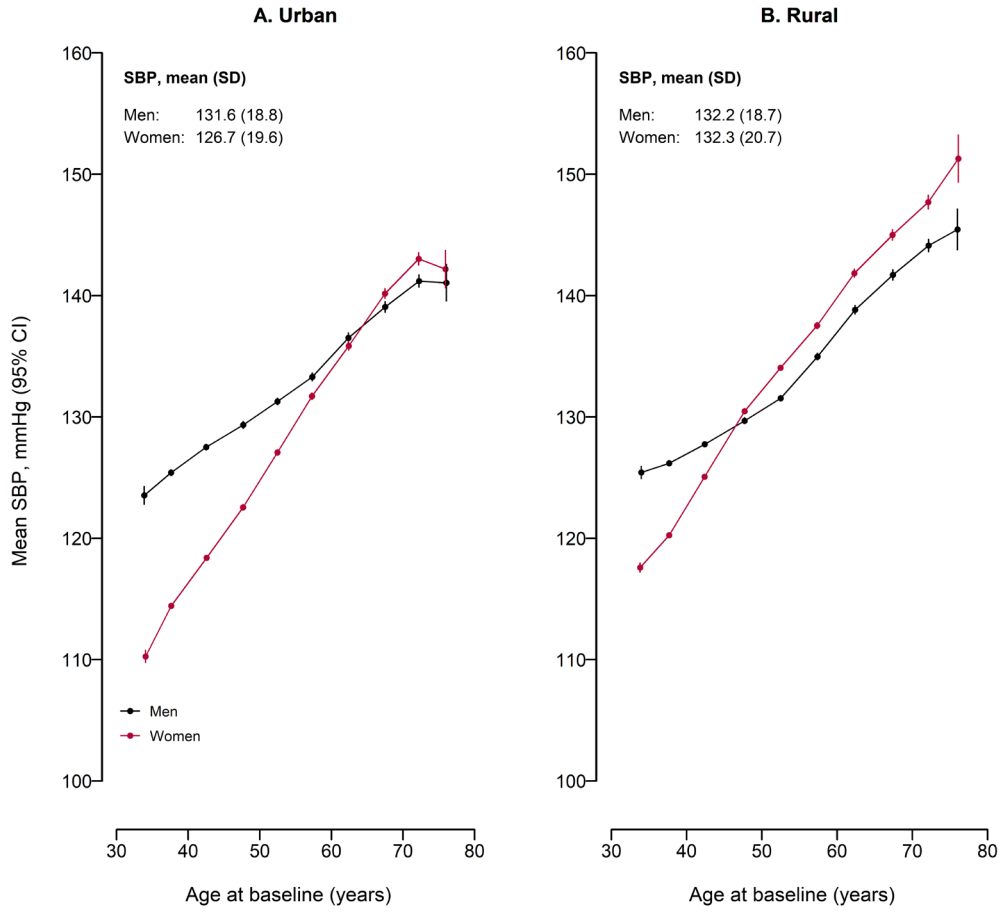
The mean (SD) levels of SBP, DBP and RHR at baseline were 131 (21) mmHg, 78 (11) mmHg and 79 (12) bpm, respectively. The mean (SD) levels of SBP for men and women were 134 (22) and 132 (24), and the mean DBP were 78 (12) and 77 (11), respectively. SBP levels increased linearly with age regardless of age and area of residence (**Figure 5.1**). At age 32, the average SBP of urban men was 14 mmHg higher than that for urban women of the same age, but this gap gradually diminished with increasing age and completely disappeared at 65 years (**Figure 5.1).** In contrast, the mean levels of SBP between rural men and women were approximately similar. At age 32 years, the mean level of SBP for rural men was 9 mmHg higher than that for rural women. This trend reversed by age 50 years, with rural women consistently having 1–2 mmHg higher mean SBP up to the age of 85 (**Figure 5.1**).

Table 5.1. Baseline characteristics of study participants by fifths of baseline systolic blood pressure.

Characteristics	Fifths of systolic blood pressure (SBP), mmHg (Range)					All
	Q1 (<113.5)	Q2 (114.0 to <123.0)	Q3 (123.5 to <132.0)	Q4 (132.5 to <146.0)	Q5 (146.5+)	
No. of participants	101,164	95,311	96,955	98,213	97,114	488,757
SBP, mean (SD), mmHg	106.1 (6.9)	118.6 (2.9)	127.6 (2.7)	138.5 (4.3)	161.9 (19.5)	130.6 (21.1)
DBP mean (SD), mmHg	67.0 (7.7)	73.0 (6.9)	77.0 (7.1)	81.8 (8.3)	91.6 (14.5)	77.7 (11.1)
Demographic and SES factors, %						
Age, mean (SD), years	47.8 (9.8)	48.8 (9.6)	50.7 (10.1)	53.8 (10.2)	57.9 (9.9)	51.6 (10.6)
Women	64.9	58.0	54.3	52.6	55.7	59.1
Rural area	48.4	54.3	60.1	60.8	63.0	56.8
Ambient outdoor temperature, mean (SD), °C	19.5 (9.6)	17.2 (9.2)	15.9 (9.3)	14.9 (9.9)	13.9 (12.5)	16.2 (10.0)
Household month income < 10,000 Yuan	29.0	28.0	27.8	28.2	29.3	28.3
Agriculture or factor worker	57.2	58.1	57.9	57.8	57.3	57.4
Lifestyle factors, %						
Ever regular smoker, Men	87.3	85.7	85.8	85.3	85.3	85.8
Ever regular smoker, Women	5.9	5.2	4.8	4.6	4.6	4.9
Current drinkers, Men	28.8	31.2	33.4	36.2	39.3	33.9
Current drinkers, Women	2.1	2.3	2.0	2.0	2.1	2.1
Total physical activity, mean (SD), MET-hour/day	21.5 (13.8)	21.6 (12.0)	21.8 (11.9)	21.5 (12.6)	21.2 (16.8)	21.5 (13.9)
Sedentary time, mean (SD), hours/day	3.0 (1.7)	3.0 (1.5)	3.0 (1.5)	3.0 (1.5)	3.0 (2.0)	3.0 (1.5)
Physical measurements, mean (SD)						
BMI, kg/m²	22.2 (3.3)	23.1 (3.2)	23.7 (3.2)	24.3 (3.6)	25.0 (5.0)	23.6 (3.4)
Waist circumference, cm	76.4 (9.8)	78.8 (9.1)	80.3 (9.1)	81.8 (9.8)	83.4 (13.2)	80.0 (9.7)
Body fat percentage, %	24.9 (6.8)	26.8 (6.4)	28.0 (6.5)	29.2 (7.2)	30.6 (9.7)	27.9 (8.4)
Resting heart rate, bpm	76.5 (12.5)	77.9 (11.4)	79.0 (11.7)	80.4 (13.6)	82.0 (19.1)	78.9 (11.8)
Prevalent conditions						
Hypertension	2.3	5.2	10.7	52.4	100.0	33.7
Taking BP-lowering drugs	1.9	4.0	6.7	11.8	22.4	10.1
Diabetes	2.9	3.8	4.9	6.4	8.5	5.4
COPD	8.1	7.4	7.1	7.0	6.6	7.1
Emphysema or bronchitis	2.8	2.6	2.6	2.5	2.2	2.5
Asthma	0.6	0.5	0.6	0.5	0.5	0.5
Self-reported poor health	10.3	8.9	8.7	9.0	10.5	9.4

Means and proportions where directly standardised by the age, sex, and all 10 CKB study areas.
BP=Blood pressure; bpm = beats per minute; COPD = Chronic obstructive pulmonary disease; DBP = Diastolic blood pressure; MET = Metabolic Equivalent Task; SBP = Systolic blood pressure; SD = Standard deviation.

Figure 5.1. Age-specific mean systolic blood pressure levels by sex and urbanicity.



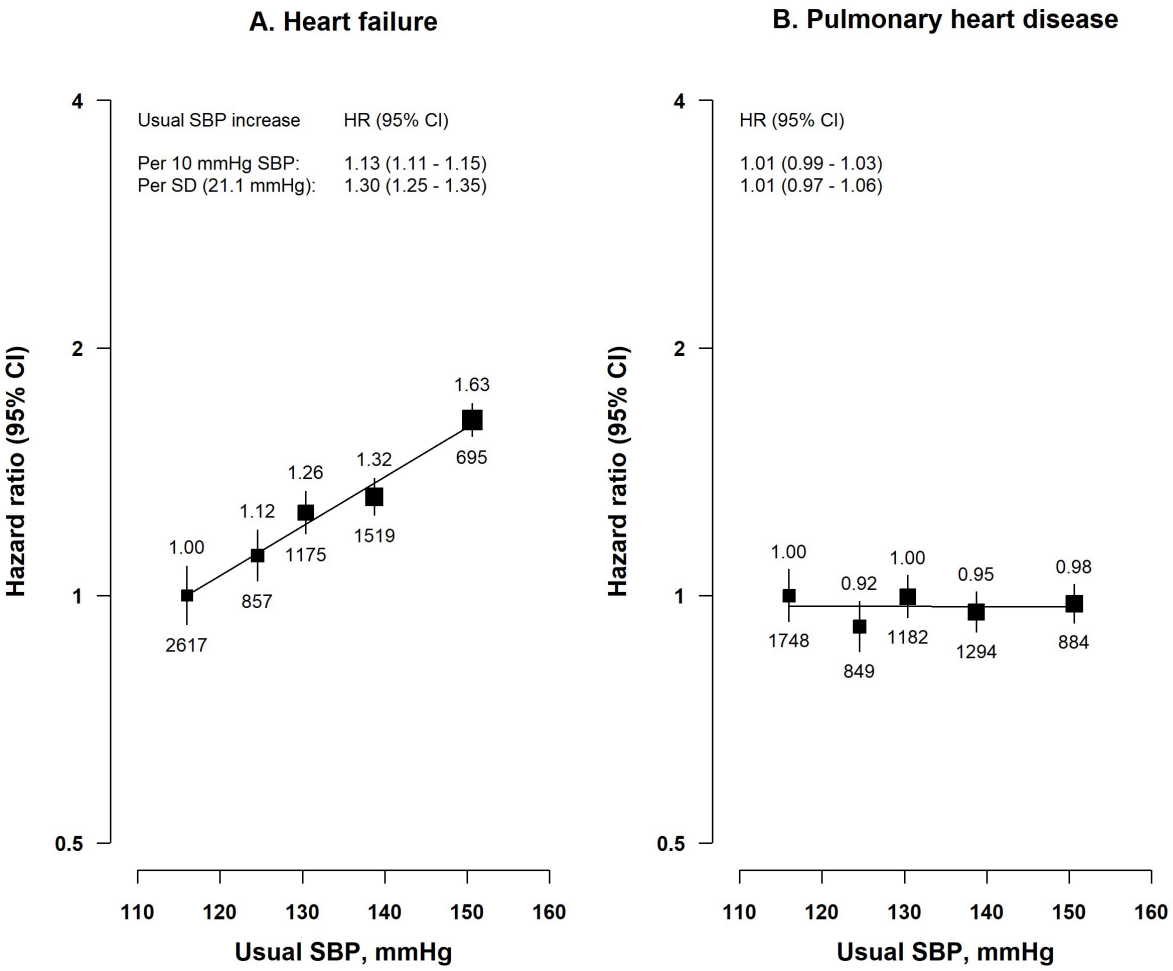
The solid circles and vertical lines represent the age-specific mean systolic blood pressure (SBP) and their 95% confidence intervals (CIs), respectively. SD = Standard deviation.

5.3.2. Associations of usual SBP with HF and PHD

During a median of 10 years of follow-up, 8,223 and 6,404 cases of HF and PHD were recorded, respectively. The mean (SD) age at diagnosis of HF was 65.6 (10.0) years, whereas that for PHD was 66.7 (9.0) years.

After adjusting for potential confounders, higher levels of SBP were positively associated with higher risks of HF with a log-linear shape across the range of usual SBP studied (**Figure 5.2**). Each 10 mmHg higher usual SBP was associated with about 13% higher risk of HF (HR 1.13, 95% CI 1.11–1.15, P_{trend} <0.0001). By contrast, there was no evidence of any association between usual SBP and PHD (**Figure 5.2**).

Figure 5.2. Associations of fifths of usual systolic blood pressure (SBP) with (A) heart failure and (B) pulmonary heart disease.

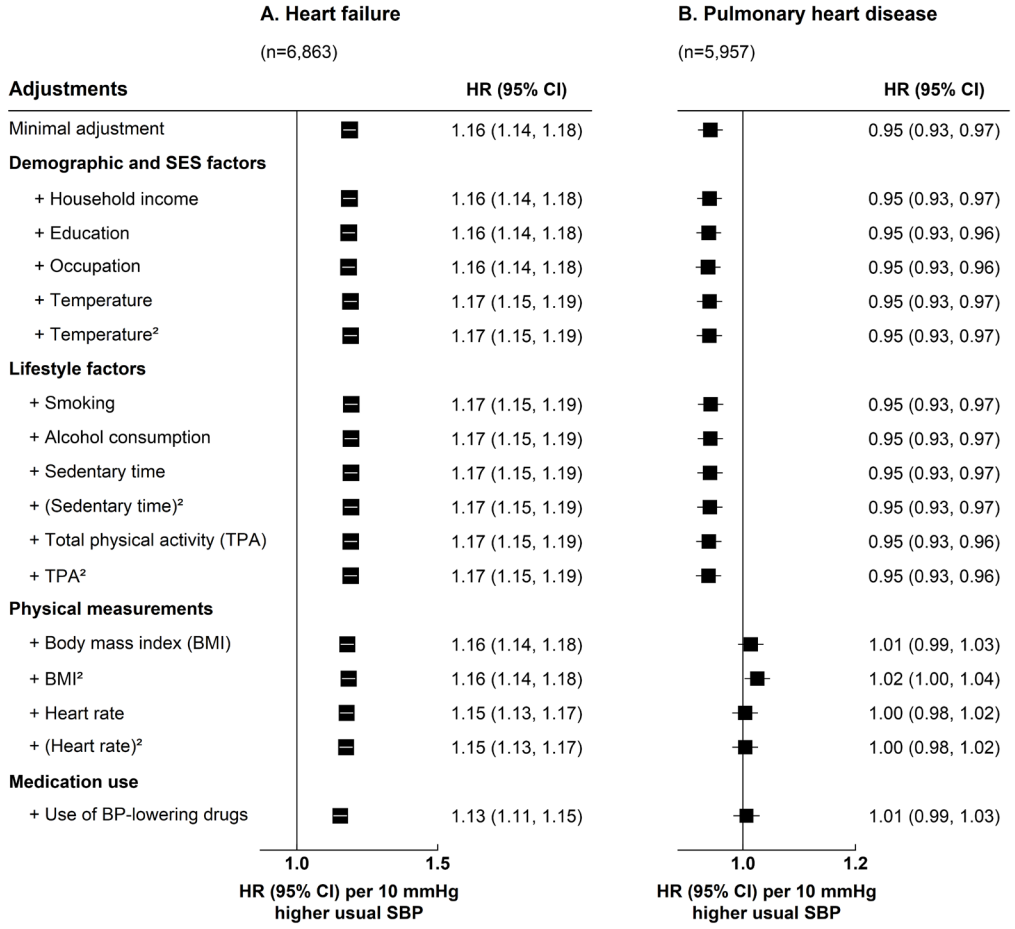


Analyses were stratified by age-at-risk, sex and study area and adjusted for baseline age, baseline age^2, income, education, occupation, temperature, temperature^2, smoking status, alcohol consumption, sedentary time, (sedentary time)^2, total physical activity (TPA), TPA^2, heart rate, (heart rate)^2, body mass index (BMI), BMI^2 and blood pressure-lowering drug use. The black squares are the quintile-specific hazard ratios (HR) weighted by the inverse of its variance. The numbers above and below the confidence intervals (CI, vertical lines) are the adjusted HRs and number of events.

5.3.3. Usual SBP with risks of HF and PHD with adjustments for confounding

After stratifying for age-at-risk, sex and all 10 study areas and adjusting for baseline age, further adjustment for demographic factors, lifestyle factors, physical measurements and use of BP-lowering drugs led to a 17.7% in the logHR of the association of usual SBP with HF **(Figure 5.3)**. By contrast, adjusting for BMI explained almost all the associations between SBP and PHD beyond demographic, socioeconomic and lifestyle factors **(Figure 5.3)**.

Figure 5.3. Associations of usual systolic blood pressure with (A) heart failure and (B) pulmonary heart disease associated with sequential adjustment.



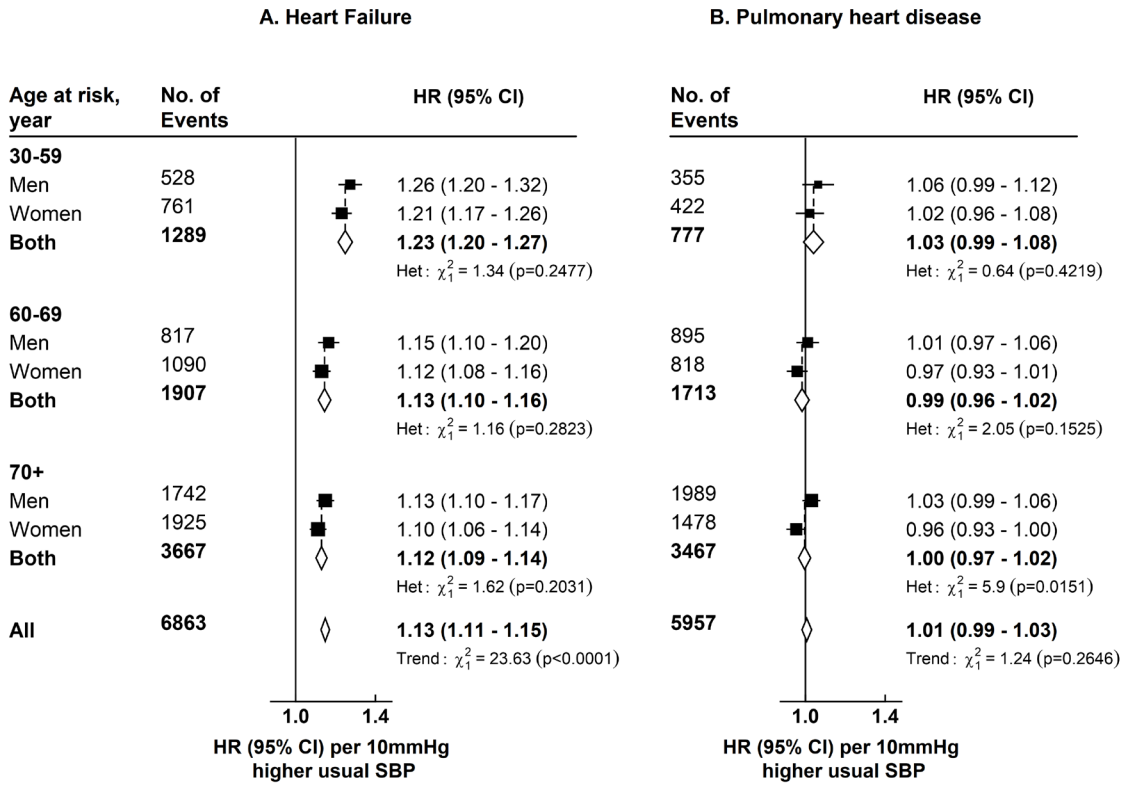
The minimally adjusted model was stratified by age-at-risk, sex and all 10 study areas and adjusted for baseline age and baseline age^2.

5.3.4. Associations of usual SBP with HF and PHD by subgroups

Figure 5.4 shows the age-sex-specific associations of usual SBP with HF and PHD. The associations with HF were similar for men and women, but the strengths of these associations were attenuated in older age groups (P_{trend}<0.0001). In the age groups 30–59, 60–69, and ≥ 80, each 10 mmHg higher usual SBP was associated with a 23% (20–27), 13% (10–16), and 12% (9–14) higher risk of HF, respectively (**Figure 5.4**). The strength of association among those ≥70 was about 45% weaker than that in the youngest age group (30–59 years) (**Figure 5.4**).

The associations of usual SBP with PHD were consistent across age and sex groups (**Figure 5.4**).

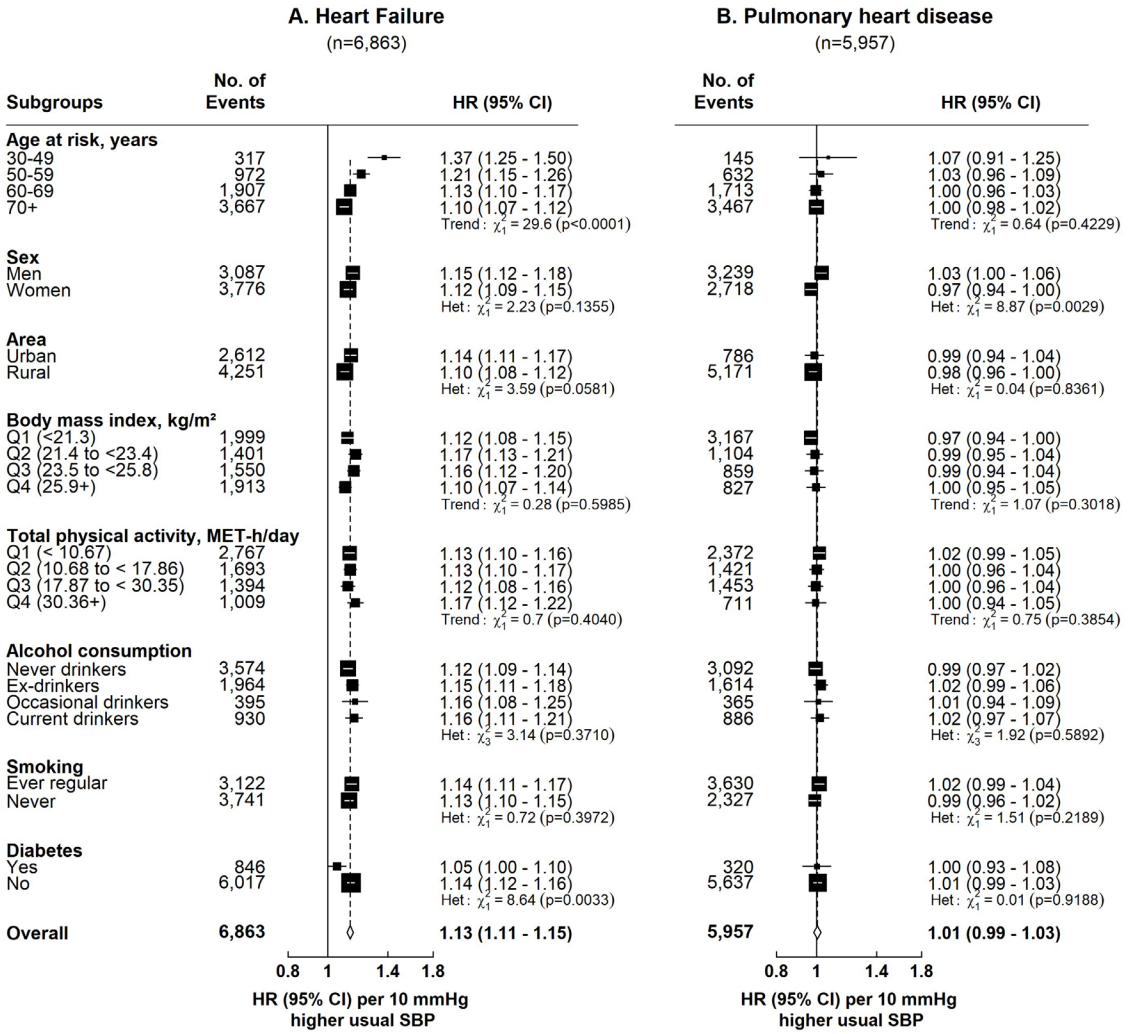
Figure 5.4. Age-sex-specific associations of usual systolic blood pressure with (A) heart failure and (B) pulmonary heart disease.



Analyses were stratified by all 10 study areas and adjusted for baseline age, baseline age^2, income, education, occupation, temperature, temperature^2, smoking status, alcohol consumption, sedentary time, (sedentary time)^2, total physical activity (TPA), TPA^2, heart rate, (heart rate)^2, body mass index (BMI), BMI^2 and BP-lowering drug use. The black square and horizontal line for each age-sex group are the inverse-variance-weighted hazard ratios (HRs) and confidence intervals. The sex-specific HRs for each age group were combined using inverse-variance-weighted fixed-effect meta-analysis.

Further subgroup analyses showed a weaker association of usual SBP with HF among participants with prevalent diabetes than in those without diabetes ($P_{\text{heterogeneity}}$ =0.0020) (**Figure 5.5**). There was no evidence of effect modification by BMI, TPA, alcohol consumption and smoking.

Figure 5.5. Association of systolic blood pressure with (A) heart failure and (B) pulmonary heart disease by subgroup.

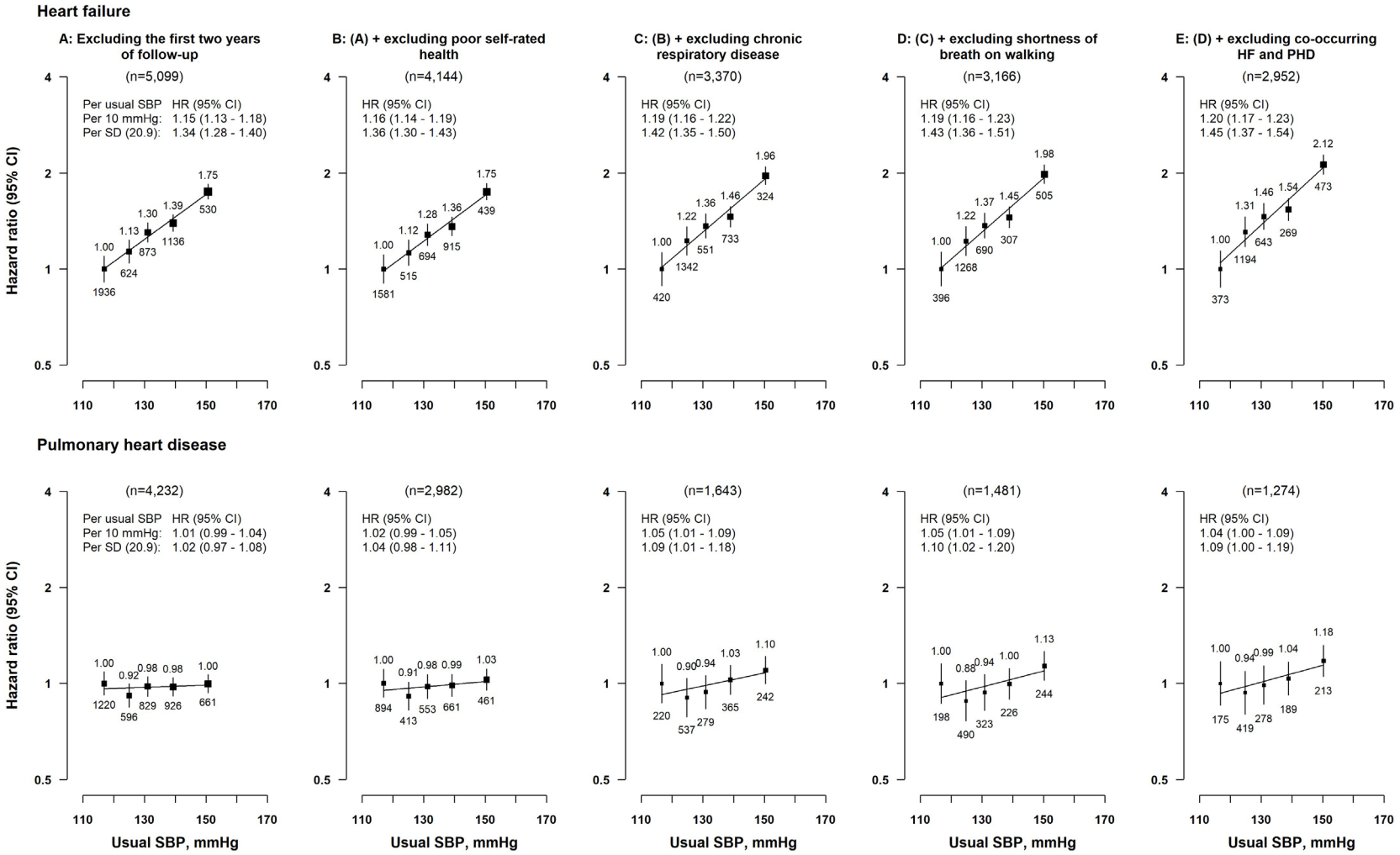


Analyses were stratified by age-at-risk, sex and study area and adjusted for baseline age, baseline age^2, income, education, occupation, temperature, temperature^2, smoking status, alcohol consumption, sedentary time, (sedentary time)^2, total physical activity (TPA), TPA^2, heart rate, (heart rate)^2, body mass index (BMI), BMI^2 and BP-lowering drug use. MET = Metabolic equivalent of task.

5.3.5. Sensitivity analyses

Sequentially excluding the first two years of follow-up and prior disease and conditions at baseline showed stronger associations of usual SBP with HF; there was, however, no strong evidence of an association of usual SBP with PHD (**Figure 5.6**).

Figure 5.6. Associations of fifths of usual systolic blood pressure (SBP) with heart failure (upper panel) and pulmonary heart disease (lower panel) with sequential exclusions.



Analyses were stratified by age-at-risk, sex and study area and adjusted for baseline age, baseline age^2, income, education, occupation, temperature, temperature^2, smoking status, alcohol consumption, sedentary time, (sedentary time)^2, total physical activity (TPA), TPA^2, heart rate, (heart rate)^2, body mass index (BMI), BMI^2 and BP-lowering drug use. The standard deviation (SD) of baseline SBP, sex-specific SBP quintiles of baseline SBP and quintile-specific mean usual SBP were recalculated for each analysis.

5.4. Discussion

5.4.1. Summary of the main findings

This chapter investigated the shape and strength of the association of usual SBP with HF and PHD in adults aged 30–79 with no prior history of heart disease, stroke or TIA and assessed the extent to which these associations were due to confounding or reverse causality. The findings demonstrated strong positive and log-linear associations of usual SBP with HF with no evidence of a threshold throughout the range of SBP studied. The strengths of these associations were similar for men and women. The associations of usual SBP with HF were attenuated with increasing age, but the association remained strong and positive even among individuals in the oldest age group. By contrast, there was no strong evidence of any association between SBP and PHD.

5.4.2. Associations of usual SBP with HF and PHD

SBP is a well-established risk factor for HF in high-income countries, but there is a lack of evidence on this association among LMIC populations. The results of the present analysis (HR per 10 mmHg higher usual SBP: 1.13, 95% CI 1.11–1.15) were similar to those reported by the APCSC (1.13, 1.06–1.21)[95], the Framingham Heart Study (1.12, 1.06–1.18)[87], a prospective cohort study of 1.25 million patients linked to healthcare databases in the UK (1.13, 1.11–1.15)[89] and a recent meta-analysis including only individuals predominantly of European ancestry (1.13, 1.10–1.16).

By contrast, the risk estimate in this study was stronger than that reported in a cohort of 5,280 Black Americans in the Jackson Heart Study[88]. It is likely that the Jackson Heart Study underestimated the strength of association between SBP and HF because it did not correct for regression dilution and adjusted for prior myocardial infarction and left ventricular hypertrophy in the multivariable analyses. Moreover, the Jackson Heart Study did not exclude prior CVD and hence could predispose the findings to reverse causality

bias. The findings of these analyses suggest that prior disease may attenuate the strength of the association between SBP and HF, as the association was stronger after sequential exclusion of participants with prior disease.

There is substantial uncertainty about the optimal level of SBP for prevention of HF and uncertainty about a possible threshold value below which lower levels of SBP might increase risk of HF. Consistent with findings in studies from high-income countries[87,89], this chapter found no evidence of a threshold or J-shaped association down to at least 115 mmHg. Two prospective cohort studies of individuals with diabetes and no prior CVD from Sweden[54] and the USA[55] reported that participants with SBP <130 mmHg had higher risks of HF compared to individuals with SBP of 130–139. The reasons for this J-shaped association in individuals with diabetes, even after excluding individuals with prior CVD, are unclear. Excluding prevalent CVD might not be sufficient to prevent reverse causality in people living with diabetes as over half of these individuals are likely to have sub-clinical CVDs that are often missed in epidemiological studies[167–169]. Hence, uncertainty persists regarding the optimum level of SBP for prevention of HF due to a lack of robust evidence.

In contrast to HF, this analysis found no evidence of an association between SBP and PHD. Sustained elevations in SBP cause HF by increasing the pressure against which the heart must overcome to supply blood to the body. The left ventricle hypertrophies to overcome the vascular resistance caused by sustained elevations of SBP[170]. However, the left ventricle eventually decompensates, leading to HF. Since the left heart is responsible for pumping against this resistance to the whole body, it is not surprising to observe higher risks of HF with high levels of SBP. In contrast, for this resistance to affect the right heart, it must first affect the left heart and the pulmonary circulation, making it implausible for SBP to cause PHD directly. Nevertheless, conventional observational studies are unable to elucidate the mechanisms underlying these

associations, and further studies are required to explore the reasons for such differences in associations between usual SBP and HF and PHD.

5.5. Summary

This chapter demonstrated a positive and log-linear association of usual SBP with incident HF with no evidence of a threshold down to at least 115 mmHg. The consistency of the shape and strength of the association across studies involving individuals from diverse populations suggests that SBP is a risk factor for HF. However, observational studies are limited by residual confounding and reverse causality. Chapter 7 will explore the extent to which the association of SBP with HF is causal. The present chapter demonstrated no evidence of any association between SBP and PHD; and there is no need for further analyses to explore any association of SBP with PHD. Chapter 6 will investigate the prospective associations of BMI and WHR with HF and PHD and assess how much of these associations are explained by SBP.

6. Associations of adiposity with risk of heart failure and pulmonary heart disease

6.1. Background

According to the WHO, obesity has developed into a pandemic affecting 1 in 8 adults worldwide. In 2019, overweight and obesity caused about five million deaths from non-communicable diseases, including HF[72].

Previous studies have reported positive associations of BMI and WHR with HF, but uncertainties persist about the shape and strength of these associations. Previous studies mainly involving individuals of European ancestry have demonstrated positive associations of BMI and WHR with HF but with a threshold below which levels of BMI or WHR no longer affect the risk of HF[97,98,104–107]. However, other studies showed a U-shaped association of BMI with HF[102]. In addition, studies conducted in non-Western countries demonstrated inverse associations of BMI with HF[79,96]. Understanding the shape of the associations of adiposity traits with HF is important to guide pharmacological and non-pharmacological strategies for prevention of HF. As demonstrated in Chapter 2 (Literature review), little is known about associations of adiposity traits with HF in non-Western populations and no study has explored the associations of adiposity with PHD.

This chapter will investigate (i) the shape and strength of the associations of BMI and WHR with HF and PHD in Chinese adults, (ii) modification of these associations by demographic, anthropometric and lifestyle factors and prior disease, (iii) the association of each adiposity measure independent of other adiposity measures, (iv) the percentage of these associations that are mediated through SBP and diabetes, and (iv) the robustness of the findings to confounding and reverse causality.

6.2. Methods

6.2.1. Study population

This analysis excluded participants with histories of rheumatic heart disease (RHD, n = 937), coronary heart disease (CHD, n=15472), stroke or transient ischaemic attack (TIA, n = 8884), cancer (n = 2578) and tuberculosis (n = 7659) at baseline.

6.2.2. Outcome definition

The outcomes evaluated in the present chapter included fatal and non-fatal incident HF (I50) and PHD (I27) for analyses of HF and PHD, respectively. Observations were censored at the first HF or PHD event, death from any cause, loss to follow-up, or the global censoring date (31 December 2018).

6.2.3. Statistical analysis

6.2.3.1. Descriptive analyses

A series of descriptive analyses with standardisation by age at baseline, sex and study area were conducted to explore how key social and demographic variables (e.g., age, sex and residence in urban and rural areas, lifestyle factors (e.g., smoking and drinking), and medical conditions (e.g., hypertension and diabetes) vary by quintiles of baseline BMI (cut-offs: 20.7, 22.5, 24.2, 26.3) and WHR (cut-offs: 0.82, 0.86, 0.90, 0.94). Categorical and quantitative variables were summarised using percentages and mean (SD) standardised by age, sex and all 10 study areas.

6.2.3.2. Association of adiposity with risks of HF and PHD

Stratified multivariable Cox proportional hazard regression models yielded hazard ratios (HRs) for the associations of sex-specific quintiles of BMI and WHR with first incident HF and PHD using time in the study as the timescale. Minimally adjusted models were stratified by age-at-risk (5-year age bands, nine categories), sex and all 10 study areas and sequentially adjusted for baseline age, baseline age^2, education (four categories: no

formal, primary, middle and high school or higher education), annual household income (four categories: <10,000, 10,000–29,999, 30,000–34,999 and ≥35,000 Yuan), occupation (two categories: agricultural vs non-agricultural workers), alcohol consumption (four categories: never, ex-, occasional, and current drinkers), smoking (three categories: never, ex-smokers/occasional and current smokers), sedentary time in hours, (sedentary time)^2, total physical activity (TPA) in metabolic task equivalent (MET)-hour, TPA^2, heart rate and (heart rate)^2.

As described in Chapter 5, potential confounders were identified based on prior knowledge. Potential confounders were assessed based on whether they led to a ≥ 10% change in the log-HR of the associations of adiposity traits with HF and PHD; no additional confounders were identified following this data-driven approach. All potential confounders selected for the present analysis were based on prior knowledge.

For continuous covariates, the LR test for non-linearity was used to assess departures from a linear trend by comparing a nested model containing the confounder in its linear form to a more complex model with an additional quadratic term of the covariate. In case of significant deviations from a linear trend, an additional quadratic term of the covariate was added to the model to account for non-linear relationships.

To explore the shape of the association of adiposity with HF and PHD and correct for regression dilution bias, the FAR method was used to calculate the variance of the log risk in each baseline-defined fifth of BMI or WHR, as described in Chapter 5. The quintile-specific FARs were then plotted against the mean levels of BMI or WHR at resurvey (usual levels).

To quantify the strength of the associations of BMI or WHR with HF and PHD, the data were restricted to the range of the adiposity trait in which the associations with HF or PHD were approximately linear. Within the range of the adiposity traits under study, the

log HRs for HF or PHD were then estimated for each 1-unit increase in the baseline adiposity trait. The log HRs (and their standard errors) were then multiplied by the SD of the adiposity trait within the range of the trait studied and divided by the RDR before exponentiating to obtain HRs per 1 SD increase in the usual values of the adiposity trait.

RDRs were estimated using Rosner's regression method with adjustments for age at baseline, sex and residence in urban and rural areas[128]. In age-sex-specific analyses, the log HRs for each age-sex group were corrected for regression dilution using age-sex-specific RDRs and then combined using inverse variance-weighted fixed-effect meta-analysis to obtain the age-specific HRs.

6.2.3.3. Quantifying the effects of confounding

A change in the percentage of log-HR of the associations between the adiposity traits with HF and PHD before and after adjusting for confounding were used to quantify the impact of adjusting for confounding, as described in Chapter 5.

6.2.3.4. Subgroup analysis

In subgroup analyses, the LR test was used to evaluate effect modification by age-at-risk, sex, urbanicity, alcohol consumption, smoking, SBP and diabetes by comparing nested models with and without an additional interaction term.

6.2.3.5. Independence of adiposity measures

To investigate the independence of associations from other adiposity measures, the associations of BMI with HF or PHD were further adjusted for WHR; the associations of WHR with HF and PHD were additionally adjusted for BMI. The impact of adjusting for other measures of adiposity on the associations of BMI and WHR with HF and PHD was estimated using percentage change in the logHR before and after adjustment.

6.2.3.6. Adjustments for intermediate factors

The LR test was extended to investigate the mediation of the associations of BMI and WHR with HF and PHD. An intermediate factor was defined as a variable that is associated with both the exposure and outcome and lies on the pathway linking the exposure to the outcome[171]. The percentage of the associations of BMI and WHR with HF and PHD explained by SBP and diabetes was calculated as the percentage change in the log-HR before and after adjusting for these intermediate factors[172,173].

6.2.3.7. Sensitivity analyses

Sensitivity analyses were conducted to evaluate the robustness of the findings to reverse causality. These analyses involved sequentially excluding (1) the first two years of follow-up, (2) poor self-reported health at baseline, (3) shortness of breath at baseline, (4) prevalent chronic lung diseases, including COPD, tuberculosis, emphysema, bronchitis, and asthma, and (5) co-occurring HF and PHD events.

The Cox proportionality assumption was assessed by visual inspection of the correlation of scaled Schoenfield residuals with the natural log of survival time, and the assumption was satisfied. All hypothesis tests were two-tailed, and two-tailed p-values less than 0.05 were considered statistically significant. Data were analysed using R version 4.2.2[163].

6.3. Results

6.3.1. Baseline characteristics of the study participants

A total of 479,553 of the 512,724 recruited at baseline were included in this analysis. In total, 33,171 participants with prevalent RHD (n=937), CHD (n=15,472), stroke or TIA (n=8,884), cancer (n=2,578) and tuberculosis (n=7,659) were excluded from this analysis.

Table 6.1 shows the baseline characteristics of the participants included in this analysis. The mean (SD) age of the participants was 51.6 (10.6); 59% were female, 57% resided in rural areas and 57% were either agricultural or factory workers (**Table 6.1**). About 86% of men (compared to 4.9% of women) were ever-regular smokers, while 34% of men (compared to 2.1%) of women were currently alcohol drinkers. About 34%, 5.4% and 7% of the participants had hypertension, diabetes and COPD.

The mean (SD) levels of BMI and WHR at baseline were 23.6 (3.4) and 0.9 (0.1), respectively. The mean levels of BMI for men and women were 23.4 (3.2) and 23.8 (3.4); the mean WHR for men and women were 0.9 (0.1) and 0.9 (0.1), respectively. The mean levels of BMI for urban and rural participants were 24.2 (3.4) and 23.1 (3.3); the mean levels of WHR for urban and rural individuals were 0.9 (0.1) and 0.9 (0.1), respectively.

Analysis of trends of baseline characteristics by quintiles of BMI indicated that participants with higher quintiles of SBP were older and more likely to be men and current drinkers and were more likely to reside in rural than urban areas and have diabetes and higher levels of RHR (**Table 6.1**).

Figure 6.1 shows the age-specific mean BMI and WHR by sex and area of residence. The mean BMI of urban women increased sharply from ~22 kg/m^2 among participants who were 30 years old, reached a threshold of ~25 kg/m^2 among individuals who were 60 years old, and then declined to a mean BMI of 23 kg/m^2 among participants aged 80

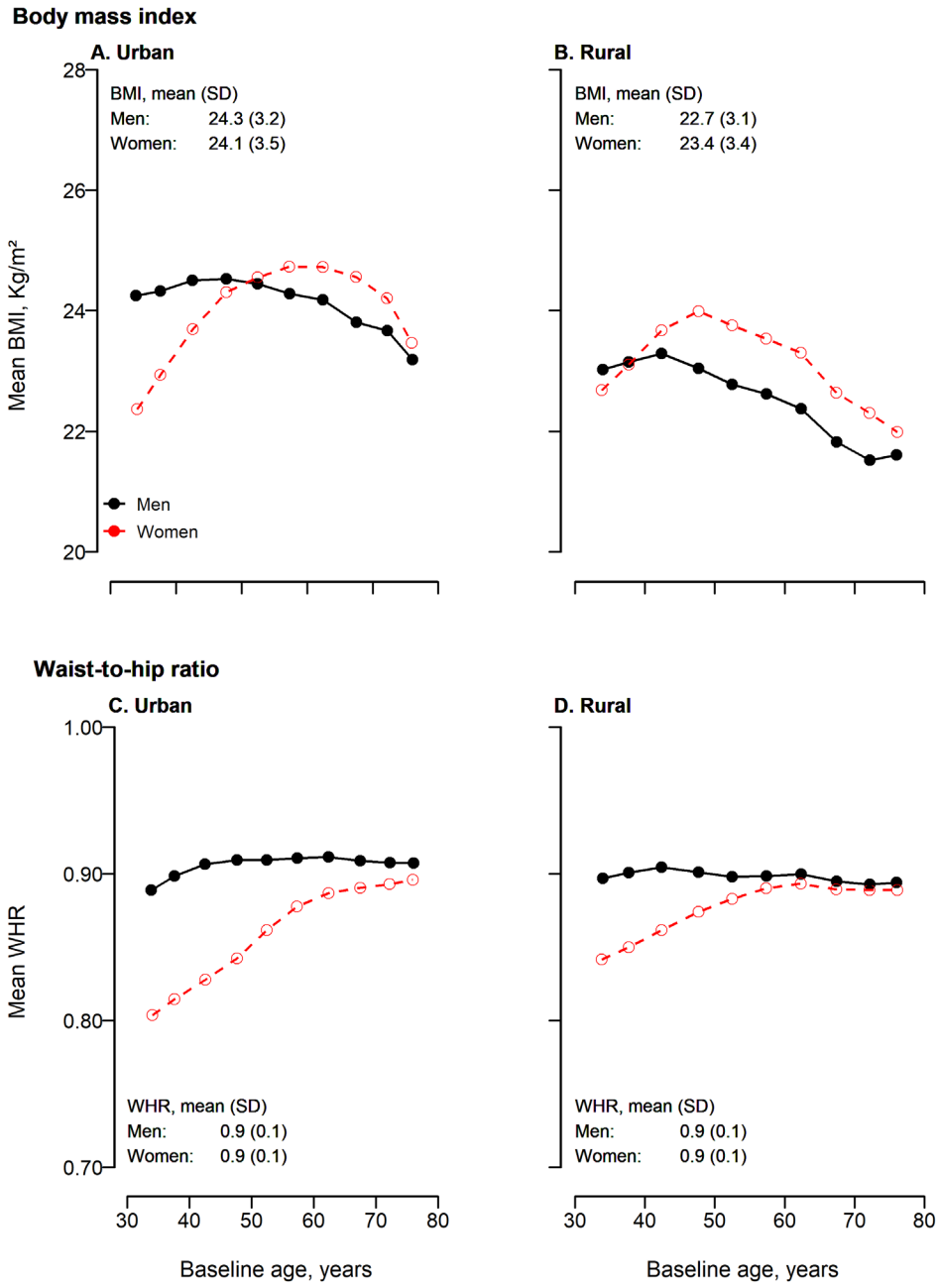
years old. By contrast, the mean BMI for urban men decreased steadily from 24 kg/m^2 in participants who were 30 years old to 23 kg/m^2 in individuals who were 80 years old (**Figure 6.1**). Similar to urban women, the mean BMI of rural women increased steeply from 22.5 kg/m^2 among those who were 30 years old and reached a threshold at 24 kg/m^2 among women aged 50 years old before decreasing sharply to 22 kg/m^2 in women who were 80 years old (**Figure 6.1**). The mean BMI of rural men decreased steadily from 23 kg/m^2 among 30-year-old men to 22 kg/m^2 among 80-year-olds (**Figure 6.1**).

Table 6.1. Baseline characteristics of study participants by fifths of baseline body mass index (BMI) and waist-to-hip ratio (WHR).

Characteristics	Fifths of BMI, kg/m² (Range)					Fifths of WHR (Range)					All
	Q1 (<20.7)	Q2 (20.8 to <22.5)	Q3 (22.6 to <24.2)	Q4 (24.3 to <26.3)	Q5 (26.4+)	Q1 (<0.82)	Q2 (0.83 to <0.86)	Q3 (0.87 to <0.90)	Q4 (0.91 to <0.94)	Q5 (0.95+)	
No. of participants	98,264	97,655	97,120	92,175	94,339	114,063	103,783	82,613	92,029	87,065	479,553
BMI, kg/m², mean (SD), mmHg	19.4 (1.2)	21.7 (0.5)	23.4 (0.5)	25.3 (0.6)	28.5 (2.1)	21.0 (2.7)	22.7 (2.6)	23.8 (2.7)	24.9 (2.9)	26.5 (4.1)	23.6 (3.4)
WHR, mean (SD), mmHg	0.83 (0.06)	0.86 (0.05)	0.88 (0.05)	0.91 (0.05)	0.94 (0.06)	0.80 (0.03)	0.85 (0.01)	0.89 (0.01)	0.92 (0.01)	0.98 (0.08)	0.88 (0.07)
Demographic and SES factors, %											
Age, mean (SD), years	51.9 (12.1)	50.8 (10.6)	51.0 (10.1)	51.4 (10.0)	51.5 (10.5)	49.5 (11.0)	50.3 (10.1)	51.3 (10.1)	52.5 (10.2)	54.8 (10.6)	51.5 (10.5)
Women	61.0	59.5	58.4	59.7	60.3	59.9	57.8	57.5	57.6	61.9	59.3
Rural area	68.0	62.2	56.6	51.5	46.6	54.3	56.6	56.7	58.5	64.2	57.0
Married	89.7	90.5	91.0	91.7	91.9	90.1	90.9	91.1	91.2	91.3	90.9
Household monthly income <10,000 Yuan	31.3	29.1	27.7	26.7	26.3	30.1	29.0	28.2	27.8	27.4	28.3
Agriculture or factory worker	61.9	60.1	57.8	55.3	51.8	61.1	59.1	57.6	56.1	54.3	57.8
Lifestyle factors, %											
Ever regular smoker, Men	80.0	76.4	73.6	71.2	70.8	75.5	74.0	73.9	73.5	75.4	74.5
Ever regular smoker, Women	3.9	3.0	3.0	2.6	2.9	3.3	3.1	3.1	3.0	3.5	3.1
Current drinkers, Men	32.7	34.9	34.6	33.9	33.4	29.8	32.8	34.5	36.5	37.2	34.0
Current drinkers, Women	2.0	2.1	2.2	2.1	2.0	1.9	2.1	2.2	2.1	2.2	2.1
Total physical activity, mean (SD), MET-hour/day	22.1 (13.3)	22.2 (12.0)	21.8 (11.7)	21.3 (12.0)	20.5 (12.9)	22.5 (13.2)	22.0 (11.9)	21.7 (12.0)	21.2 (12.4)	20.5 (14.4)	21.6 (13.9)
Sedentary time, mean (SD), hours/day	2.9 (1.6)	2.9 (1.4)	3.0 (1.4)	3.1 (1.4)	3.1 (1.6)	2.9 (1.6)	3.0 (1.4)	3.0 (1.4)	3.0 (1.5)	3.1 (1.8)	3.0 (1.5)
Physical measurements, mean (SD)											
Waist circumference, cm	69.7 (5.8)	75.4 (4.9)	79.8 (4.9)	84.4 (5.2)	91.8 (7.4)	70.1 (5.6)	76.4 (5.0)	80.7 (5.2)	84.9 (5.9)	91.5 (8.9)	80.1 (9.7)
Resting heart rate, bpm	78.9 (13.5)	78.1 (11.7)	78.4 (11.4)	79.0 (11.8)	80.3 (13.3)	77.8 (12.9)	78.2 (11.6)	78.7 (11.6)	79.4 (12.1)	80.9 (14.4)	78.9 (11.8)
Prevalent conditions											
Hypertension	20.5	26.7	32.2	39.1	50.3	24.0	29.6	34.1	38.5	45.5	33.7
Diabetes	3.3	4.2	5.2	6.2	8.1	2.3	3.6	4.8	6.4	10.4	5.4
COPD	9.5	7.1	6.3	5.8	5.7	8.1	7.3	6.6	6.5	6.6	7.0
Emphysema or bronchitis	3.1	2.1	2.2	2.2	2.5	2.6	2.3	2.2	2.4	2.8	2.4
Asthma	0.6	0.4	0.4	0.5	0.6	0.4	0.5	0.5	0.5	0.7	0.5
Self-reported poor health	11.8	8.9	8.0	8.1	9.1	10.3	9.0	8.6	8.8	9.7	9.2

Means and proportions were directly standardised by the age, sex and all 10 CKB areas.
Bpm: beats per minute; COPD: Chronic obstructive pulmonary disease; MET: Metabolic Equivalent Task; SD: Standard deviation; Q1–5: First to fifth quintile

Figure 6.1. Distribution of age-specific mean levels of baseline body mass and waist-to-hip ratio by sex and area of residence.

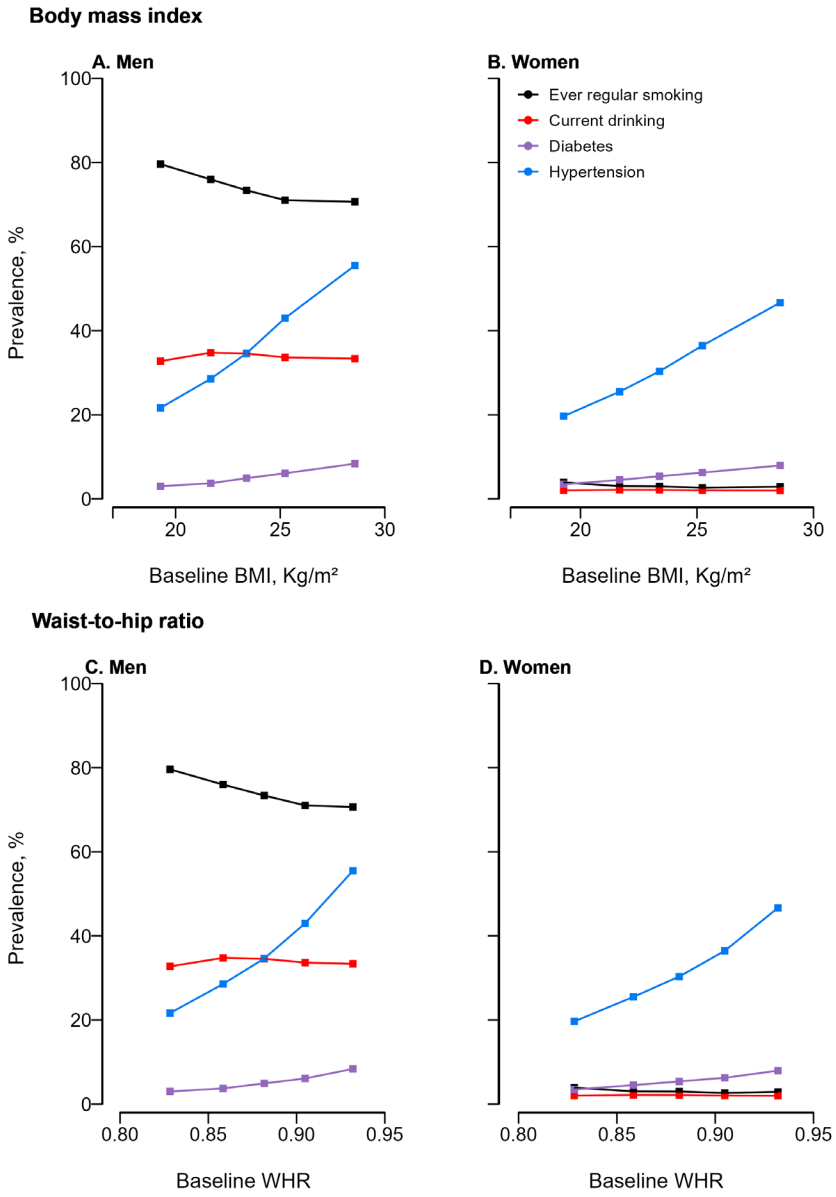


The mean WHR for urban and rural women increased from 0.8 among 30-year-old women to 0.9 in 80-year-old women. By contrast, the mean WHR for urban and rural men was constant at around 0.9 throughout the age range studied (**Figure 6.1**).

Figure 6.2 shows similar trends in the prevalence of ever-regular smoking, current drinking, hypertension and diabetes by quintiles of baseline BMI and WHR for men and

women. For men, the prevalence of ever-regular smoking was highest (80%) among those in the lowest quintile of BMI (~19 kg/m^2) and decreased steadily ~70% among those in the highest quintile (~28 kg/m^2) (**Figure 6.2**).

Figure 6.2. Prevalence of smoking, drinking, hypertension and diabetes by sex and quintiles of body mass index and waist-to-hip ratio.



The prevalence estimates were standardised by age and all ten study areas.

In addition, the prevalence of hypertension increased steeply from ~20% among those in the lowest BMI quintile to ~58% among those in the highest BMI quintile. The prevalence

of current drinking did not vary by BMI levels. Finally, the prevalence of diabetes increased modestly with increasing levels of BMI. Similar trends were observed for levels of WHR (**Figure 6.2**). For women, the trends in the prevalence of hypertension and diabetes were similar to those in men, with the exception of ever-regular smoking and current drinking that had a negligible prevalence in women (**Figure 6.2**).

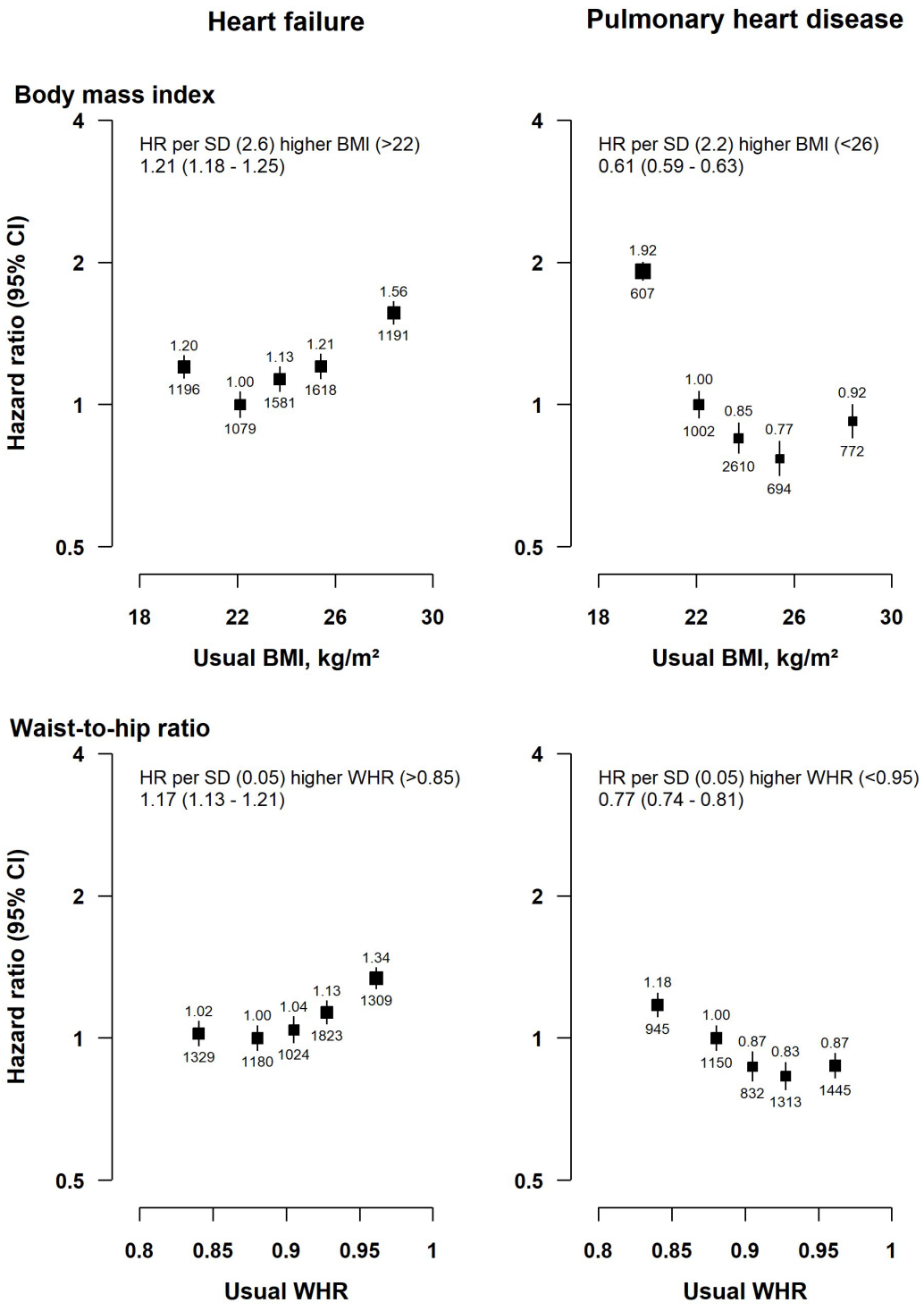
6.3.2. Associations of usual BMI and WHR with HF and PHD

During a median of 10 years of follow-up, 6,665 and 5,685 cases of HF and PHD were recorded, respectively. The mean (SD) age at diagnosis of HF was 65.1 (10.4) years, whereas that for PHD was 66.7 (9.0) years.

After adjusting for potential confounders, BMI and WHR had a J-shaped association with HF. (**Figure 6.3**). Higher levels of BMI and WHR were positively log-linearly associated with higher risks of HF down to a threshold of 22 kg/m^2 for BMI and 0.86 for WHR. Below this threshold, higher levels of BMI were associated with lower risks of HF, whereas higher levels of WHR were not associated with HF. At BMI levels >22 kg/m^2, each 1 SD (2.6 kg/m^2) higher level of BMI was associated with a 21% higher risk of HF (HR 1.21; 95% CI, 1.18–1.25, p<0.0001). At WHR levels >0.85, each 1 SD (0.05) higher WHR was associated with a 17% higher risk of HF (1.17, 1.13–1.21, p<0.0001).

In contrast, BMI and WHR had a reverse J-shaped association with PHD. Higher levels of BMI and WHR were associated with lower risks of PHD up to a threshold of ~24.8 kg/m^2 for BMI and ~0.94 for WHR (**Figure 6.3**). Beyond these thresholds, higher levels of BMI were associated with higher risks of PHD, whereas WHR was not associated with PHD (**Figure 6.3**). For BMI levels <26 kg/m^2, each 1 SD (2.6 kg/m^2) higher level of BMI was associated with a 39% lower risk of PHD (0.61, 0.59–0.63, p<0.0001). For WHR levels <0.95, each 1 SD (0.05) higher WHR was associated with a 28% lower risk of HF (0.77, 0.74–0.81, p<0.0001).

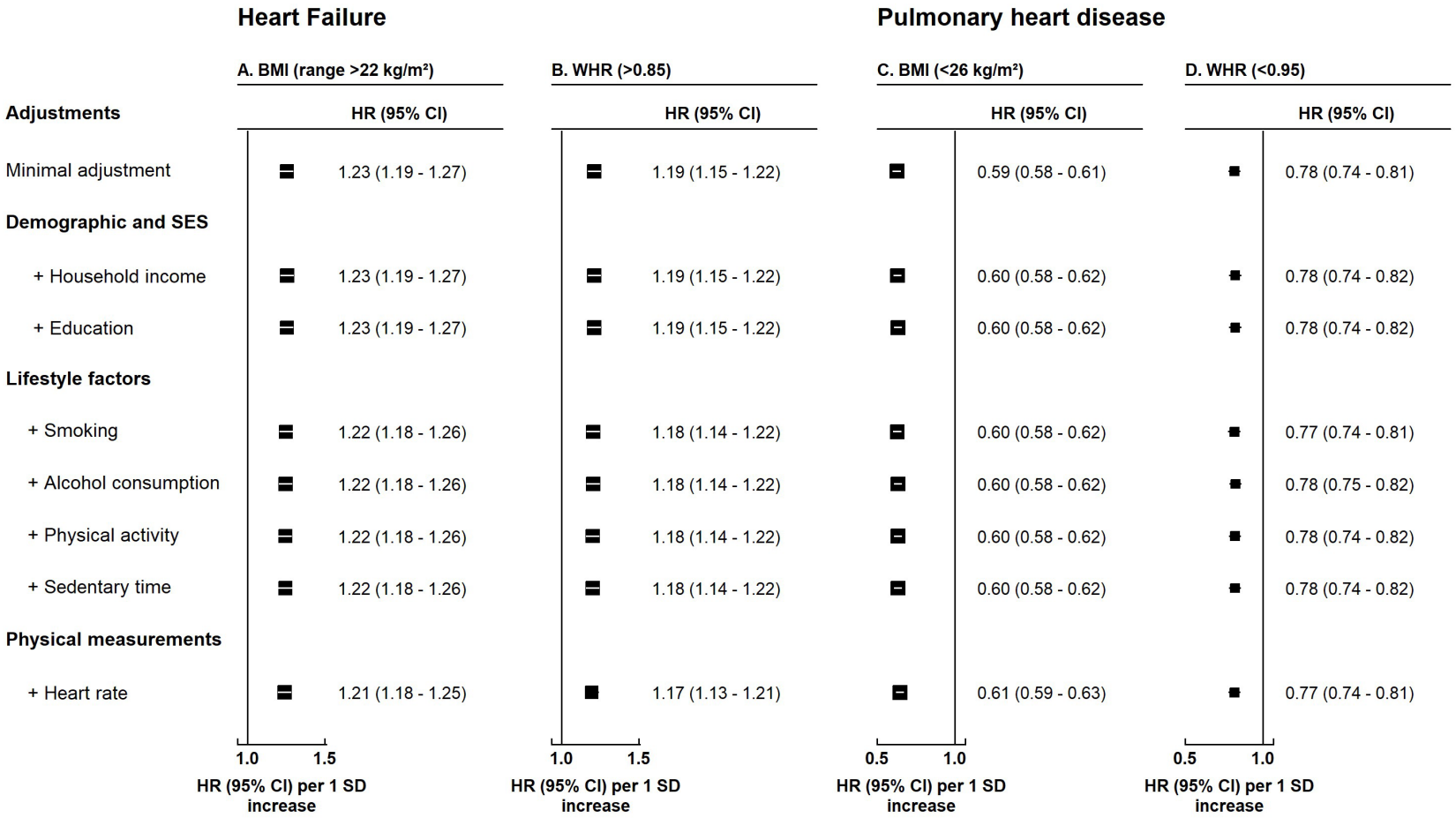
Figure 6.3. Associations of fifths of usual body mass index and waist-to-hip ratio with heart failure and pulmonary heart disease.



Analyses were stratified by age-at-risk, sex and study area and adjusted for baseline age, baseline age^2, income, education, occupation, smoking, alcohol drinking, sedentary time, TPA and heart rate. The black squares are the quintile-specific floating absolute risk (FAR) weighted by the inverse of its variance. The numbers above and below the FAR confidence intervals (CIs, vertical line) are the FARs and number of events.

After accounting for age, sex and study area further adjustments did not materially change the associations of BMI and WHR with HF and PHD (**Figure 6.4**). The percentage changes in the logHR for the associations of BMI and WHR with HF were 7.9% and 9.7%, respectively. For PHD, the logHR for the associations with BMI reduced by 6.3% but increased by 5.2% for WHR (**Figure 6.4**).

Figure 6.4. Associations of usual body mass index and waist-to-hip ratio with heart failure and pulmonary heart disease with sequential adjustment.



The minimally adjusted model was stratified by age-at-risk, sex and all 10 CKB areas and adjusted for baseline age and baseline age^2.

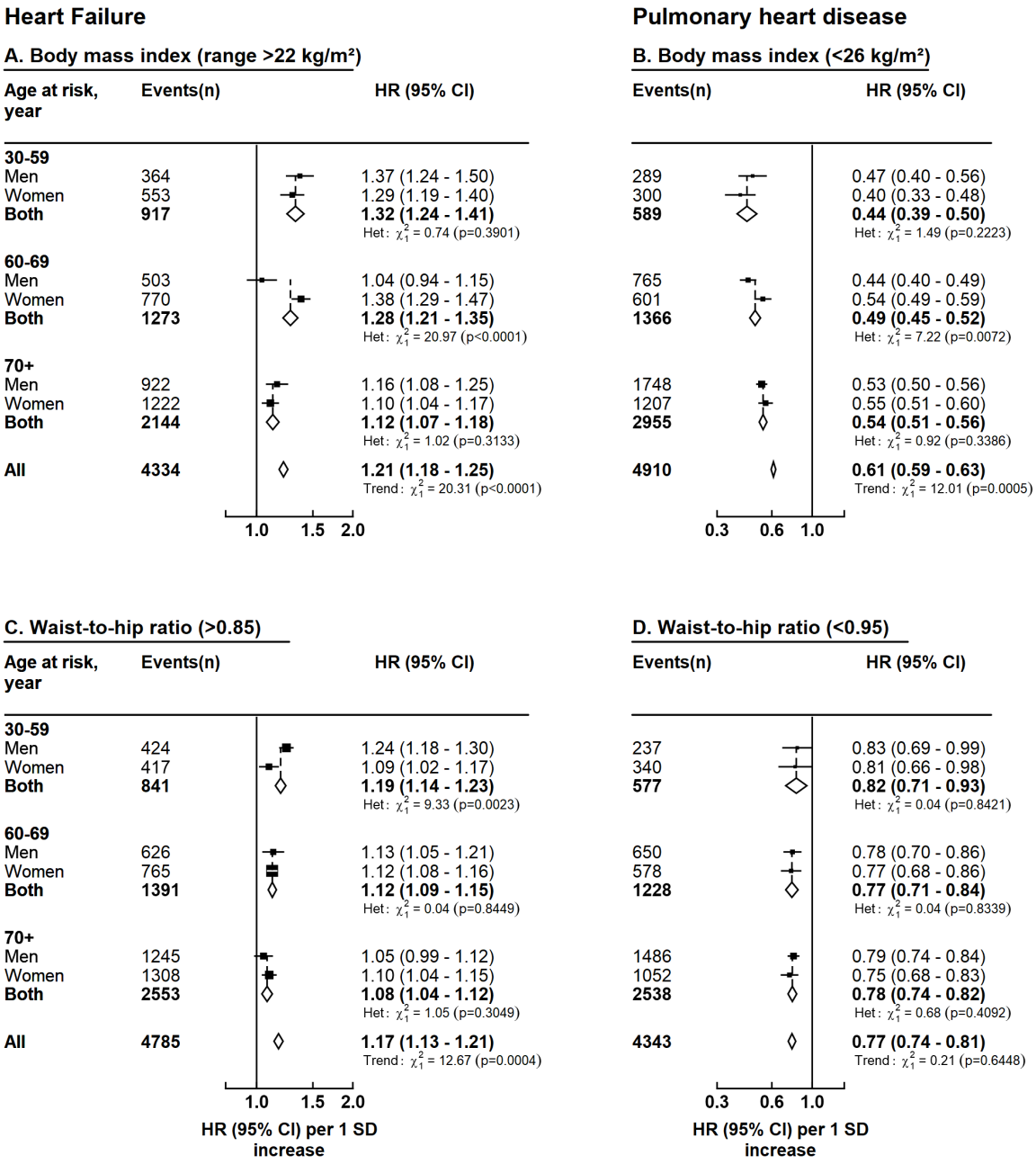
6.3.3. Association of usual BMI and WHR with HF and PHD by subgroup

Figure 6.5 shows the age-sex-specific associations of usual BMI and WHR with HF and PHD. The age-specific relative risks of HF for BMI and WHR were weaker at older ages (P_{trend}<0.0001), although the absolute risk was higher at older ages. Each 1 SD higher usual levels of BMI and WHR was associated with higher risks of HF in the following age groups: 30–59 years (BMI: HR 1.33 [95% CI 1.25–1.41]; WHR: 1.19 [1.15–1.23]), 60–69 years (BMI: 1.27 [1.21–1.34]; WHR: 1.12 [1.09–1.15]) and ≥80 years (BMI: 1.12 [1.07–1.17]; WHR: 1.08 [1.04–1.12]) (**Figure 6.5**).

Broadly, the age-specific relative risks were similar for men and women, except for the 60–69-year age group, where women (1.37, 1.29–1.46) had a higher risk of HF for per 1 SD increase in usual BMI compared to men (1.03, 0.93–1.14) (**Figure 6.5**). In contrast, among those in the 30–59-year age group, the risk of HF per 1 SD higher usual SD was greater in men (1.25, 1.19–1.31) than in women (1.11, 1.04–1.17) (**Figure 6.5**).

The age-specific associations of usual BMI with PHD attenuated with age, but there was no evidence of modification of the association of usual WHR and PHD with age (**Figure 6.5**). The risk of PHD per 1 SD increase in usual BMI was about 14% weaker among participants in the 30–59 years group (0.44, 0.38–0.50) compared to participants in the 70+ years group (0.52, 0.50–0.55) (**Figure 6.5**). The age-specific associations of usual BMI and WHR with PHD were similar for men and women, except for the 60–69 year age group in which men (0.42, 0.39–0.47) had a stronger risk for each higher SD increase in BMI compared to women (0.51, 0.46–0.57) (**Figure 6.5**).

Figure 6.5. Age-sex specific associations of usual body mass index and waist-to-hip ratio with heart failure and pulmonary heart disease.

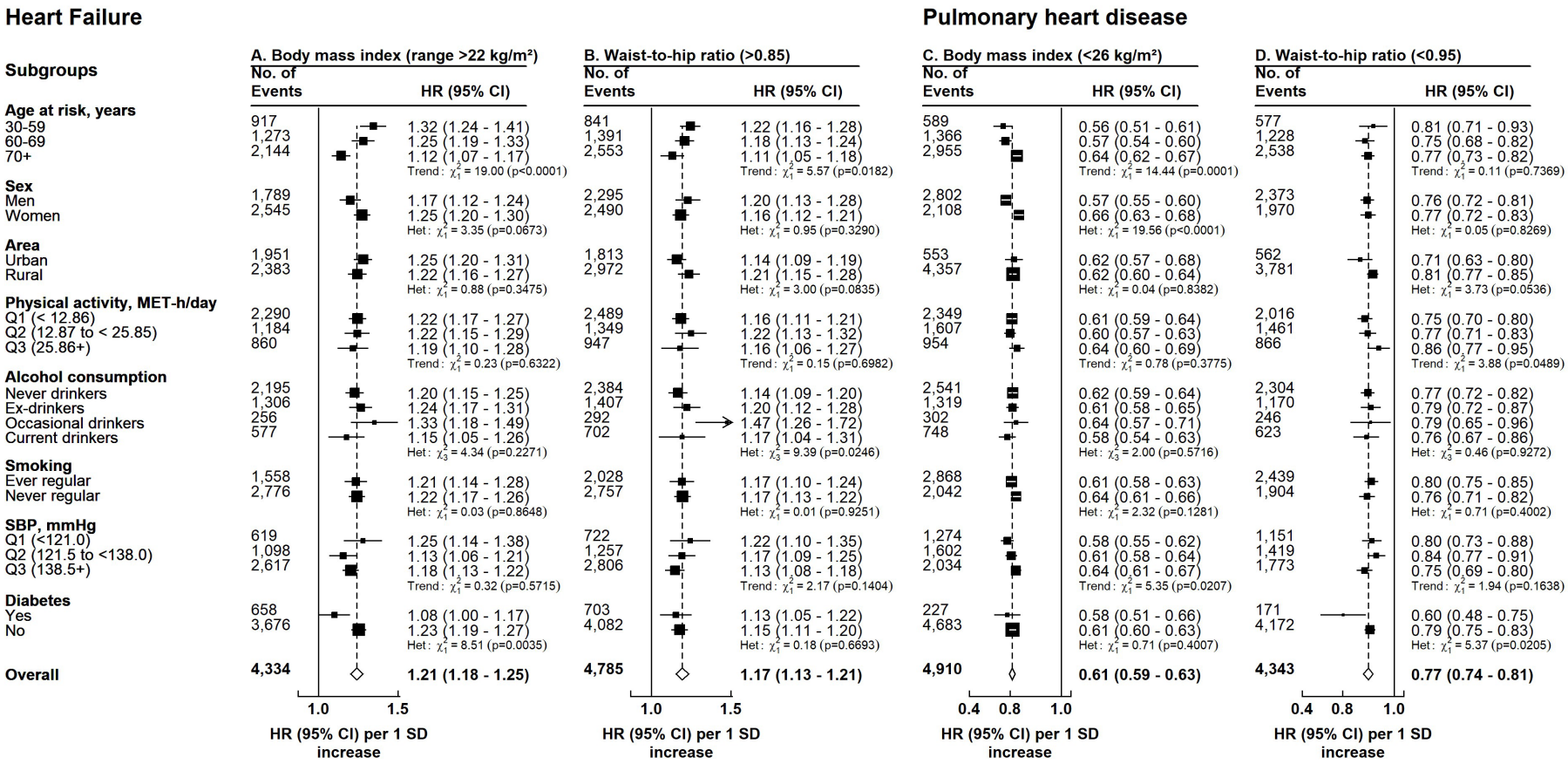


All analyses were stratified by study area and adjusted for baseline age, baseline age^2, income, education, occupation, smoking, alcohol consumption, sedentary time, TPA and heart rate. The black square and horizontal lines are the hazard ratios weighted by the inverse of their variance and the 95% confidence intervals (CIs), respectively. The standard deviations (SDs) for each adiposity trait were recalculated for each age-sex stratum. Het = Heterogeneity test for sex.

Figure 6.6 shows the associations of BMI and WHR with HF and PHD by additional subgroups. The association of BMI with HF was weaker in participants with diabetes (1.08, 1.00–1.17) than in those without diabetes (1.23, 1.19–1.27) at baseline ($P_{heterogeneity}$ = 0.0035). The association of WHR with HF was strongest among occasional drinkers (1.47, 1.26–1.72) compared to never (1.14, 1.09–1.20), ex- (1.20, 1.12–1.28) and current drinkers (1.17, 1.04–1.31) ($P_{heterogeneity}$ = 0.0246). There was no evidence of modification of the association of BMI and WHR with HF by sex, area, physical activity, smoking status and SBP (**Figure 6.6**). In addition, there was no evidence of modification of the association of WHR with HF by diabetes.

For PHD, the association of BMI with HF was weaker with higher levels of SBP (P_{trend} = 0.0207) and was stronger in men (0.57, 0.55–0.60) than in women (0.66, 0.63–0.68) ($P_{heterogeneity}$ <0.0001) (**Figure 6.6**). The association of usual WHR and PHD was weaker with higher levels of physical activity (P_{trend} = 0.0489) and was weaker in participants with diabetes (0.60, 0.48–0.75) than in those without diabetes (0.79, 0.75–0.83) (**Figure 6.6**).

Figure 6.6. Associations of body mass index and waist-to-hip ratio with heart failure and pulmonary heart disease by subgroup.

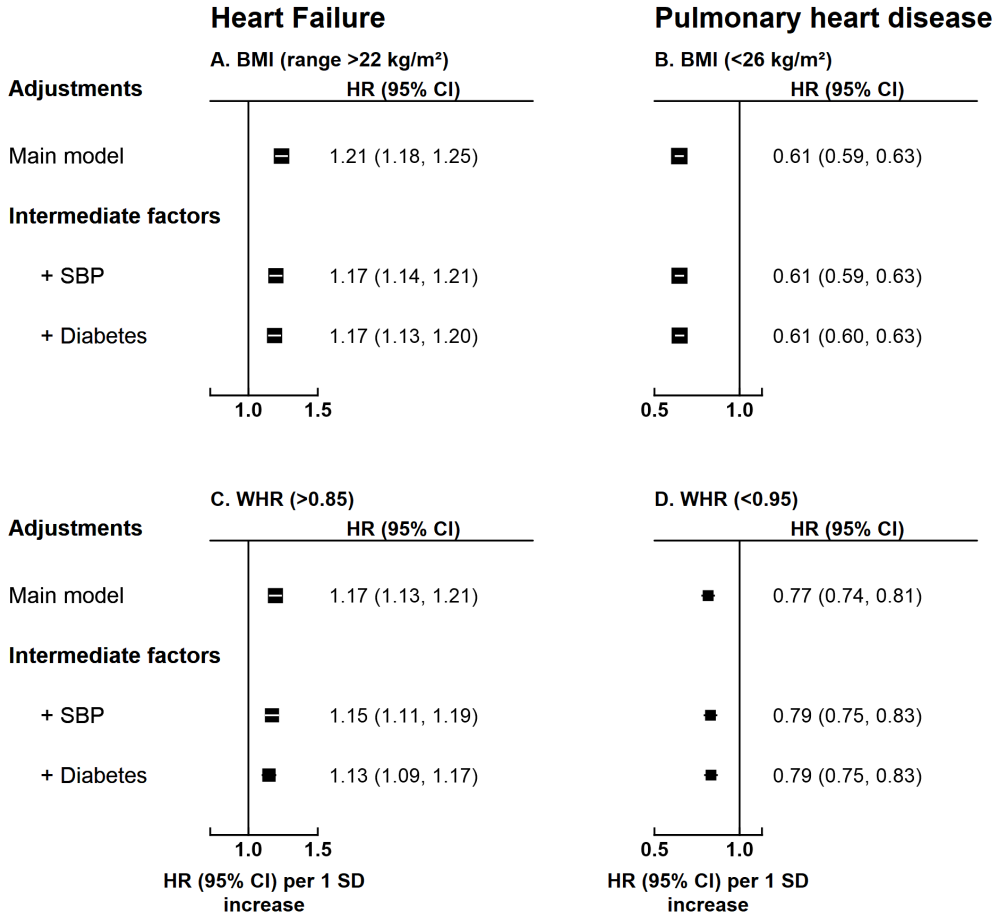


Where applicable, analyses were stratified by age-at-risk, sex, and study area and adjusted for baseline age, baseline age², income, education, occupation, smoking status, alcohol, sedentary time, physical activity, and heart rate
Diabetes included self-reported and screen-detected diabetes. SBP = systolic blood pressure, Het = Heterogeneity, MET = Metabolic equivalent of task

6.3.4. Associations of BMI and WHR with HF and PHD with adjustments for intermediate factors

Figure 6.7 shows the associations of BMI and WHR with HF and PHD with further adjustments for SBP and diabetes. For BMI, further adjustments for SBP and diabetes reduced the logHR by 17.6% for the association with HF but did not alter the association with PHD (**Figure 6.7**). For WHR, additional adjustments for SBP and diabetes reduced the logHR by 22.2% for the association with HF and by 1% for the association with PHD (**Figure 6.7**).

Figure 6.7. Associations of body mass index and waist-to-hip ratio with heart failure and pulmonary heart disease with adjustments for intermediate factors.



The main model was stratified by age-at-risk, sex, and all 10 study areas and adjusted for baseline age, baseline age^2, income, education, occupation, smoking status, alcohol, sedentary time, physical activity, and heart rate. SBP: Systolic blood pressure; SD: standard deviation.

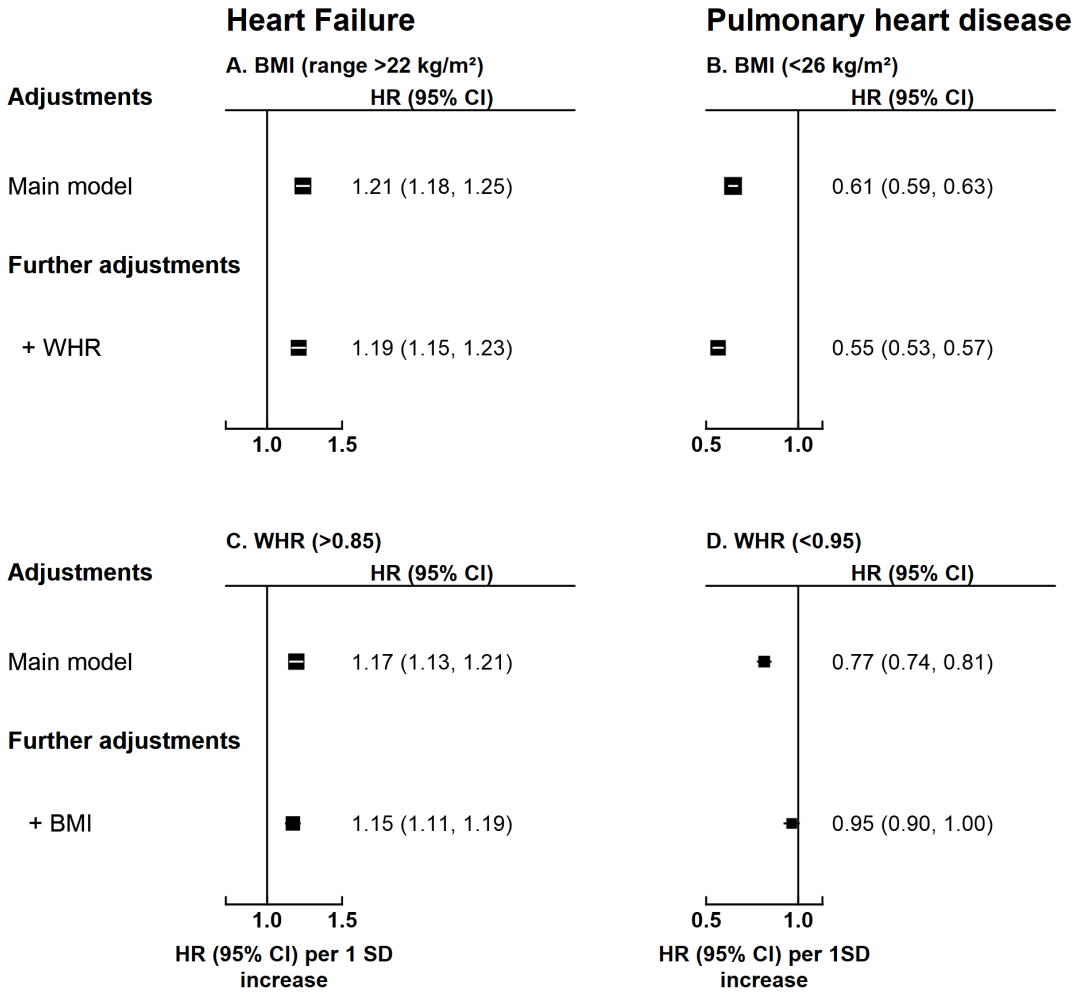
6.3.5. Association of BMI and WHR with HF and PHD independent of other adiposity traits

In order to investigate the independence of the associations of BMI and WHR with HF and PHD from other adiposity measures, the associations from the fully-adjusted Cox regression models were mutually for BMI and WHR (**Figure 6.8**). Further adjustments for WHR resulted in 8.7% reduction in the logHR of the association of BMI with HF. Additional adjustments for BMI reduced the logHR of the association of WHR with HF by 27.8% (**Figure 6.8**). Regarding PHD, further adjustments for WHR increased the logHR by 20.9%. By contrast, further adjustment for BMI completely attenuated the association between WHR and PHD (**Figure 6.8**).

6.3.6. Sensitivity analyses

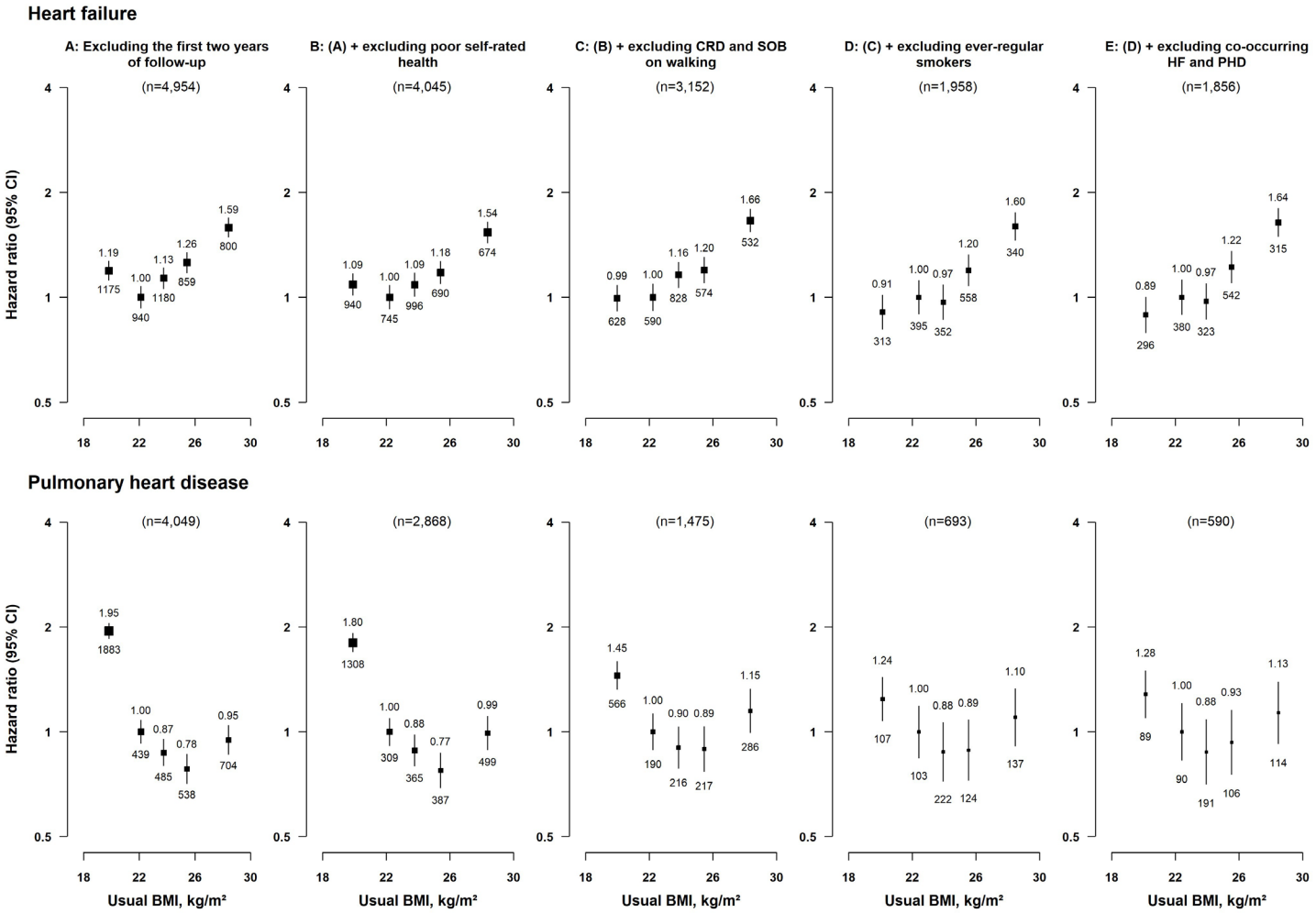
Sensitivity analyses involving sequential exclusions of the first few years of follow-up, individuals with poor self-rated health and chronic conditions and ever-regular smokers revealed a stronger association of BMI with HF (**Figure 6.9**). Although a threshold was observed at a BMI level of 22 kg/m^2, BMI was not associated with a higher risk of HF below this threshold as observed in the main analysis. By contrast, the association between BMI and PHD attenuated substantially towards the null. The findings for the associations of WHR with HF and PHD were similar to those observed for BMI after making the same exclusions (**Figure 6.10**).

Figure 6.8. Associations of BMI and WHR with HF and PHD with mutual adjustment for each adiposity trait.



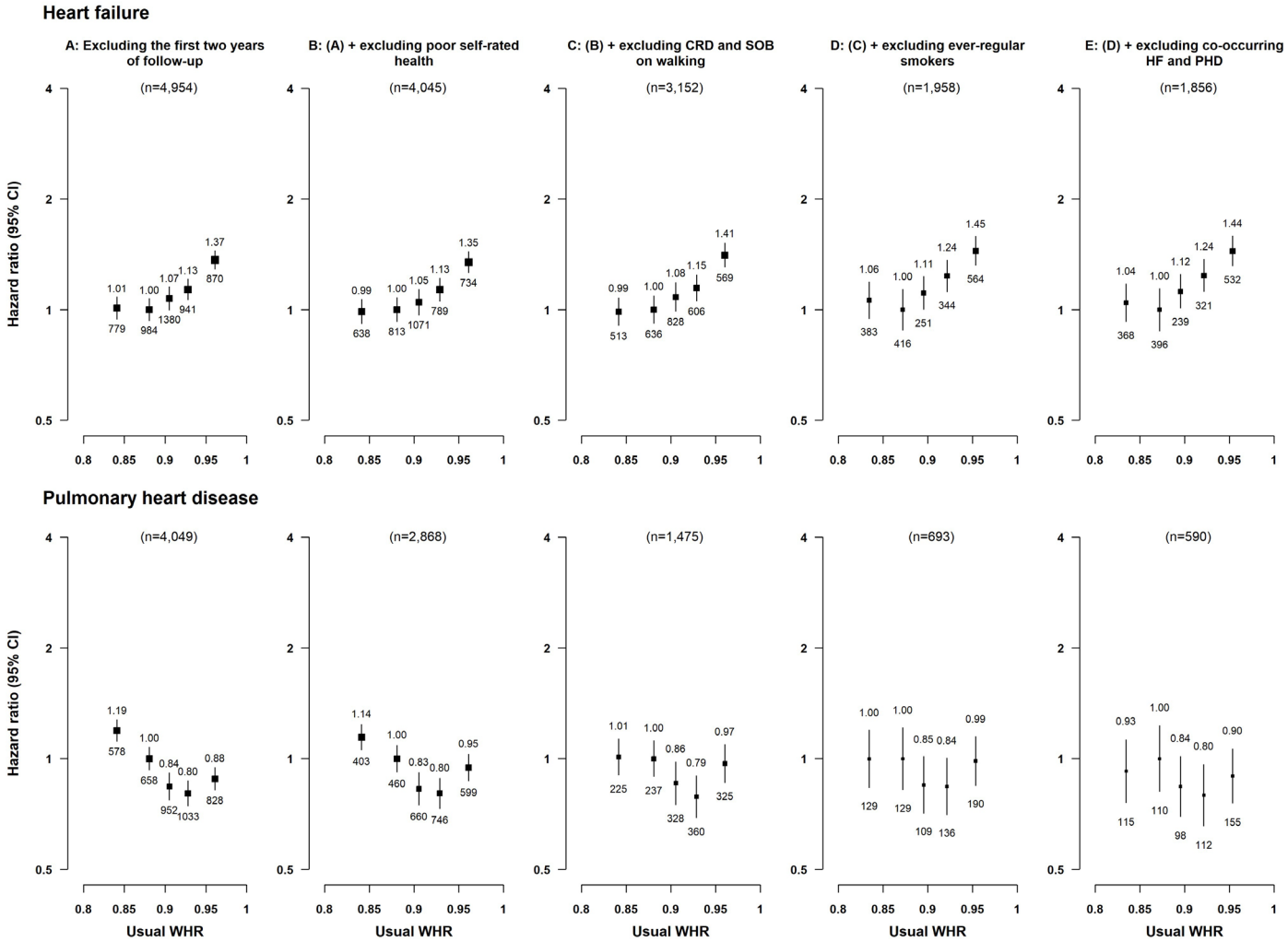
The main model was stratified by age-at-risk, sex, and all 10 study areas and adjusted for baseline age, baseline age^2, income, education, occupation, smoking status, alcohol, sedentary time, physical activity, and heart rate. BMI: Body mass index; WC: waist circumference; WHR Waist-to-hip ratio; SD: Standard deviation.

Figure 6.9. Associations of fifths of usual BMI with HF and PHD after excluding the first two years of follow-up and prior medical conditions.



All analyses were stratified by age-at-risk, sex, and study area and adjusted for baseline age, baseline age², income, education, occupation, smoking status, alcohol, sedentary time, physical activity, and heart rate. Analyses excluding ever-regular smoker did not include additional analysis for smoking status in the regression analyses. The quintile-specific mean usual body mass index (BMI) were recalculated for each analysis. CRD = Chronic respiratory disease; SOB = Shortness of breath.

Figure 6.10. Associations of fifths of usual WHR with HF and PHD after excluding the first two years of follow-up and prior medical conditions.



Conventions as in Figure 6.9. CRD = Chronic respiratory disease; SOB = Shortness of breath

6.4. Discussion

6.4.1. Association of BMI and WHR with HF

This chapter investigated the shape and strength of the associations of BMI and WHR with HF and PHD overall and by subgroup. In addition, it evaluated the mediating effects of SBP and diabetes on the overall associations. The present analysis showed J-shaped associations of BMI with HF: each 1 SD higher level of BMI (2.6 kg/m^2) was associated with a 22% higher risk of HF (HR 1.22, 95% CI 1.18–1.25) at BMI levels >22 kg/m^2. Adjusting for an increasing number of confounders did not materially alter the strength of the associations of BMI and WHR with HF and PHD, indicating that the observed associations are less likely to be explained by residual confounding.

The strength of the association of BMI with HF was similar to studies conducted in Western countries. Djousse et al., 2012 reported a 22% higher risk of HF per 1 SD higher level of BMI in the Cardiovascular Health Study[107]. In addition, Van Lieshout et al., 2011 demonstrated similar estimates in Dutch men (1.21, 1.09–1.34) and women (1.19, 1.07–1.31) for every 1 SD higher level of BMI[98]. However, previous studies have demonstrated a higher risk of HF per 1 SD higher level of BMI, with HRs ranging from 1.34–1.49[75,100,101,104,110]. By contrast, other studies have shown a weaker association of BMI with HF than reported in the present analysis.

A recent meta-analysis that investigated the association of BMI with HF demonstrated substantial heterogeneity between studies included in the meta-analysis[75]. In that meta-analysis, the association of BMI with HF was weaker at older ages and in men, studies that did not exclude participants with prior CVD at baseline and studies with shorter follow-up periods[75]. The present analysis showed weaker associations of BMI with HF at older ages, although the association was similar in men and women. The discrepant results are also likely due to model misspecification whereby studies assume a linear association of BMI with HF when the underlying shape of the association is non-linear or

studies adjust for a mediator such as BP in the main analysis, leading to more conservative estimates. About 38% and 49% of the association of BMI and WHR with HF were mediated through SBP and diabetes, respectively. Hence, studies adjusting for such mediators are likely to observe weaker associations compared to studies that did not adjust for these mediators.

The present analysis found a J-shaped association between BMI and HF; there was a positive log-linear association between usual BMI and HF down to a BMI threshold of 22 kg/m^2. Previous studies conducted in Western countries also found a non-linear association of BMI with HF with a threshold at a BMI level of 25 kg/m^2 below which there is no association with HF[75,104,174]. In contrast, the two published prospective cohort studies conducted in the Asia-Pacific Region reported an inverse non-linear association of BMI with HF[95,96]. The non-linear association between BMI and HF is most likely due to reverse causality, as those in the lowest BMI groups are more likely to have underlying medical conditions that predispose them to HF. This reverse causality bias is stronger in studies that include fatal endpoints, such as the two studies from the Asia-Pacific Region that included only fatal HF cases[95,96]. The present analysis suggests that the association of BMI with HF might be linear or at least similar to those reported in Western populations and that the reported non-linear associations could be due to reverse causality; the inverse association in the lowest quintile of BMI disappeared with exclusion of the early years of follow-up and prior medical conditions to mitigate reverse causality bias.

Similar to previous studies investigating the association of WHR and HF, this analysis found a positive log-linear association of WHR with HF down to a threshold at ~0.85, below which there was no observed association between WHR[75,97,99–101]. The shape of the association of WHR with HF was consistent in sensitivity analyses excluding the first

few years of follow-up and prior medical conditions, suggesting that the results for WHR were less likely to have been affected by reverse causality.

The strength of the association of WHR with HF was consistent with the findings of studies conducted in Western populations. Each 1 SD (0.05) higher usual WHR was associated with a 17% higher risk of HF at WHR levels >0.85 (1.17, 1.13–1.21), which was consistent with the findings of a recent meta-analysis of studies conducted in Western countries (HR per 0.05-unit increase in WHR: 1.153; 95% CI, 1.148–1.159)[75]. By contrast, the ARIC study reported stronger associations of WHR with HF (HR per 0.05 increase: 1.40, 95% CI 1.33–1.48)[100], whereas Levitan et al., 2009 reported a weaker association among Swedish men (1.06, 1.02–1.11)[101]. These discrepancies could be due to biased risk estimates from model misspecification resulting from not accounting for non-linear relationships between WHR and HF. Levitan et al., 2009 did not investigate the shape of the association of WHR with HF and, hence, assumed a linear relationship[101]. The ARIC study assessed the risk of HF across tertiles of WHR[100]; however, using only three categories is insufficient to reliably investigate the shape of associations. The stronger association reported in the ARIC study could be due to insufficient adjustment for major confounders. Although the study adjusted for most major confounders, it did not account for income and physical activity, which are relevant in Western populations[100].

The present analysis found similar risks of HF for general adiposity (BMI) and central adiposity (WHR), consistent with the findings of previous studies[97,100,101]. However, previous studies have mostly compared BMI and waist circumference as measures of general and central adiposity, with the findings suggesting that compared to waist circumference, BMI is more strongly associated with HF. Although BMI resulted in a greater reduction in the logHR than WHR when both traits were mutually adjusted for

each other, the associations of BMI and WHR with HF in the present analysis were similar.

6.4.2. Associations of BMI and WHR with PHD

This is the first study to report the association of BMI and WHR with PHD. The results showed a reverse J-shaped association of usual BMI and WHR levels with PHD. At BMI levels <26 kg/m^2, each 1 SD higher levels of BMI were associated with a 39% lower risk of PHD. By contrast, WHR had a weaker association with PHD than BMI; each 1 SD higher usual levels of WHR were associated with a 23% higher risk of PHD. The associations of BMI and WHR with PHD were unchanged after adjusting for a wide range of confounders and the inverse associations of BMI and WHR with PHD were unaltered.

Further adjustments for mediators (SBP and diabetes) did not alter the strength of the associations of WHR and BMI with PHD, indicating that, if the present associations are causal, BMI and WHR might affect PHD risk via mechanisms not involving SBP and diabetes. The findings suggest that BMI, but not WHR, was independently associated with PHD after accounting for other measures of adiposity as the association of WHR with PHD attenuated to the null after adjusting for BMI. In contrast, the association of BMI with PHD remained significant after adjusting for WHR and waist circumference. Efforts to mitigate reverse causality on the associations of BMI and WHR with PHD substantially attenuated the shape of the associations to the null, suggesting that the observed associations were likely to be due to reverse causality.

6.4.3. Strengths and limitations

The present analysis had several strengths, such as a large number of HF and PHD cases that permitted reliable investigation of the shape of associations, including participants with underweights. The details of data collected in the CKB allowed for

rigorous adjustment for major confounders. The availability of resurvey data also permitted correction for measurement errors in the main exposures.

The findings are limited by the inability of anthropometric indicators of central adiposity including WHR to discriminate between different types of visceral adiposity that might be associated with the risk of HF. In addition, the findings are limited by unmeasured confounding. However, adjustments for major confounders reduced the associations of BMI and WHR with HF and PHD by less than 25%. Hence, further adjustment of unmeasured confounders is unlikely to explain the observed associations. However, the associations of BMI and WHR with PHD observed in the present analysis could be due to reverse causality bias. Exclusion of the first two years of follow-up and prior medical conditions attenuated the strength of these associations of BMI and WHR with PHD to the null, suggesting that these associations may not be causal.

6.5. Summary

This Chapter showed positive log-linear associations of BMI and WHR with HF with evidence of a threshold below which there was an inverse association (for BMI) or a null association (for WHR). Sensitivity analyses to mitigate reverse causality suggested a positive log-linear between BMI and HF. However, the shape of the association of WHR with HF was unchanged after such sensitivity analyses. By contrast, BMI and WHR were inversely associated with PHD, with evidence of a threshold above which there was a positive association (for BMI) or a null association (for WHR). The associations of BMI and WHR with PHD were attenuated to the null after sensitivity analyses to minimise reverse causality.

The prospective associations presented in Chapters 5 and 6 are chiefly limited by residual confounding and reverse causality. Chapters 7 and 8 will investigate the extent

to which these prospective associations are causal using approaches that minimise bias from confounding and reverse causality.

7. Genetically-predicted systolic blood pressure and risk of heart failure

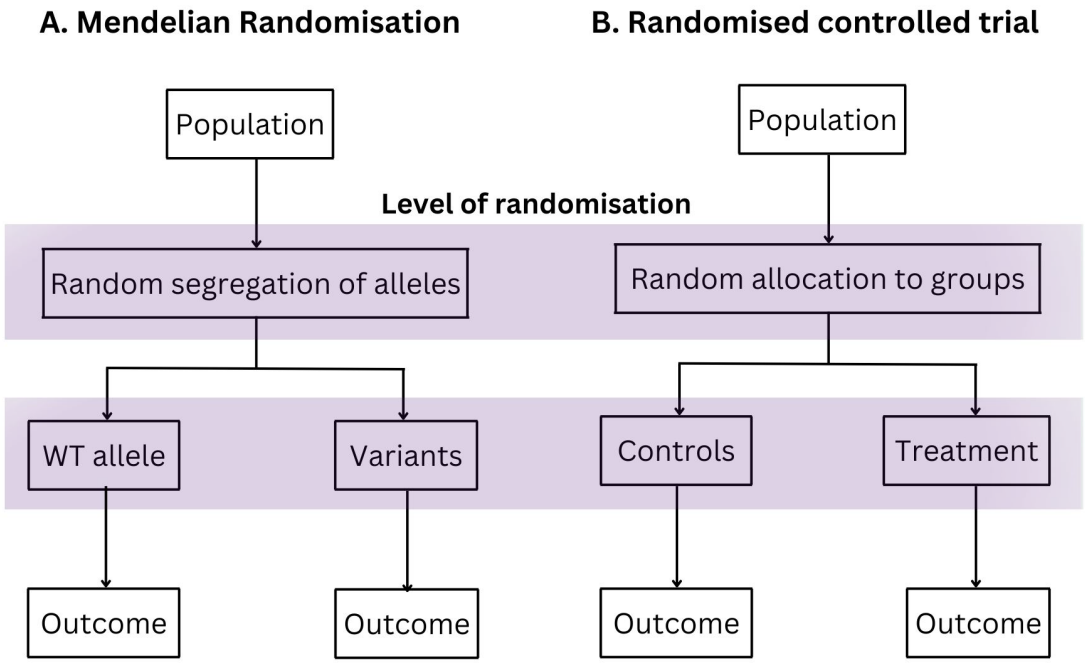
7.1. Background

The findings from Chapter 5 showed that higher levels of usual SBP were log-linearly and positively associated with higher risks of HF. These findings are consistent with a recent systematic review and meta-analysis, which reported higher relative risks of HF with higher levels of SBP[94]. However, findings of observational studies can be biased by confounding and reverse causality. These limitations can be mitigated using Mendelian randomisation (MR) to provide reliable evidence of causal associations. The findings from prospective observational studies of SBP and HF, which are also supported by MR analyses, are likely to be causal.

MR relies on an instrumental variable (IV) analysis. IV analysis, proposed about a century ago, uses a source of variation in the exposure that is uncorrelated with the outcome to make causal inferences in the presence of residual confounding[175,176]. In MR, genetic variants (usually single nucleotide polymorphisms [SNPs]) are used as IVs to investigate whether the relationship between an exposure and an outcome is causal[177,178]. MR relies on Gregor Mendel's first and second laws of genetic inheritance[179–182]. The first law, the law of segregation, implies that at each point in an autosomal genome, one copy of an allele from parents is transmitted to the offspring in a random manner[181,182]. The second law, the law of independent assortment, states that these alleles are passed down to offspring independently of one another[181,182]. The concept of MR is important for causal inference because bias from unmeasured confounding is less likely owing to the random allocation of genetic variants from parents to offspring; MR can be viewed as a natural experiment, analogous to an RCT, whereby participants have been randomly assigned to an intervention at birth (**Figure 7.1**)[177,183].

In addition, MR minimises the effects of reverse causality because the occurrence of disease cannot modify the genetic variation[177,181,183].

Figure 7.1: Similarities between Mendelian randomisation and randomised controlled trials[177].



In randomised controlled trials, individuals in the target population are randomly allocated to a control or intervention group, permitting the assessment of the intervention effect in the absence of confounding. In Mendelian randomisation, individuals in a population are randomly allocated to receive the more common allele (wild-type [WT] allele) or its variant from their parents. This randomisation is asymptotic at the population level, permitting the investigation of causal effects in the absence of confounding.

To be a valid instrument to investigate the causal association between an exposure (X) and an outcome (Y) using MR, the genetic variant (Z) must satisfy the three core IV assumptions (**Figure 7.2**)[176,177,184]. These assumptions are required to provide a reliable test of the null hypothesis of whether there is a causal effect[176,177,184].

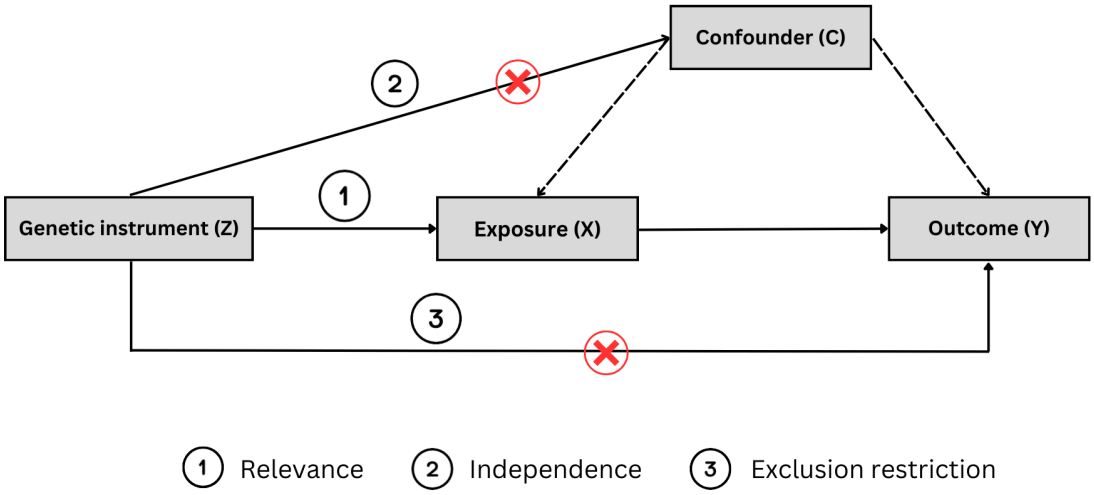
Relevance: The genetic variant should be associated with the exposure[177,184,185]. In the context of MR, this includes genetic variants robustly associated with the exposure at genome-wide significance (p-value $<5 \times 10^{-8}$) and have been replicated in independent

samples from the study population[185,186]. The genetic variant is, however, not required to be the causal variant: as long as it is correlated with a causal variant, which is a valid IV, then the genetic variant is also a valid instrument[184]. Ensuring the genetic IV is robustly associated with the exposure is important to prevent weak instrument bias. Weak instrument bias occurs when the genetic IV is weakly associated with the exposure, and in a one-sample MR setting, the causal estimate is biased towards the (confounded) observational association[178,187,188].

Independence: The genetic variant is not associated with any confounders[177,184,185]. The use of IVs is to mitigate the confounding. However, even the best-fitting model will be associated with some of the confounders of the exposure and outcome by the play of chance, which can lead to spurious association[177,184].

Exclusion restriction: The variant is not associated with the outcome except through the exposure[177,184,185]. A violation of this assumption, which is the most likely in practice, can lead to horizontal pleiotropy, biasing the MR estimate[177,184,185]. Horizontal pleiotropy occurs when the genetic IV influences the outcome via mechanisms other than its effect on the exposure[177,184,185].

Figure 7.2: Core assumptions of Mendelian randomisation.



MR can be defined as one-sample or two-sample. In one-sample MR, data on the genetic variants, exposure and outcome are measured in the same study population[178,185]. In contrast, two-sample MR requires data from two samples: one contains data on genetic variants and exposure, whereas the other contains data on genetic variants and outcome[178,185].

Few studies, mostly including individuals of European descent, have used MR to investigate the causal relationship between SBP and HF[189–193]. All studies demonstrated higher risks of HF with higher levels of genetically-predicted SBP. Most of the studies used a two-sample MR approach. Lian et al., 2023 conducted a two-sample MR analysis in individuals of European ancestry[190] using gene-SBP associations from a meta-analysis of about one million individuals from the UK Biobank and the International Consortium of Blood Pressure (ICBP)[194] and gene-HF associations from about one million individuals (including ~47000 HF cases) in the HEart failure Molecular Epidemiology for Therapeutic targetS (HERMES) consortium: an international collaboration comprising 57 population-based case-control studies, cohorts and RCTs

designed to investigate the genetic causes of HF and inform novel therapeutic approaches[195]. The study demonstrated that each 10 mmHg higher genetically-predicted SBP was associated with a 24% higher odds of HF (OR 1.24, 95% CI 1.19–1.29)[190]. Ciofani et al., 2023 used the same datasets as Lian et al. to investigate the causal associations between SBP and HF[193]; it is, therefore, not surprising that they found similar results to Lian et al. (OR per 10 mmHg higher genetically-predicted SBP=1.17, 95% CI 1.07–1.28). Wan et al., used data on the gene-SBP association from the ICBP and gene-HF data from the UK Biobank. They reported a 28% higher odds of HF for each 10 mmHg higher genetically-predicted SBP (OR 1.28, 95% CI 1.18–1.42)[189]. Consistent with studies conducted in individuals of European ancestry, Ciofani et al., 2023 also reported a higher but stronger risk of HF per 10 mmHg increase in genetically-predicted SBP among East Asians (OR 1.63, 95% CI 1.48–1.79)[193].

Previous MR studies suggested that SBP was a risk factor for HF, but majority of these studies were conducted in individuals of European ancestry. This chapter will (1) investigate the causal relationship between genetically-predicted SBP and HF and (2) evaluate the extent to which this association is due to horizontal pleiotropy.

7.2. Methods

7.2.1. Data source of SNPs for SBP

SNPs for the MR analyses were identified from a meta-analysis of data of about one million individuals of European ancestry in the UK Biobank and ICBP[194]. The UK Biobank involved about 500,000 participants and the ICBP included data from 77 independent studies of 299,024 participants. Studies included in the meta-analysis measured BP using standard operating procedures to ensure consistency across cohorts. Validated and calibrated automated BP devices or manual sphygmomanometers were used to take ~2–3 BP readings, and the average of these readings was used for analysis[194]. Genetic associations were replicated using data from

220,520 individuals in the Million Veteran Program in the US and 28,742 people in the Estonian Genome Centre, University of Tartu[194]. The study demonstrated good concordance of the SNP effects on SBP across samples of European, African and South Asian ancestry, highlighting the applicability of the findings across ancestry[194].

SNPs were selected from the UK Biobank and ICBP meta-analysis rather than an East Asian-specific GWAS because the larger sample size provides a greater statistical power to identify weak associations, enabling broader coverage of the common genetic variations across the genome[196]. Moreover, this approach serves to externally validate these SNPs in an East Asian population. SNPs were not selected based on the discovery GWAS conducted in the CKB as this would lead to "winner's curse" (whereby genetic associations are overestimated when first discovered) and exacerbate weak instrument bias[178,187,188]. Hence, selecting genetic variants robustly associated with the exposure in an external GWAS can minimise the risk of weak instrument bias and improve the external validity of the MR results[178].

7.2.2. Statistical analysis

7.2.2.1. Selection of SNPs for genetic instrument

Lead SNPs within each genomic locus, excluding the human leukocyte antigens (HLA) region on chromosome 6, robustly associated with the UK Biobank and ICBP meta-analysis by Evangelou et al., were selected and retrieved in the CKB. Evangelou et al. defined a lead SNP as the SNP with the smallest p-value within each genomic locus after excluding SNPs in linkage disequilibrium (LD: genetic correlation between SNPs) using an r^2 threshold of ≥0.1[194].

Differences in LD patterns between populations can cause SNPs identified as independent in Europeans to be correlated in East Asians[197,198]. Correlation between SNPs in an MR analysis reduces the reliability of the causal estimate as it violates the

second and third MR assumptions, leading to horizontal pleiotropy[177,184,185]. In addition, correlations between SNPs inflate the standard errors of MR associations, increasing the risk of type I error[178]. Hence, to ensure the genetic variants included in this analysis were independent, LD clumping was performed using the *"ld_clump"* package[199]. The package uses data from the 1000 Genome Project (Phase 3) to calculate population-specific LD between SNPs[200]. The 1000 Genome Project Phase 3 provides a detailed characterisation of over 88 million common genetic variations in individuals from diverse ancestry (Africa, the Americas, Europe, East Asia and South Asia) using whole-genome sequencing[200]. Genetic variants correlated ($r^2 \geq 0.001$) with the lead SNP and within a 10 MB genomic window of the lead SNP in East Asian populations were excluded.

Of the 521 SNPs identified in the UK Biobank and ICBP meta-analysis, seven were monomorphic in East Asians and were excluded. In addition, 174 SNPs in LD were excluded. Moreover, 44 SNPs with a minor allele frequency <0.005 in the CKB were excluded to minimise the risk of weak instrument bias[201].

7.2.2.2. Correcting the effect of BP-lowering medication on SBP

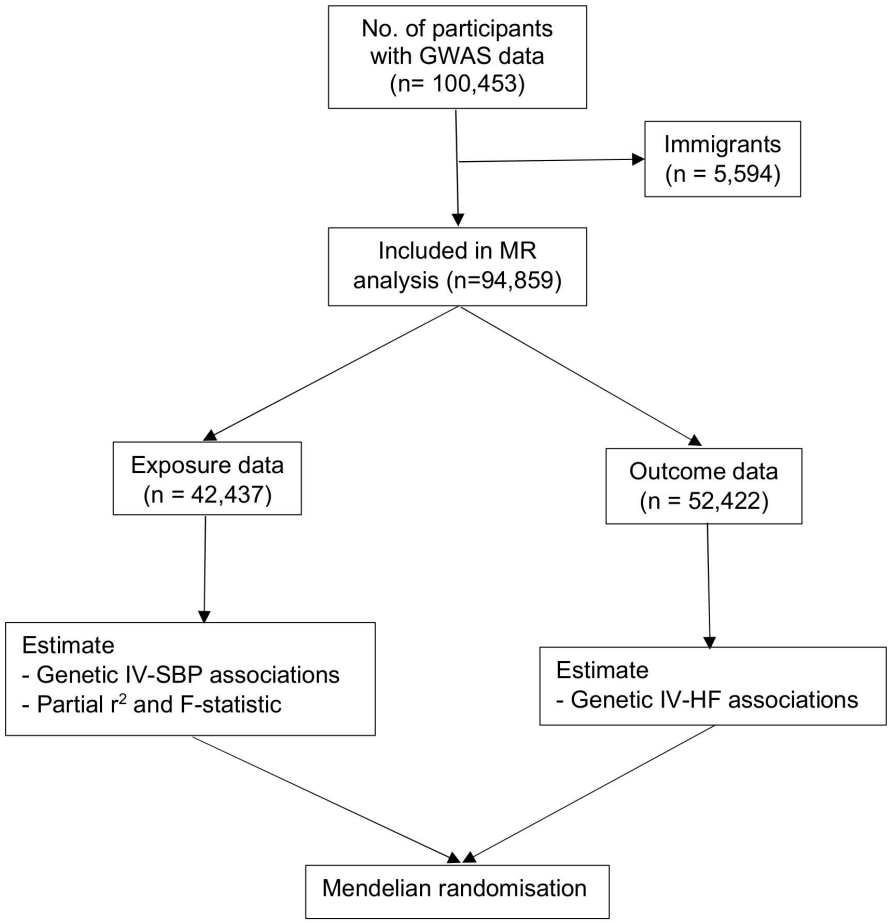
Baseline SBP values were corrected for BP-lowering medication use by adding 15.0 mmHg to the observed baseline SBP values. This correction factor was obtained from the mean reduction in SBP due to BP-lowering medications reported from randomised controlled trials and had been used in previous studies to correct SBP values for medication use[152,153,194,202,203].

7.2.2.3. Instrument strength

Before MR analysis, the CKB data used for the present analysis was randomly divided into two independent subsets based on the 10 CKB study areas. The exposure data was used to estimate the associations between genetic IVs and SBP and the F-statistic and partial r^2 of genetic IVs (**Figure 7.3**). The outcome data was used to estimate the

associations of genetic IVs with HF (**Figure 7.3**). Each subset included five of the 10 CKB study areas to account for the substantial genetic diversity across these areas. The subset of data used to estimate the associations between the genetic IV and SBP included the areas Hunan, Liuzhou, Qingdao, Haikou and Sichuan. In contrast, Harbin, Suzhou, Gansu, Henan and Zhejiang were included in the subset used to estimate the associations between the genetic IVs and HF (Total HF events = 1160).

Figure 7.3. Flowchart showing the partitioning of datasets for Mendelian randomisation analysis.



GWAS: Genome-wide association study, IV: Instrumental variable, SBP: Systolic blood pressure, HF: Heart failure.

The F-statistic quantifies the strength of the genetic instrument, and a value of ≥ 10 indicates a strong instrument and a lower likelihood of weak instrument bias[177,178,204].

The F-statistic and conditional variance in SBP explained (partial r^2) by the genetic IVs were estimated to assess the relevance assumption. Linear regression was used to regress measured SBP (dependent variable) on each SNP (independent variable) separately for each CKB study area (five study areas) and for men and women (10 strata), with adjustments for age, age^2, genotype array and area-specific principal components. The area-sex-specific estimates were then combined using inverse-variance-weighted fixed-effect meta-analysis. The SNP-specific F-statistic and partial r^2 were calculated as the median of area-sex-specific F-statistic and partial r^2, respectively, and the overall F-statistic and partial r^2 were obtained by summing the SNP-specific F-statistic and r^2.

Owing to the genetic diversity across CKB areas[125], performing area-specific analyses with careful adjustment for area-specific principal components is essential to prevent confounding by population stratification[205,206]. Although reverse causality bias is unlikely in MR studies because the occurrence of diseases such as HF cannot alter germline genetic variations, causes of HF like rheumatic heart disease and coronary heart disease that influence the SBP levels can cause reverse causation[207]. Hence, analyses of the genetic IVs with SBP were restricted to individuals without rheumatic heart disease, coronary heart disease, stroke or TIA to mitigate reverse causality.

7.2.2.4. Mendelian randomisation analysis

The causal estimate was calculated using the IVW multiplicative random-effects (mre), which improves statistical power by accounting for both within (precision of each SNP-specific estimate)- and between (heterogeneity across estimates)-variant variability[178,208,209]. To provide valid causal estimates, the IVW mre method relies on the balanced pleiotropy assumption: the pleiotropic effects are symmetrically distributed around zero[178,208,209]. This method can also quantify the degree of heterogeneity in the genetic IV-exposure association, which could indicate pleiotropy or other violations of

MR assumptions. The Wald ratio method was used to estimate the MR estimate for each SNP[178,210], and the SNP-specific estimates were then pooled into an IVW mre to calculate the overall causal estimate.

To calculate the Wald ratio, measured SBP was first regressed on each genetic IV using linear regression to estimate the mean change in SBP per additional unit of genetic IV. Separate regressions were fitted for each of the five CKB areas and for men and women with adjustments for age, age^2, BMI, BMI^2, genotype array, case ascertainment status and area-specific principal components. All analyses were restricted to individuals without rheumatic heart disease, coronary heart disease, stroke or TIA at baseline. The area-sex-specific betas were then combined using inverse-variance-weighted fixed-effect meta-analysis to obtain an overall beta ($\hat{\beta}_x$).

Second, Cox proportional hazard regression yielded log HR for the risk of HF per additional unit of the genetic IV. Regressions between each genetic IV and HF were fitted separately for each study area (five areas) and for men and women with adjustment for age, age^2, genotype array, case ascertainment status, BMI, BMI^2, and area-specific principal components and stratification by age-at-risk (10-year age-at-risk bands). The area-sex-specific log HRs were then combined using inverse-variance-weighted fixed-effect meta-analysis to obtain an overall log HR ($\hat{\beta}_y$). Analyses of the genetic IV-outcome associations were not restricted to individuals with rheumatic heart disease, coronary heart disease, stroke or TIA because there is no risk of reverse causality.

The ratio estimate was then calculated as $\hat{\beta}_y / \hat{\beta}_x$ and is interpreted as the log HR of HF per 1 mmHg increase in genetically-predicted SBP. The standard error of the ratio estimate was calculated from the second-order term of the delta method that accounts for uncertainties in the denominator of the ratio estimate as follows[211]:

$$\sqrt{\frac{se(\hat{\beta}_y)^2}{(\hat{\beta}x)^2} + \frac{(\hat{\beta}_y)^2 \times se(\hat{\beta}x)^2}{(\hat{\beta}x)^4}}$$

The SNP-specific ratio estimates were then pooled using IVW mre meta-analysis to obtain the overall causal estimate. The overall causal estimate (and standard error) was multiplied by 10 before exponentiating to obtain the HR (and standard error) for HF per 10 mmHg increase in genetically-predicted SBP.

7.2.2.5. Sensitivity Analyses

Heterogeneity in genetic IV-specific effects, which can result from varying biological pathways, weak instruments, non-linearity and effect modification, suggests a violation of the core MR assumptions and can lead to unreliable and biased causal inference[178,212].

Therefore, robust methods addressing various sources of bias such as horizontal pleiotropy are recommended to assess the validity and reliability of the findings from primary MR analyses. Various robust methods have been proposed for MR, but it will be excessive to evaluate all of these methods as sensitivity analyses. Hence, the weighted median, Mendelian randomisation robust adjusted profile score (MR-RAPS) and Mendelian Randomisation using constrained Maximum Likelihood (MR-cML) methods were used as sensitivity analyses to assess the robustness of the MR estimates under different assumptions, including the majority valid, Instrument Strength Independent of Direct Effect (InSIDE) and plurality valid assumptions, respectively. The underlying assumptions, strengths and limitations of these methods are summarised in **Table 7.1**.

MR-RAPS assumes the InSIDE assumption—the effect of the genetic IV on the exposure is independent of its direct effect on the outcome—after excluding strongly pleiotropic variants[178,213,214]. The median-based method assumes that at least 50% of the genetic IVs used for causal inference are valid[208]. The median estimates the causal

effect by calculating some measure of central tendency and is, therefore, robust to the presence of outliers[178,215,216]. Finally, the MR-cML method makes the plurality assumption, which states that most of the genetic IVs estimate the true causal effect more than any other quantity[178,217]. The MR-cML relies on constraint maximum likelihood and model averaging to select and remove invalid variants and make estimations[178,217]. The chief strengths of this method include the ability to make valid causal inferences about the presence of both uncorrelated pleiotropy and correlated pleiotropy[178,217].

Table 7.1: Robust methods used to assess the reliability of the Mendelian randomisation causal estimates.

Method	Assumption	Strengths and limitations
MR-RAPS	InSIDE (in the absence of outliers)	Downweights outliers but sensitive to violations of the assumption of balanced pleiotropy.
Weighted median	Majority valid	Outlier robust but sensitivity to the removal or addition of genetic variants.
MR-cML	Plurality valid	Likelihood-based estimation approach and robust to violations of all the three instrumental variable assumptions.

InSIDE: Instrument Strength Independent of Direct Effect
IVW: Inverse-variance weighted
MR-cML: Mendelian Randomisation using constrained Maximum Likelihood
MR-RAPS: Mendelian randomisation robust adjusted profile score

7.3. Results

7.3.1. Baseline characteristics of participants included in the genetic analysis

Of the 100,453 participants included in the GWAS in the CKB, 5,594 immigrants were excluded. **Table 7.2** summarises the baseline characteristics of the non-immigrants with GWAS data. The mean (SD) age, SBP, BMI and WHR of the participants were 53.7 (11.0) years, 133.3 (21.7) mmHg, 23.6 (3.5) kg/m^2 and 0.9 (0.1), respectively (**Table 7.2**). About 40% of the participants were ever-regular smokers, 16% were current drinkers and 13% were taking BP-lowering drugs (**Table 7.2**). About 86% and 6.3% of men and women were ever-regular smokers, respectively. In addition, 33.9% and 2.3% of men and women were current drinkers. The baseline characteristics of participants included in the genetic analysis were representative of the CKB population (**Table 7.2**).

Table 7.2: Characteristics of CKB participants with genetic data compared to the entire CKB population.

Characteristics	Genetic analysis			Baseline (ALL)		
	Men	Women	All	Men	Women	All
No. of participants	40,614	54,245	94,859	210,205	302,519	512,724
Demographic and SES factors, %						
Age, mean (SD), years	54.7 (11.2)	52.9 (10.7)	53.7 (11.0)	52.9 (10.9)	51.4 (10.4)	52.0 (10.7)
Living in rural area	60.3	56.9	58.6	56.7	55.1	55.9
Ambient outdoor temperature, mean (SD), °C	15.4 (9.6)	15.7 (9.5)	15.6 (10.5)	15.9 (9.2)	16.3 (9.0)	16.1 (10.1)
Married	92.3	86.5	89.2	93.3	88.4	90.6
Middle school education or higher	65.6	67.6	66.3	67.7	68.4	67.8
Household month income <10,000 Yuan	29.5	33.6	31.9	26.0	29.8	28.2
Agriculture or factor worker	61.5	47.6	54.1	62.3	50.7	55.8
Lifestyle factors, %						
Ever regular smoker	85.9	6.3	40.3	85.6	5.2	38.1
Current drinkers	33.9	2.3	15.5	33.4	2.1	14.9
Total physical activity, mean (SD), MET-hr/day	21.1 (13.3)	19.0 (10.2)	19.9 (13.8)	22.4 (13.4)	20.1 (10.4)	21.1 (13.9)
Sedentary time, mean (SD), hours/day	3.1 (1.4)	2.9 (1.4)	3.0 (1.5)	3.1 (1.4)	2.9 (1.4)	3.0 (1.5)
Physical measurements, mean (SD)						
BMI, kg/m²	23.4 (3.1)	23.8 (3.4)	23.6 (3.5)	23.5 (3.1)	23.8 (3.3)	23.7 (3.4)
SBP, mean (SD), mmHg	135.0 (20.3)	133.3 (21.7)	133.9 (22.7)	132.3 (18.8)	130.5 (20.2)	131.1 (21.3)
DBP mean (SD), mmHg	80.2 (11.7)	77.7 (11.3)	78.7 (11.7)	79.2 (11.1)	76.8 (10.7)	77.8 (11.2)
Resting heart rate, bpm	78.0 (12.3)	79.9 (11.6)	79.0 (12.0)	77.8 (12.1)	79.7 (11.5)	78.9 (11.8)
Waist circumference, cm	81.9 (9.2)	79.1 (9.3)	80.2 (10.0)	82.1 (9.2)	79.1 (9.1)	80.3 (9.8)
Waist-to-hip ratio	0.9 (0.1)	0.9 (0.1)	0.9 (0.1)	0.9 (0.1)	0.9 (0.1)	0.9 (0.1)
Body fat percentage, %	21.9 (5.9)	32.2 (7.2)	27.7 (8.7)	22.1 (5.9)	32.1 (6.9)	27.9 (8.4)
Prevalent conditions						
Hypertension	41.5	39.7	40.3	36.3	34.7	35.2
Taking BP-lowering drugs	11.3	13.8	12.6	10.2	12.4	11.5
Diabetes	6.3	7.1	6.6	5.5	6.3	5.9
COPD	11.0	8.4	9.7	8.2	6.4	7.2
Emphysema or bronchitis	4.4	3.2	3.8	2.9	2.3	2.6
Asthma	0.7	0.7	0.7	0.6	0.5	0.5
Shortness of breath on walking	6.1	8.9	7.7	4.5	6.9	5.9
Self-reported poor health	10.1	13.0	11.7	8.8	11.5	10.4

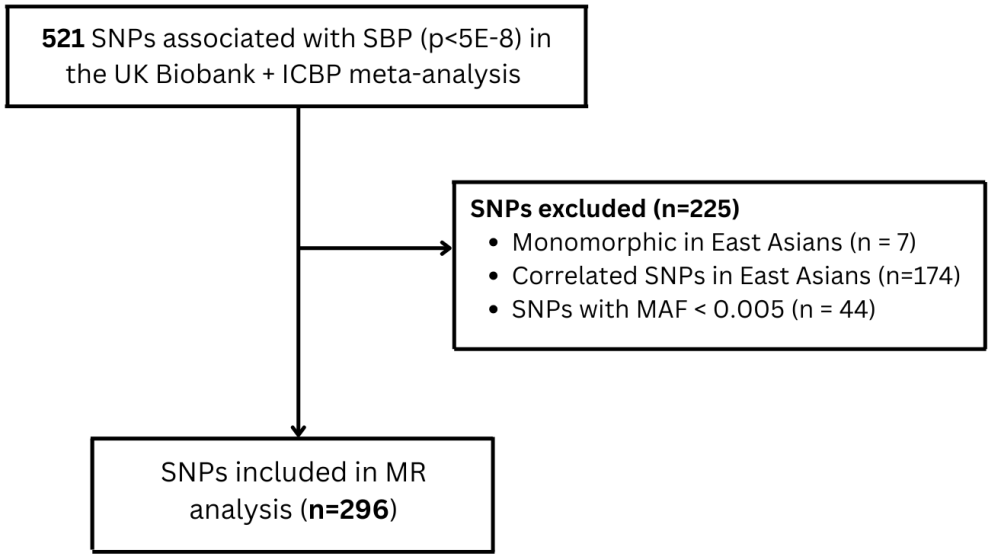
Means and proportions where directly standardised by the age, sex, and all 10 CKB areas.
bpm: beats per minute; COPD: Chronic obstructive pulmonary disease; DBP: Diastolic blood pressure; MET: Metabolic Equivalent Task; SD: Standard deviation; SES: Socioeconomic factors; SOB: Shortness of breath

7.3.2. Strength of genetic instruments for SBP

Among the 521 SNPs associated with SBP at genome-wide significance in the UK Biobank and ICBP meta-analysis, 225 SNPs were excluded, leaving 296 SNPs for the present analysis (**Figure 7.4**). The reasons for exclusion included SNPS that were monomorphic in East Asians (n=7), in LD (n=174) or had a minor allele frequency <0.005 (n=44). **Appendix 1** shows the genetic variants included in this analysis.

Each additional copy of the genetic IV was associated with a 0.32 mmHg increase in SBP (beta = 0.32, 95% CI 0.32–0.33). The genetic IVs explained 4.0% of the total variability in SBP, corresponding to an F-statistic of 625.2.

Figure 7.4: Flowchart for selection of single nucleotide polymorphisms included in the analysis.

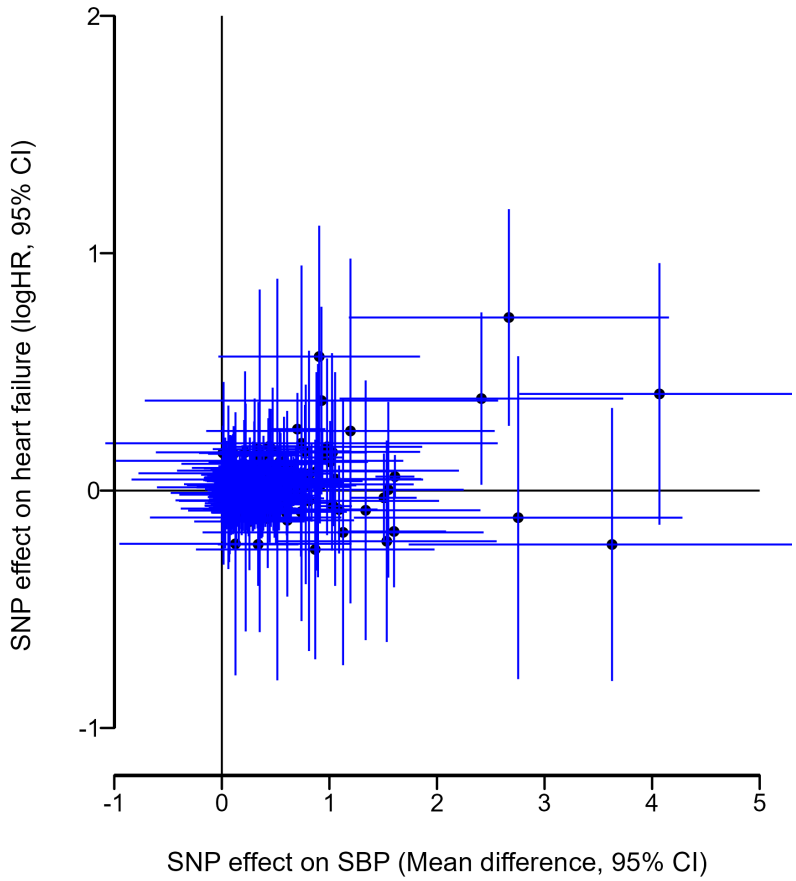


ICBP: International Consortium for Blood Pressure; GRS: Genetic risk score; MAF: Minor allele frequency; W*i*: Weight for the i^{th} single nucleotide polymorphism (SNP).

7.3.3. Association of genetically-predicted SBP and risk of HF

This analysis included 1160 HF cases. Genetically-predicted SBP was significantly positively associated with a higher risk of HF. Each 10 mmHg higher levels of genetically-predicted SBP were associated with an 18% higher risk of HF (HR 1.18, 95% CI 1.01–1.39, p=0.0356). **Figure 7.5** shows the scatterplot for the SNP associations with SBP and HF in the CKB. There was no evidence of heterogeneity of the SNP-specific effect (Q= 318.2, df= 295, $P_{heterogeneity}$= 0.1680).

Figure 7.5. Scatterplot of SNP-SBP and SNP-HF associations in the CKB.



7.3.4. Sensitivity analyses

MR-RAPS demonstrated consistent results with the IVW mre (**Table 7.3**), suggesting that any variants with extreme effects are less likely to bias the overall causal estimate. The weighted median and MR-cML methods demonstrated consistent strengths of effect albeit a loss of statistical power using these methods as reflected by the wider confidence intervals (**Table 7.3**).

Table 7.3. Associations genetically-predicted with risk of heart failure using robust methods.

Methods	HR (95% CI)	P-value
IVW mre	1.18 (1.01–1.39)	0.0356
Weighted median	1.19 (0.91–1.55)	0.2120
MR-cML	1.15 (0.93–1.43)	0.2040
MR-RAPS	1.24 (1.02–1.42)	0.0264

IVW mre: Inverse-variance-weighted multiplicative random-effects
MR-cML: Mendelian Randomisation using constrained Maximum Likelihood
MR-RAPS: Mendelian randomisation robust adjusted profile score

7.4. Discussion

This chapter investigated the associations of genetically-predicted SBP with risks of HF among Chinese adults. The results showed a positive association of genetically-predicted SBP with HF. Each 10 mmHg higher levels of genetically-predicted SBP were associated with an 18% (HR 1.18, 95% CI 1.01–1.39) higher risk of HF. The associations of genetically-predicted SBP with HF were consistent across various pleiotropy-robust MR methods.

7.4.1. Association of genetically-predicted SBP with risk of HF

The findings of the present analyses are consistent with previous studies conducted on individuals of European ancestry. Lian et al., 2023 and Ciofani et al., 2023 analysed data

from the ICBP and HERMES consortia and reported 24% (OR 1.24, 95% CI 1.19–1.29) and 17% (1.17, 1.07–1.28) higher risk of genetically-predicted HF for each 10 mmHg higher risk of genetically-predicted SBP, respectively[190,193]. Consistent findings were reported by Wan et al., 2021 among individuals of European ancestry (OR 1.28, 95% CI 1.18–1.42)[189]. The findings corroborated with a previous study by Clarke et al., using genetic risk scores among Chinese adults; the study showed a 32% higher risk of HF for each additional 10 mmHg higher risk of genetically-predicted SBP (HR 1.32, 95% CI 1.13–1.54)[192].

By contrast, Ciofani et al., 2023 reported stronger associations between genetically-predicted SBP and the risk of genetically-predicted HF[193]. Each 10 mmHg higher levels of genetically-predicted SBP were associated with a 63% higher risk of genetically-predicted HF (OR 1.63, 95% CI 1.48–1.79). This discrepancy is likely due to over-fitting from sample overlap as the study used data on genetic associations of the exposure and outcome derived from Biobank Japan, potentially resulting in an overestimation of the causal effect[193].

The positive association of genetically-predicted SBP levels with risks of HF is consistent with the results of the prospective associations presented in Chapter 5 and a recent meta-analysis of observational studies, which included mostly participants from Western countries[94].

7.4.2. Strength and limitations

The chief strength of the present analysis includes the availability of individual-level data for MR, which enabled careful control of confounding and reverse causality from prevalent medical conditions such as coronary heart disease on measured SBP. By mitigating reverse causality, MR enabled reliable estimation of the causal association of genetically-predicted SBP with risks of HF. Furthermore, the use of genetic variants

discovered from the largest GWAS consortium on BP enabled a broad coverage of genetic variants associated with SBP. Moreover, the conditional F-statistic of the genetic IVs was high, thereby minimising the possibility of weak instrument bias.

However, the present analysis had several limitations. The MR approach assumes that the effect of genetic IV on the outcome is entirely through the exposure and not through alternative pathways (horizontal pleiotropy). While this assumption is not testable, the results of pleiotropy-robust methods suggested that the MR estimates are unlikely to be biased by horizontal pleiotropy.

7.5. Summary

The findings of the present analysis were consistent with the observational associations presented in Chapter 5, which demonstrated a positive association between SBP and HF. Combined with genetic evidence from previous studies, the findings of this chapter highlight the importance of higher levels of SBP as a causal risk factor for HF in Chinese adults. The next chapter will also use MR analyses to further investigate the causal relevance of adiposity traits for HF and PHD.

8. Genetically-predicted adiposity traits and risk of heart failure and pulmonary heart disease

8.1. Background

Chapter 6 demonstrated directionally opposing associations of BMI and WHR with HF and PHD. BMI and WHR were positively and log-linearly associated with higher risks of HF, with evidence of a threshold of BMI or WHR below which there was an inverse association (for BMI) or a null association (for WHR). The findings persisted after sensitivity analyses that excluded events occurring during the first few years of follow-up in addition to prior diseases. By contrast, BMI and WHR were inversely associated with PHD; however, these associations attenuated to the null after excluding prior diseases and events occurring during the first few years of follow-up, suggesting that the associations in the observational analysis were likely to be due to reverse causality. This chapter uses Mendelian randomisation (MR) analyses to investigate the causal relevance of the associations of BMI and WHR with HF and PHD, respectively.

Previous MR studies on the association of adiposity and risk of HF have been limited to individuals of European ancestry[77,218–221], with these studies indicating that genetically-predicted BMI and WHR were associated with higher risks of HF. Fall et al., 2013 reported one of the earliest MR studies investigating the causal association between BMI and HF[219]. The authors used genetic variants from the fat-mass- and obesity-associated gene (*FTO*) locus, which is specific for obesity and explains 0.34% of the variability in the BMI phenotype, as proxies for BMI[219]. Fall et al., 2013 reported a 19% higher risk of HF per 1 kg/m^2 higher levels of genetically-predicted BMI (HR 1.19, 95% CI 1.03–1.39)[219]. Hagg et al., 2015 used data from the European Network for Genetic and Genomic Epidemiology (ENGAGE) consortium and showed a 47% higher risk of HF per 1 SD (4.5 kg/m^2) increase in genetically-predicted BMI[218]. In 2021, Kim et al. and Larsson et al. published meta-analyses of MR studies of adiposity and risk of HF[77,220].

Kim et al., 2021 reported a 12% higher risk of HF per 1 kg/m^2 higher levels of genetically-predicted BMI (OR 1.12, 1.08–1.17)[77], whereas Larsson et al., 2021 reported a 62% higher risk of HF per 1 SD higher level of genetically-predicted BMI (OR 1.62, 1.50–1.75)[220].

A single previous study reported the association of genetically-predicted WHR with HF[221]. Hong et al., 2024 used genetic data for WHR from the Genetic Investigation of ANthropometric Traits (GIANT) consortium and for HF from the HERMES consortium and reported a 34% higher risk of HF per 1 SD higher levels of genetically-predicted WHR (OR 1.34, 1.11–1.62)[221].

There is limited available evidence on the associations of BMI and WHR with HF in non-European populations. Furthermore, no study evaluated the association of genetically-predicted BMI and WHR with PHD. This chapter aimed to investigate the associations of genetically-predicted BMI and WHR with HF and PHD in the CKB.

8.2. Methods

8.2.1. Data sources of SNPs for BMI and WHR

SNPs for adiposity traits were obtained from a trans-ethnic meta-analysis of the sex-specific GWAS summary statistics from the CKB and UK Biobank using rank-inverse-normal-transformed (RINT) residuals of BMI and WHR measured at baseline[222]. The UK Biobank sample included 443,359 participants of European ancestry. Participants were excluded if: they had withdrawn consent; their genotypic and self-reported sex did not match; did not have both imputation and genotype information; and had high heterozygosity[222].

The residual values were obtained by regressing the adiposity traits against age and age^2; for the UK Biobank, additional adjustments were made for the assessment centre and 40 principal components[222]. These analyses excluded participants with chronic

obstructive pulmonary disease (COPD) at baseline to mitigate reverse causality bias. The transformations were conducted separately by sex in the UK Biobank and by sex and study area in the CKB to account for differences in the distribution of these adiposity traits by sex in the UK Biobank and by sex and study area in the CKB[222]. The lead variants within each genomic locus were defined as variants with the smallest p-value in the UK Biobank, CKB or trans-ethnic meta-analysis[222].

8.2.2. Selection of SNPs for the analysis

The present analysis excluded SNPs that were monomorphic in East Asians, SNPs in LD (r^2 threshold ≥0.001) and SNPs with a minor allele frequency < 0.005 to minimise weak instrument bias[201]. Additionally, for WHR, SNPs associated with BMI were excluded.

8.2.3. Instrument strength

Before MR analysis, the CKB data under investigation was divided into two independent subsets as described in **Chapter 7**. The strength of the genetic IVs for each adiposity trait (BMI and WHR) was quantified using the F-statistic. The variance in the adiposity trait explained by the genetic IVs was assessed using the partial r^2. Linear regression was used to regress the adiposity trait on each SNP stratified by each of the five study areas and by men and women (10 strata). All analyses were adjusted for age, age^2, genotype array and area-specific principal components. In order to mitigate reverse causality, analyses of associations between genetic IVs and adiposity were restricted to participants without COPD, tuberculosis or cancer at baseline.

The SNP-specific F-statistic and partial r^2 were estimated as the median of the area-sex-specific estimates. The overall partial r^2 and F-statistic were calculated by summing the SNP-specific partial r^2 and F-statistic, respectively. The area-sex-specific regression coefficients (betas) were combined using inverse-variance-weighted fixed-effect meta-

analysis to estimate the overall association of the genetic IV with the measured adiposity trait ($\hat{\beta}_x$).

8.2.4. Mendelian randomisation

The causal estimate was calculated by pooling SNP-specific Wald ratios using an IVW multiplicative random-effects (mre) model as described in **Chapter 7**. To calculate the SNP-specific Wald ratios, the strength of the association of each SNP with a measured adiposity trait was calculated by regressing the adiposity trait on the SNP with adjustment for age, age^2, genotype array, case ascertainment status and area-specific principal components within each study area-sex stratum. The area-specific betas were pooled using inverse-variance-weighted fixed-effect meta-analysis to obtain the overall beta ($\hat{\beta}_x$), which represents the mean change in the adiposity trait per additional copy of the genetic IV. The analyses were conducted in participants without COPD, tuberculosis or cancer at baseline to minimise reverse causality.

Next, Cox proportional hazard regression yielded log HRs for the association between each genetic IV and incident HF or PHD after stratification for age-at-risk (10-year age bands) and sex and adjustments for age, age^2, genotype array, case ascertainment status and area-specific principal components. Analyses were conducted separately for each of the five study areas, and the area-specific log HRs were combined using inverse-variance-weighted fixed-effect meta-analysis to obtain the overall log HR ($\hat{\beta}_y$), which represents the risk of HF for each additional copy of the genetic IV.

The ratio estimate, calculated as $\hat{\beta}_y / \hat{\beta}_x$, represents the log HR of the outcome (HF or PHD) per unit increase in the genetically-predicted adiposity trait (BMI or WHR)[211]. To account for uncertainties in the denominator of the ratio estimate, the standard error of the ratio estimate was calculated from the second-order term of the delta method[211].

The SNP-specific ratio estimates were then pooled using IVW mre meta-analysis to obtain the overall causal estimate. The IVW estimates and their standard errors were multiplied by the SD of the adiposity traits to permit comparisons of the contributions of the two traits to HF or PHD risk. The scaled estimates were exponentiated to obtain the HRs of the outcomes per 1 SD higher level of the genetically-predicted trait.

8.2.5. Sensitivity analyses

The robustness of the overall MR findings to horizontal pleiotropy was assessed using MR robust adjusted profile score (MR-RAPS)[178,213,214], weighted median[178,215,216] and MR using constraint Maximum-likelihood (MR-cML)[178,217]. The MR estimate was compared before and after removing outlying SNPs to evaluate the impact of horizontal pleiotropy on the overall MR estimates[223]. Two-tailed p-values <0.05 were considered statistically significant.

8.3. Results

8.3.1. Characteristics of included participants

Table 7.2 in the previous chapter compared the characteristics of participants in the genetic analysis against those in the observational analysis. The mean (SD) BMI and WHR of the participants included in the genetic analysis were 23.6 (3.5) kg/m^2 and 0.9 (0.1). The mean BMI, WHR and waist circumference were similar for men and women. However, the mean (SD) body fat percentage was greater in women [32.2 (7.2) %] than in men [31.9 (5.9) %]. Men were more likely to have COPD at baseline (11.0%) than women (8.4%).

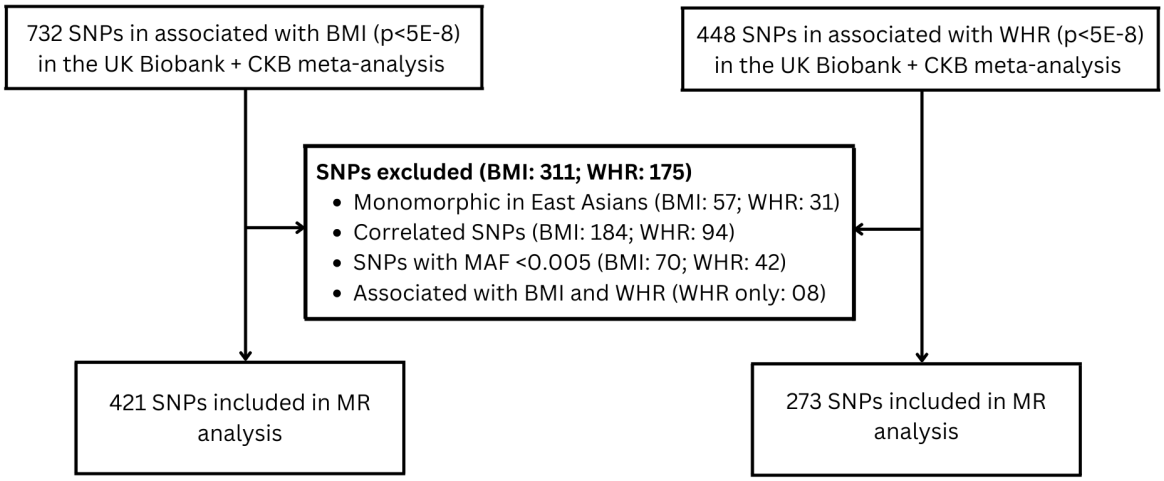
8.3.2. Strength of genetic instruments for BMI and WHR

This analysis considered 732 and 448 genetic variants associated with BMI and WHR at genome-wide significance ($p<5x10^{-8}$) in the CKB-UK Biobank trans-ethnic meta-analysis

for inclusion (**Figure 8.1**). Variants that were monomorphic in East Asians (BMI: 57; WHR: 31), correlated variants (BMI: 241; WHR: 121) and variants with a MAF<0.005 (BMI: 70; WHR: 42) in the CKB were excluded. In addition, for WHR, 08 variants associated with BMI were excluded. Finally, 421 and 273 variants were used as IVs for the MR analysis for BMI and WHR, respectively (**Figure 8.1**). **Appendices 2 and 3** show the genetic variants included in the analyses for BMI and WHR, respectively.

Each additional copy of the genetic IVs for BMI and WHR was associated with 0.06 kg/m^2 and 0.001 higher levels of BMI and WHR, respectively. The overall partial r^2 were 13.5% and 7.5%, corresponding to F-statistics of 540.4 and 307.1, respectively.

Figure 8.1. Flowchart for the selection of single nucleotide polymorphisms included in the analyses.



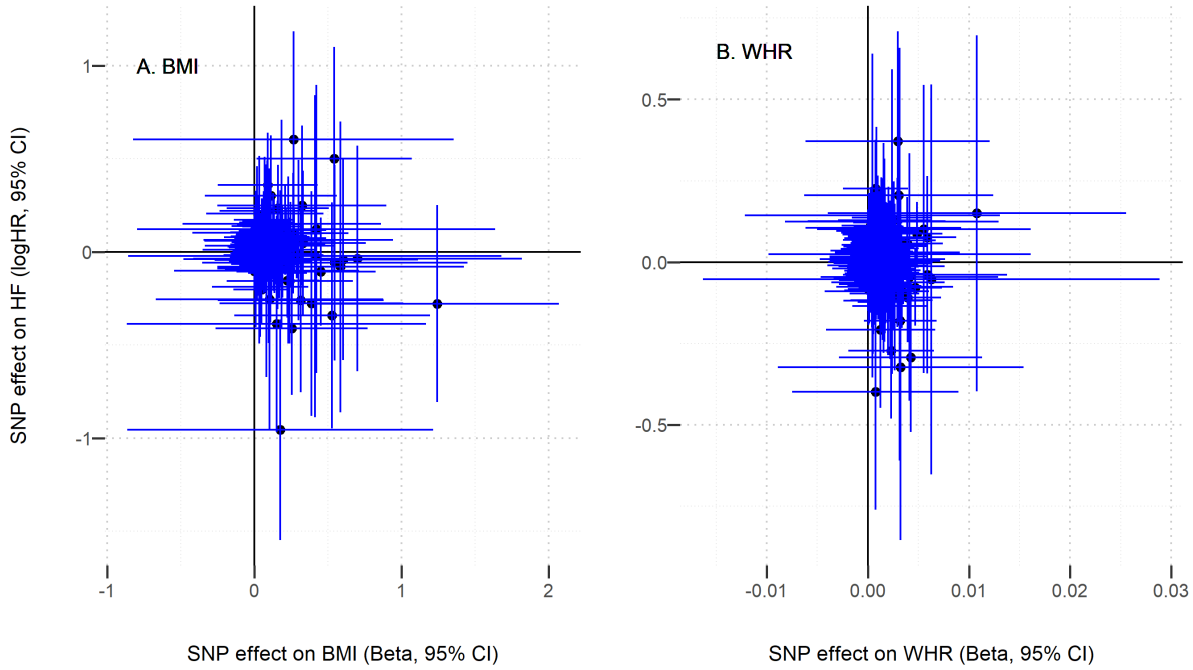
MAF: Minor allele frequency.

8.3.3. Associations of genetically-predicted BMI and WHR with HF and PHD

In total, 1160 and 1104 cases of HF and PHD were included in this MR analysis. Genetically-predicted BMI was positively associated with HF but not with PHD. There was a 39% (HR 1.39, 95% CI 1.11–1.75, p=0.0044) higher risk of HF per 1 SD increase

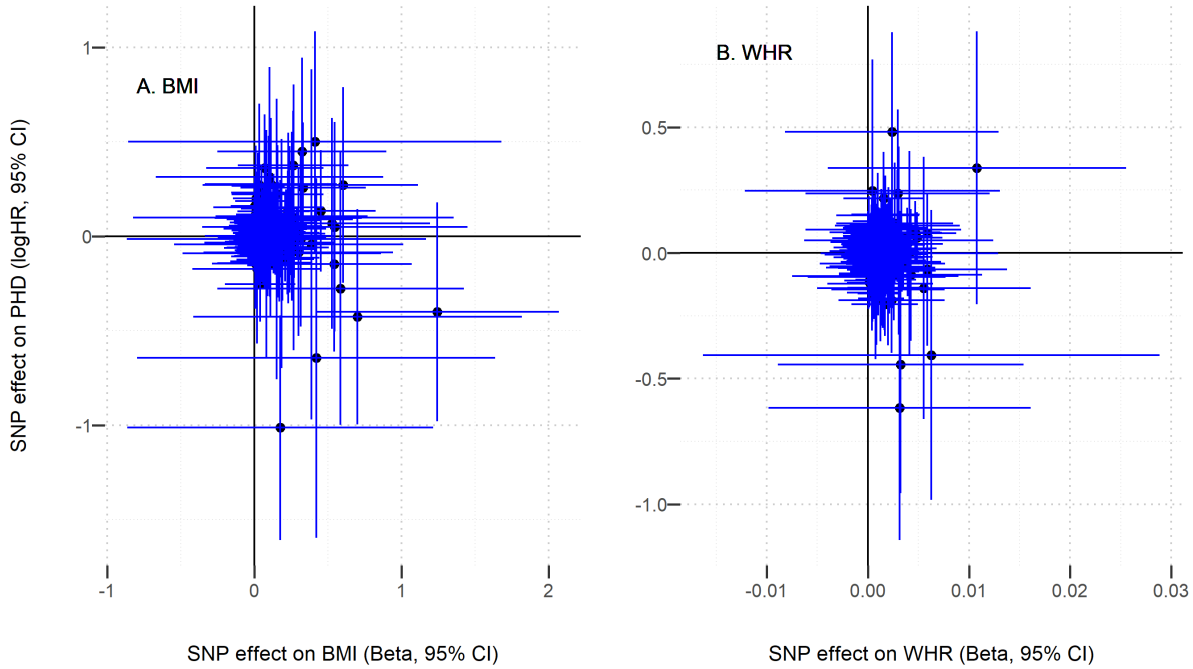
in genetically-predicted BMI with no evidence of heterogeneity between SNP-specific effects (Q=459.6, $P_{heterogeneity}$=0.0890). By contrast, there was no association of genetically-predicted BMI with PHD (HR 1.16, 95% CI, 0.94–1.43, p=0.1685) and no evidence of heterogeneity between SNP-specific effects (Q=264.8, $P_{heterogeneity}$=0.6119). Genetically-predicted WHR was not associated with HF (HR 1.08, 95% CI, 0.73–1.60, p=0.6897) and PHD (HR 0.91, 95% CI, 0.63–1.32, p=0.6128). There was no evidence of heterogeneity between SNP-specific effects for genetically-predicted WHR with HF (Q=460.7, $P_{heterogeneity}$=0.0886) and PHD (Q=285.2, $P_{heterogeneity}$=0.2785). **Figure 8.2** and **Figure 8.3** show the scatterplot of the SNP-adiposity associations with SNP-HF and SNP-PHD associations in the CKB.

Figure 8.2. Scatterplot of SNP-adiposity and SNP-HF associations in the CKB.



Beta: Mean difference; BMI: Body mass index; WHR: Waist-to-hip ratio

Figure 8.3. Scatterplot of SNP-adiposity and SNP-PHD associations in the CKB.



Beta: Mean difference; BMI: Body mass index; WHR: Waist-to-hip ratio

8.3.4. Sensitivity analyses

Table 8.1 presents the associations of genetically-predicted BMI and WHR with HF and PHD using pleiotropy-robust methods. For the associations of genetically-predicted BMI with HF, MR-RAPS, weighted median and MR-cML showed directionally consistent results with the IVW mre method. For the associations of genetically-predicted BMI and PHD, the results from different pleiotropy-robust methods were consistent with the main analysis. Likewise, the associations of genetically-predicted WHR with HF and PHD, when using robust methods, were consistent with the main IVW mre analysis (**Table 8.1**).

Table 8.1. Associations of genetically-predicted BMI and WHR with risks of HF and PHD using robust methods.

	Heart failure		Pulmonary heart disease	
Method	**HR* (95% CI)**	**P-value**	**HR (95% CI)**	**P-value**
Body mass index				
Weighted median	1.27 (0.89–1.82)	0.1857	1.04 (0.73–1.50)	0.1857
MR-cML	2.04 (0.71–5.87)	0.1851	1.15 (0.37–3.57)	0.8099
MR-RAPS	2.48 (1.34–4.60)	0.0039	1.36 (0.78–2.39)	0.2829
IVW mre	*1.39 (1.11–1.75)*	*0.0044*	*1.16 (0.94–1.43)*	*0.1685*
Waist-to-hip ratio				
Weighted median	0.87 (0.53–1.44)	0.5895	1.26 (0.77–2.07)	0.5895
MR-cML	0.81 (0.08–8.17)	0.8594	1.81 (0.22–14.96)	0.5832
MR-RAPS	0.62 (0.10–3.75)	0.6023	2.51 (0.51–12.28)	0.2571
IVW mre	*1.08 (0.73–1.60)*	*0.6897*	*0.91 (0.63–1.32)*	*0.6128*

CI: Confidence interval; HR: Hazard ratio; IVW mre: Inverse-variance weighted multiplicative random-effects; MR-cML: Mendelian Randomisation using constrained Maximum Likelihood; MR-RAPS: Mendelian randomisation robust adjusted profile score.

* HR per 1 SD higher adiposity trait.

8.4. Discussion

8.4.1. Summary of the main findings

This chapter aimed to evaluate the causal associations of BMI and WHR with HF and PHD. The findings showed a positive association of genetically-predicted BMI with HF

but found no strong evidence of any such association with PHD. In contrast, there was no strong evidence of any association of genetically-predicted WHR with HF or PHD.

8.4.2. Associations of genetically-predicted BMI with HF and PHD

Consistent with findings from previous studies in individuals of European ancestry, this Chapter demonstrated positive associations between higher levels of genetically-predicted BMI and HF. The findings of the present study (HR per 1 SD higher genetically-predicted BMI: 1.43; 95% CI, 1.31–1.57) were similar to those in European ancestry populations reported by Hagg et al., 2015[218]. Hagg et al., 2015 constructed a GRS using 32 SNPs and demonstrated a 47% higher risk of HF per 1 SD higher levels of genetically-predicted BMI (HR 1.47; 95% CI 1.35–1.60). The findings of the present study were also consistent with the results of Kim et al., 2021 (OR per 1 SD increase: 1.66; 1.41–2.03)[224], Lind et al., 2021 (OR per 1 SD increase: 1.57; 1.48–1.68)[225] and Fall et al., 2013 (OR per 1 SD increase: 2.19; 1.14–4.26)[219] conducted in individuals of European ancestry.

However, the association of genetically-predicted BMI with HF in the present study was weaker than the findings of a meta-analysis by Benn et al., 2022 involving individuals of European descent in the Copenhagen City Heart Study, Copenhagen General Population Study, GIANT, HERMES and UK Biobank. Benn et al., 2022 reported a 4.4-fold higher odds of HF per 1 SD higher level of genetically-predicted BMI (OR 4.40; 2.93–6.58)[224]. Comparison with other studies was not possible because the authors did not provide the unit of change of the exposure: for instance, whether the risk of HF was associated with a 1 kg/m^2 higher BMI[221,226]. Li et al., 2024 reported a positive association of genetically-predicted BMI with HF, but it is difficult to validate the findings of the latter study as the authors neither provided information about the dataset used for the analysis nor the unit of change of genetically-predicted BMI associated with the reported risk of HF[227].

By contrast, the present analyses did not observe any association of genetically-predicted BMI with PHD, consistent with the observational analysis (**Chapter 6**) that demonstrated attenuation of the association of BMI with PHD after excluding prior disease and the first few years of follow-up. Hence, the present analysis does not provide sufficient evidence of any association of genetically-predicted BMI with PHD.

8.4.3. Association of genetically-predicted WHR with HF and PHD

MR studies investigating the causal associations of WHR with HF and PHD are limited. The present analysis demonstrated no evidence of any association of genetically-predicted WHR with HF. This finding is consistent with those of Hong et al., 2024 who showed no association of WHR with HF after accounting for BMI[221], indicating that any association between WHR and HF in the observational analysis was probably mediated through BMI. In addition, there was no significant association between genetically-predicted WHR with PHD. This finding is in line with the results of the prospective associations in Chapter 6, which showed an attenuation of the inverse association of WHR with PHD to the null after excluding the first few years of follow-up and prior disease. These findings indicate that the observational associations could be due to confounding and reverse causality bias.

8.4.4. Strengths and limitations

The present study had several limitations including the possibility of weak instrument bias, which would cause a loss of power and bias the MR association towards the confounded observational association. The F-statistics for the instruments for BMI and WHR were below 10, which increases the probability of weak instrument bias. MR results could be biased by horizontal pleiotropy. Although the analyses adjusted for potential confounders, horizontal pleiotropy cannot be completely excluded.

The strengths of this study included the availability of individual-level data, which permitted careful adjustments for confounding and exclusion of participants with COPD, tuberculosis and cancer at baseline to minimise reverse causality. The use of strong instruments minimised weak instrument bias.

8.5. Summary

The present analysis used genetic IVs to proxy the effects of higher levels of BMI and WHR with risk of HF and PHD. Consistent with the results of the observational analyses in Chapter 6, genetically-predicted BMI was associated with higher risks of HF. By contrast, there was no evidence of any association between genetically-predicted BMI and PHD. In addition, the MR showed no evidence of any association of genetically-predicted WHR with HF and PHD.

9. Discussion and conclusions

9.1. Introduction

Heart failure is a global health problem that has been estimated to affect about 56 million people worldwide in 2019, with a majority of the cases living in China[20]. The prevalence of HF is predicted to increase owing to the increasing prevalence of the risk factors for HF, including older age, obesity, hypertension and physical inactivity[20,23–25]. HF is associated with a 5-year mortality risk of 50%, a high risk of rehospitalisation and substantial economic costs for both individual patients and the healthcare system[1,25]. In contrast, PHD remains a poorly understood and less widely studied disease worldwide. The available epidemiological evidence on PHD has mostly been inferred from autopsy reports of patients with COPD and other chronic lung diseases[21,32]. The prognosis of PHD is much worse than that for HF, and the 1-year mortality risk is estimated to be over four-fold greater than that for HF[1,25].

Although prospective cohort studies and MR studies suggest causal associations of higher levels of SBP and adiposity with HF, uncertainties persist regarding the shape and strength of these associations, especially in non-Western populations such as China. In addition, no study has investigated the shape and strength of the associations of SBP and adiposity with PHD. Using data from the CKB, this thesis investigated the shape and strength of the associations of SBP and adiposity with HF and PHD. The aims of this thesis were to (1) estimate the hospitalisation and mortality rates of HF and PHD overall and by subgroups, (2) investigate the shape and strength of the associations of SBP and adiposity with HF and PHD in observational analyses and evaluate the extent to which these associations are due to intermediate factors, residual confounding and reverse causality, and (3) evaluate the causal relevance of the most promising associations from the observational analyses using MR.

9.2. Summary of the main findings in relation to the thesis objectives

Table 9.1 summarises the main findings of this thesis. Chapter 4 investigated the rates of first hospitalisation for HF and PHD and all-cause mortality rates after the first hospitalisations for HF and PHD, respectively. The hospitalisation rates for HF and PHD were high, and the rates increased log-linearly with older age. The hospitalisation rates for HF and PHD were greater in men than in women and in rural compared to urban areas. HF and especially PHD were associated with high short-term and long-term mortality. HF and PHD mortality rates were higher in older than in younger adults and in men than in women, with no rural-urban differences.

Chapter 5 investigated the shape and strength of the associations of usual SBP with HF and PHD. The findings demonstrated strong positive and log-linear associations of SBP with HF with no evidence of a threshold throughout the range of SBP studied. By contrast, there was no evidence of any material association of SBP with PHD. Chapter 7 used MR analysis to further explore the causal relevance of the associations of SBP with HF risk. Consistent with the observational associations presented in Chapter 5, the MR results demonstrated a positive association of higher levels of genetically-predicted SBP with higher risks of HF.

Chapter 6 explored the observational associations of adiposity (BMI and WHR) with HF and PHD. There were positive log-linear associations of BMI and WHR with HF with evidence of a threshold below which there was an inverse association (threshold for BMI: 22 kg/m^2) or a null association (threshold for WHR: 0.85), respectively. Analyses adjusting for other adiposity traits to evaluate the associations of BMI and WHR with HF independent of other adiposity measures demonstrated that the strengths of the associations of BMI and WHR with HF were similar. SBP and diabetes accounted for about 18% and 22% of the associations of BMI and WHR with HF, respectively. Unlike the associations with HF, BMI and WHR were inversely associated with PHD with

evidence of a threshold above which there was a positive association (threshold for BMI: 26 kg/m^2) or a null association (threshold for WHR: 0.95). However, after sensitivity analyses to minimise reverse causality, the associations of BMI and WHR with PHD attenuated to the null, suggesting that the observational associations were accounted for by reverse causality.

Since observational associations are chiefly limited by residual confounding and reverse causality, Chapter 8 used MR analysis that minimises these biases to investigate the causal relevance of the associations of BMI and WHR with HF and PHD. For BMI, genetically-predicted BMI was associated with higher risks of HF, consistent with the results of the observational analyses reported in Chapter 6**.** By contrast, there was no evidence of any association between genetically-predicted BMI and PHD. There was also no evidence of any association of genetically-predicted WHR with HF and PHD.

These findings provide robust evidence that SBP and BMI are risk factors for HF in Chinese adults owing to the consistent strength and direction of the associations across geographic regions, strong dose-response relationships, evidence of temporality, biological plausibility and evidence from MR. Evidence of a causal association between SBP and HF is also supported by RCTs showing reductions in the risk of HF by lowering SBP in individuals at high risk of CVD[228,229].

Table 9.1. Summary of the main findings of this thesis

Objectives	Design	Sample	Main findings	
			Heart failure	**Pulmonary heart disease**
Hospitalisation rates for HF and PHD (**Chapter 4**)*.	Prospective cohort	For HF: 511,960 For PHD: 512,243	**H**ospitalisation rates: 124.7 (121.9–127.6) per 10^5 py. Rates were ↑ Urban areas, older age, higher number of comorbidities ↓ Higher levels of education.	**H**ospitalisation rates: 98.8 (96.3–101.4) per 10^5 py. Rates were: ↑ Urban areas, older age and higher number of comorbidities ↓ Higher levels of education.
28-day case-fatality ratio (CFR) for HF and PHD (**Chapter 4**).	Prospective cohort	7,461 and 5,925 first HF and PHD hospitalisations.	The **28-day CFR** (18%) increased with older age.	The **28-day CFR** for PHD (53%) were high and ↑ Increased with older age and higher number of comorbidities ↑ Lower levels of education
5-year cumulative mortality risk (CMR) for HF and PHD (**Chapter 4**)	Prospective cohort	7,461 and 5,925 first HF and PHD hospitalisations.	The **5-year CMR** for HF (50%) were ↑ Men, older age, greater number of comorbidities ↓ Lower levels of education	The **5-year CMR** for PHD (70%) were ↑ Men, older age, greater number of comorbidities ↓ Lower levels of education
Associations of SBP with HF and PHD (**Chapters 5 & 7**)	Prospective cohort	488,755 participants with no prior history of heart disease, stroke or TIA.	HR per 10 mmHg higher usual SBP: 1.13 (95% CI 1.11–1.15).	HR per 10 mmHg higher usual SBP: 1.01 (95% CI 0.99–1.03).
	MR	94,859 included in the genetic sub-study	HR per 10 mmHg higher genetically-predicted SBP: 1.18 (95% CI 1.01–1.39).	Not applicable
Associations of genetically-predicted BMI and WHR with HF and PHD (**Chapter 6 & 8**)	Prospective cohort	479,553 participants with no prior rheumatic heart disease, coronary heart disease, stroke or TIA, cancer and tuberculosis.	**BMI** HR per 1 SD higher usual BMI: 1.21 (95% CI 1.18–1.25) at BMI >22 kg/m^2 **WHR** HR per 1 SD higher usual WHR: 1.17 (95% CI 1.13–1.21) at WHR >0.85.	**BMI** 0.61 (95% CI 0.59–0.63) at BMI <26 kg/m^2. **WHR** 0.77 (95% CI 0.74–0.81) at WHR levels >0.94
	MR	94,859 included in the genetic sub-study	**BMI** HR per 1 SD higher genetically-predicted BMI: 1.39 (95% CI 1.11–1.75). **WHR** HR per 1 SD higher genetically-predicted WHR: 1.08 (95% CI 0.73–1.60).	**BMI** HR per 1 SD higher genetically-predicted BMI: 1.16 (95% CI 0.94–1.43). **WHR** HR per 1 SD higher genetically-predicted WHR: 0.91 (95% CI 0.63–1.32).

BMI: Body mass index; GRS: Genetic risk score; GWAS: Genome-wide association study; py: person-years; SBP: Systolic blood pressure; TIA: transient ischaemic attack; WHR: Waist-to-hip ratio. * Analyses excluded first incident cases of HF and PHD diagnosed as outpatients.

9.3. Implications of the findings for population health

Despite the availability of new treatments for HF, the prognosis of the condition remains poor[1,2]. The global prevalence of HF is predicted to increase due to population growth, improvement in life expectancy, changes in the treatment of ischaemic heart disease and the increasing prevalence of modifiable risk factors including hypertension, obesity, diabetes and sedentary lifestyles[23,24]. Prioritising prevention of HF is, therefore, required to attain the sustainable development goals (SDGs) by improving population health (SDG 3) and reducing the economic burden of HF, thereby preventing poverty (SDG 1) and hunger (SDG 2)[230].

9.3.1. Targeted interventions to reduce HF and PHD hospitalisation and mortality

The high hospitalisation rates among specific groups, such as older adults and individuals with lower levels of education, highlight the need for health education programs to prevent HF and PHD that are tailored towards older individuals and individuals with less formal education, using culturally relevant educational tools[231–233]. In addition, this thesis highlights the need to raise public awareness of the diagnosis, prevention, complications and treatment of HF and PHD. Furthermore, enhanced population screening of individuals with comorbidities such as COPD or diabetes who are at high risk of HF and PHD using tools such as electrocardiography, handheld echocardiography and cardiac biomarkers may be considered to reinforce prevention, early diagnosis and treatment of HF and PHD[1,231,234,235].

There is a need for government authorities to improve availability, accessibility and affordability of evidence-based pharmacologic and non-pharmacologic preventive measures for HF, especially for older adults and individuals with comorbidities, low socioeconomic backgrounds and individuals living in underserved communities. Such interventions can include improving access to facilities for physical exercise, making

healthy foods affordable and subsidising essential medicines for treating HF risk factors, including BP and glucose-lowering medications[236].

In addition, there is a need to improve the affordability and accessibility of heart transplantation and ventricular assistive devices[237]. Heart transplantation remains the mainstay of treatment for patients with end-stage HF despite the availability of new treatments. However, only 660 HF transplants of the 3760 cases registered in 2019 were conducted in China[238]. Investing in scientific innovation and technology such as total artificial heart transplants and ventricular assistive devices could bridge the gap in the absence of heart transplants. In addition to promoting a culture of organ donation, there is a need to improve the affordability and accessibility of heart transplantation and ventricular assistive devices[237].

9.3.2. Primary prevention for HF and PHD

9.3.2.1. Systolic blood pressure

The Chinese government implemented nationwide programmes to improve healthy lifestyles and hypertension awareness, treatment and control[239]. These programmes led to a 5% reduction in the prevalence of hypertension between 2010 (29.6%) and 2018 (24.7%), but awareness (38%), treatment (35%), and control rates (12%) remained poor[149]. Chinese and international guidelines recommend SBP levels <140 mmHg as the target for preventing HF[239,240]. The findings of this thesis provide additional evidence that lower levels of SBP (down to 115 mmHg) than the current targets advocated by Chinese and international guidelines are associated with lower risks of HF. Strategies to implement lower target levels of SBP could have substantial benefits for primary prevention of HF.

Intensive SBP-lowering interventions have been shown to yield additional benefits in reducing the risk of cardiovascular diseases below the recommended BP targets. The

findings of this study provide additional support for the results of the SPRINT trial, which demonstrated that intensive SBP control (targeting SBP <120 mmHg) in individuals at high risk of CVD but without diabetes was associated with lower risk of a composite of CVD (including HF) than standard SBP control (targeting SBP <140 mmHg)[241]. Nevertheless, studies have shown that intensive SBP-lowering could increase the risks of serious adverse events, including acute renal injury and stroke[241–243]. However, a recent meta-analysis of nine RCTs involving ~33,000 participants demonstrated that intensive BP lowering was associated with a 26% reduced risk of non-fatal stroke[244]. Intensive BP lowering remains an ongoing area of research that can improve current hypertension control rates and prevent more cases of HF.

9.3.2.2. Adiposity

The prevalence of obesity in China has increased 3-fold between 2004 and 2018 (8.1%), reflecting China's rapid socioeconomic development[245]. Since 2010, China has implemented interventions to prevent obesity and non-communicable diseases, including the China Healthy Lifestyle for All Initiative and the National Demonstration Areas for Comprehensive Prevention[245–247]. Although most Chinese guidelines recommend having a healthy weight, it is unclear what specific targets should be met to prevent HF[73]. This thesis found that after sensitivity analyses to mitigate confounding and reverse causality, higher levels of BMI are associated with higher risks of HF down to a BMI level of ~20 kg/m^2, with no evidence of an inverse association at lower BMI levels. This finding contrasts with previous studies conducted in the Asian-Pacific Region, which showed an inverse association of BMI with HF[79]. The finding of the present thesis suggests that advocating for lower BMI targets could improve the benefits of primary prevention of HF. Unlike BMI, this thesis found no association of WHR with HF, indicating that BMI is a more relevant risk factor for HF than WHR in this population.

Weight loss interventions aimed at preventing HF should target BMI to optimise the benefits of such interventions.

This study showed that SBP and diabetes potentially accounted for about one-fifth of the association of BMI with HF. Although maintaining a healthy weight should be the primary management strategy for preventing overweight and obesity, interventions to prevent elevated SBP and diabetes among individuals with obesity should be considered to prevent HF, especially for individuals who are unable to undertake physical activity.

9.4. Strengths and limitations of the study

9.4.1. Strengths

This study had several strengths including the large number of events that permitted the estimation of hospitalisation rates, the strength of associations with high precision, and the shape of associations. The large number of events also enabled reliable assessment of non-trivial associations. Second, baseline exposures were carefully measured to minimise systematic error using standardised protocols across all study centres, delivering standardised training to staff involved in data collection and use of regularly calibrated devices for exposures such as SBP, BMI and WHR. A validated digital sphygmomanometer was used to measure BP to minimise systematic error; at least two measurements were taken five minutes apart to minimise random error. Third, the availability of repeat measurements permitted the correction of regression dilution bias. Fourth, loss to follow-up is a potential source of bias in prospective cohort studies[248]. The ability of the CKB to follow up with participants passively through linkage to electronic records, supplemented by active follow-up, ensured a negligible loss to follow-up rate of <1% in this study. Fifth, the collection of different adiposity traits enabled the assessments of the contribution of general and central adiposity to HF and PHD risks. Lastly, the availability of genotyping data on 100,000 CKB participants permitted the use of MR to mitigate reverse causality and investigate causal associations.

9.4.2. Limitations

This study also had several limitations that should be considered while interpreting the findings.

9.4.2.1. Selection bias

The non-probabilistic sampling of participants into the CKB, the low participation rate (~30%) and healthy volunteer effect[249]—volunteers who participate in population-based studies are healthier, more educated and from a higher social class than individuals from the population of interest—can result in selection bias, limiting the generalisation of the results of this thesis to the whole of China. However, the CKB study was not designed to be representative of the Chinese population; the study selected study areas and participants to maximise the ability to investigate the risk factors for major diseases and minimise loss to follow-up rates. Study areas were selected based on social diversity, population stability, differences in risk factor exposure, burden of major diseases and the quality of local morbidity and mortality registers[117]. The careful attention to detail of the study design to address various sources of bias when investigating epidemiological associations should ensure that the findings are valid. Selection bias is unlikely to affect the epidemiological associations of this thesis as participants were not recruited based on BP, adiposity or prior disease status.

9.4.2.2. Confounding

Although prospective cohort studies are susceptible to residual confounding, this thesis used several methodological approaches to control for these limitations. First, the multivariable regression analyses were adjusted for the major confounders based on information from the literature. In addition, continuous confounders were not categorised to prevent loss of information and minimise residual confounding[250–252]. Confounders were modelled in their linear (for example, BMI) form or an additional non-linear term was added to the model (for example, BMI^2) depending on whether the confounders had

a linear or non-linear relationship with the outcome. Furthermore, controlling for major confounders did not substantially alter the associations of SBP and WHR with HF, indicating that additional adjustments by any unmeasured confounders are unlikely to explain the remaining observed associations. Hence, the associations of SBP and BMI with HF in this thesis are less likely to be affected by residual confounding.

9.4.2.3. HF and PHD case ascertainment

HF and PHD cases in CKB have not been adjudicated, and hence, the possibility of some misclassification of these cases cannot be excluded. However, HF diagnosis in China follows international guidelines and is based on patient medical history, clinical examination, laboratory examination and cardiac imaging[2]. Similarly, the diagnosis of PHD in China is based on the presence of potential causes of PHD (e.g., COPD); clinical syndrome of right HF, including bilateral pedal oedema; and structural or functional cardiac abnormalities on cardiac imaging[2]. In this thesis, misclassifications in the outcomes are most likely to be non-differential, which can bias the slope of associations towards the null[253].

The heterogeneity in the associations between SBP and adiposity across different causes of HF or PHD suggests that studies treating HF or PHD as a single entity may underestimate the true strength of the association[86]. Because these associations are likely to vary depending on the specific cause of HF or PHD, combining them into a single analysis can mask potentially important relationships with specific subtypes. This presents a significant challenge for studies investigating the risk factors for both diseases.

This study is likely to have underestimated the HF and PHD hospitalisation rates and mortality risks as some hospitalisation events might not be indexed by all participating

hospitals. In addition, it is possible to have missed mild cases of HF and PHD, which participants might not have deemed necessary for a hospital consultation.

9.4.2.4. Reverse causality

Observational studies are limited by reverse causality. Although MR mitigates reverse causality because the outcome cannot alter germline genetic variations, outcomes (such as PHD) that can affect the exposure (e.g., BMI) can introduce reverse causality in MR studies. In observational analyses, several approaches were used to minimise reverse causality by excluding the first few years of follow-up and prior disease. In the MR analyses, prior medical conditions (e.g., tuberculosis and COPD) that affect the exposure (e.g., BMI) were excluded before estimating the genetic IV-exposure association. However, this study cannot completely exclude the possibility of reverse causality.

9.4.2.5. Horizontal pleiotropy in MR

The MR approach assumes that the effect of genetic IVs on the outcome is entirely through the measured exposure (vertical pleiotropy) and not through alternative pathways that do not involve the exposure (horizontal pleiotropy)[185]. Although MR analyses in this thesis used uncorrelated SNPs at genetic IVs and adjusted for confounders, bias from pleiotropy cannot be excluded fully. Pleiotropy-robust methods were used to assess the robustness of the main causal estimate, but these methods were likely limited by low statistical power from small sample size.

9.4.2.6. Non-linear associations in MR

Although the three core assumptions of MR are sufficient for testing hypotheses of a causal association, correct specification of the shape of the association of the exposure and the outcome is needed to obtain reliable estimates of causal effects. However, classic MR studies assume a linear relationship between the exposure and outcome.

Extensions to linear MR have been proposed to investigate non-linear associations and have been used to investigate the shape of the association between BMI and mortality[254]. These methods rely on one of two key assumptions to be met. First, the effect of the genetic IV on the exposure should be the same for everyone in the population under study, known as the homogeneity assumption[255–257]. Second, the effect of the genetic IV on the exposure is either increasing or decreasing for everyone in the population, referred to as the monotonicity assumption[255,256]. The residual method, which involves stratifying individuals in the population by the residuals from regressing the exposure on the genetic IV[255], relies strongly on the homogeneity assumption and has been shown to provide biased estimates if this assumption is violated[255,258,259]. Hence, the non-parametric doubly ranked method was recently developed and does not rely strongly on the homogeneity assumption[255]. However, this method has also been shown to provide implausible findings. For instance, the method was shown to produce biased estimates in negative control analyses where the association is expected to be null[260]. The origin of the bias remains unclear and is still an evolving area of research[258]. This thesis did not investigate non-linear associations using MR owing to the uncertainties around current non-linear MR methods.

9.5. Opportunities for future research

9.5.1. Expanding on the risk factors for HF and PHD

Although the risk factors of HF have been extensively studied in Western countries, understanding the risk factors in non-Western countries is important for developing context-specific public health education and awareness messages and guiding local clinical practices and policies for prevention of HF in these countries. In addition, despite the high 5-year mortality rate of PHD, this thesis did not find evidence of any important association of SBP or adiposity with PHD. It is worthwhile to explore the causal relevance of other potential risk factors for HF and PHD, including diabetes, smoking,

alcohol drinking, blood lipids, atrial fibrillation, physical activity and autoimmune disease. Future studies could explore other measures of adiposity, including imaging-derived adiposity measures, and their independent contribution to HF and PHD risks.

9.5.2. Risk predictions of HF and PHD

Prediction models for incident HF and PHD in China are limited owing to the lack of relevant data from prospective cohort studies. A search of PubMed and Google Scholar showed no study on risk prediction models for incident HF or PHD using the following search strategies: *("risk prediction" OR predictive) AND ("heart failure" OR "pulmonary heart disease" OR "cor pulmonale") AND inciden* AND China.* All of the studies included in a systematic review and meta-analysis conducted by Yang et al., 2015 were from the US and Europe[261]. The CKB provides a unique opportunity to combine conventional with genetic and proteomic determinants of incident HF and PHD to optimise existing prediction models for use in the Chinese population. Building context-specific prediction models could help improve the prediction of individuals who are likely to develop incident HF or PHD to guide early and targeted interventions for preventing HF or PHD and their complications.

9.5.3. Investigating drug targets

The ability to integrate genetic, proteomic and phenotypic data permits the investigation of drug targets for HF or PHD and their risk factors in individuals of East Asian ancestry. For instance, Lian et al., 2023 used a drug-target MR approach to investigate the associations of the genetically-proxied effects of antihypertensive drugs with the risk of genetically-predicted HF among individuals of European ancestry[190]. A similar approach could be used in individuals of East Asian ancestry. The findings from such studies could provide insights into differences in the effects of antihypertensive medications between East Asians and Europeans and, therefore, inform primary prevention

strategies for HF. Similar approaches could be used to identify new drugs or repurpose existing drugs for prevention of HF or PHD.

In a recently published study, Shiyang et al. 2024 used proteomic data on individuals of European ancestry from the UK Biobank to investigate the causal relevance of 3622 plasma proteins with PHD in an MR paradigm[262]. That study identified the calcium/calmodulin-dependent protein kinase I protein, which is involved in the calcium signalling pathway[263], as being causally associated with a higher risk of PHD. The association of this protein with the risk of PHD was replicated using data from the FinnGen consortium. Future studies in diverse studies are needed to investigate the proteomic signature of PHD and identify potential drug targets to treat the condition.

9.5.4. Investigating the mechanisms linking risk factors to HF and PHD

Genetic, proteomic and phenotypic data can be leveraged to assess possible mechanisms through which a risk factor affects the outcome by investigating proteins or biological pathways that are upregulated or downregulated in the presence of a risk factor and how this leads to disease[264]. Such approaches can be used to identify novel drug targets or understand the mechanism by which interventions prevent disease outcomes.

9.5.5. Clinical trials for intensive SBP and adiposity lowering for HF prevention

The findings from this study suggest added benefits of reducing SBP and BMI below the recommended thresholds for prevention of HF. Similar to the SPRINT trial[241], randomized controlled trials can be conducted using pharmacological (e.g., semaglutide) and non-pharmacological (e.g., physical activity) interventions to investigate the effects of maintaining BMI below the recommended targets to reduce the risk of HF among individuals at high risk of cardiovascular diseases. However, some concerns of such trials would be the impact of intensive body weight lowering on patients at risk of

cardiovascular disease. Individuals with diabetes or hypertension who have increased risks of HF could be suitable groups for trials of intensive weight-lowering for prevention of HF in addition to other disease outcomes.

9.6. Conclusions

This thesis demonstrated high hospitalisation and high mortality rates for HF and PHD in Chinese adults, with substantial differences by age, study area, education and comorbidities. The thesis also investigated usual SBP and adiposity (BMI and WHR) as risk factors for HF and PHD. SBP was positively and log-linearly associated with HF, with no evidence of a threshold across the range of SBP studied. However, SBP was not associated with PHD. BMI was positively associated with HF at BMI levels >22 kg/m^2. However, BMI was not causally associated with PHD. SBP and diabetes mediated over a third of the association of BMI with HF. Furthermore, there was no evidence of any causal association of WHR with HF and PHD.

The findings of this thesis highlight the need for health educational programmes on HF and PHD targeted towards older adults and individuals with lower levels of education and greater numbers of comorbidities. In addition, current recommendations to maintain a healthy weight and improve the diagnosis, awareness and treatment of hypertension and diabetes are expected to reduce the incidence of HF. However, based on the findings of this thesis, setting lower targets for SBP and weight will have added benefits in preventing HF. Future studies are needed to investigate the risk factors for HF and PHD in diverse populations. Genetic and proteomic data could be leveraged to investigate the risk factors for HF and PHD, understand the mechanisms of epidemiological associations and discover or repurpose existing drug targets for prevention of HF and PHD.

References

1. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. European Heart Journal. 2021;42:3599–726.

2. Heart Failure Group of Chinese Society of Cardiology of Chinese Medical Association, Chinese Heart Failure Association of Chinese Medical Doctor Association, Editorial Board of Chinese Journal of Cardiology. [Chinese guidelines for the diagnosis and treatment of heart failure 2018]. Zhonghua Xin Xue Guan Bing Za Zhi. 2018;46:760–89.

3. Bozkurt B, Coats AJ, Tsutsui H, Abdelhamid M, Adamopoulos S, Albert N, et al. Universal Definition and Classification of Heart Failure: A Report of the Heart Failure Society of America, Heart Failure Association of the European Society of Cardiology, Japanese Heart Failure Society and Writing Committee of the Universal Definition of Heart Failure. Journal of Cardiac Failure. 2021;27:387–413.

4. Konstam MA, Kiernan MS, Bernstein D, Bozkurt B, Jacob M, Kapur NK, et al. Evaluation and Management of Right-Sided Heart Failure: A Scientific Statement From the American Heart Association. Circulation. 2018;137:e578–622.

5. Fauci A, Braunwald E, Kasper D, Hauser S, Longo D, Jameson J, et al. Harrison's Principles of Internal Medicine, 17th Edition. Mcgraw-hill; 2008.

6. Lee G, McKee G, Protopapas A, Jones ID. Anatomy and physiology of the healthy heart. In: Jennings C, Astin F, Fitzsimons D, Lambrinou E, Neubeck L, Thompson DR, et al., editors. ESC Textbook of Cardiovascular Nursing. Oxford University Press; 2021. p. 0.

7. Bozkurt B. Contemporary pharmacological treatment and management of heart failure. Nat Rev Cardiol. 2024;:1–11.

8. Ge Z, Li A, McNamara J, dos Remedios C, Lal S. Pathogenesis and pathophysiology of heart failure with reduced ejection fraction: translation to human studies. Heart Fail Rev. 2019;24:743–58.

9. Houston BA, Brittain EL, Tedford RJ. Right Ventricular Failure. N Engl J Med. 2023;388:1111–25.

10. Kemp CD, Conte JV. The pathophysiology of heart failure. Cardiovascular Pathology. 2012;21:365–71.

11. Schwinger RHG. Pathophysiology of heart failure. Cardiovasc Diagn Ther. 2021;11:263–76.

12. Metra M, Teerlink JR. Heart failure. The Lancet. 2017;390:1981–95.

13. Clark AL. Right heart failure. In: Clark AL, Gardner RS, McDonagh TA, editors. Oxford Textbook of Heart Failure. Oxford University Press; 2022. p. 401–10.

14. Kelder JC, Cramer MJ, van Wijngaarden J, van Tooren R, Mosterd A, Moons KGM, et al. The Diagnostic Value of Physical Examination and Additional Testing in Primary Care Patients With Suspected Heart Failure. Circulation. 2011;124:2865–73.

15. Gohar A, Rutten FH, den Ruijter H, Kelder JC, von Haehling S, Anker SD, et al. Mid-regional pro-atrial natriuretic peptide for the early detection of non-acute heart failure. European Journal of Heart Failure. 2019;21:1219–27.

16. Yoo B-S. Clinical Significance of B-type Natriuretic Peptide in Heart Failure. J Lifestyle Med. 2014;4:34–8.

17. Gorter TM, Van Veldhuisen DJ, Bauersachs J, Borlaug BA, Celutkiene J, Coats AJS, et al. Right heart dysfunction and failure in heart failure with preserved ejection fraction: mechanisms and management. Position statement on behalf of the Heart Failure Association of the European Society of Cardiology: Right ventricular function in HFpEF. Eur J Heart Fail. 2018;20:16–37.

18. Rao VN, Andrews J, Applefeld WN, Gray JM, Molinger J, Felker GM, et al. Quantitative Blood Volume Analysis and Hemodynamic Measures of Vascular Compliance in Patients With Worsening Heart Failure. Journal of Cardiac Failure. 2022;28:1469–74.

19. McKee PA, Castelli WP, McNamara PM, Kannel WB. The natural history of congestive heart failure: the Framingham study. N Engl J Med. 1971;285:1441–6.

20. Yan T, Zhu S, Yin X, Xie C, Xue J, Zhu M, et al. Burden, Trends, and Inequalities of Heart Failure Globally, 1990 to 2019: A Secondary Analysis Based on the Global Burden of Disease 2019 Study. Journal of the American Heart Association. 2023;12:e027852.

21. Bragazzi NL, Zhong W, Shu J, Abu Much A, Lotan D, Grupper A, et al. Burden of heart failure and underlying causes in 195 countries and territories from 1990 to 2017. European Journal of Preventive Cardiology. 2021;28:1682–90.

22. Lloyd-Jones DM, Larson MG, Leip EP, Beiser A, D'Agostino RB, Kannel WB, et al. Lifetime Risk for Developing Congestive Heart Failure: The Framingham Heart Study. Circulation. 2002;106:3068–72.

23. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, et al. Heart Disease and Stroke Statistics—2017 Update. Circulation. 2017;135:e146–603.

24. Celermajer DS, Chow CK, Marijon E, Anstey NM, Woo KS. Cardiovascular Disease in the Developing World: Prevalences, Patterns, and the Potential of Early Disease Detection. Journal of the American College of Cardiology. 2012;60:1207–16.

25. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovascular Research. 2022;:cvac013.

26. GBD Compare. Institute for Health Metrics and Evaluation. http://vizhub.healthdata.org/gbd-compare. Accessed 31 Oct 2023.

27. Levy D, Kenchaiah S, Larson MG, Benjamin EJ, Kupka MJ, Ho KKL, et al. Long-Term Trends in the Incidence of and Survival with Heart Failure. N Engl J Med. 2002;347:1397–402.

28. Jones NR, Roalfe AK, Adoki I, Hobbs FDR, Taylor CJ. Survival of patients with chronic heart failure in the community: a systematic review and meta-analysis. European Journal of Heart Failure. 2019;21:1306–25.

29. Cai A, Zheng C, Qiu J, Fonarow GC, Lip GYH, Feng Y, et al. Prevalence of Heart Failure Stages in the General Population and Implications for Heart Failure Prevention: Reports from the China Hypertension Survey 2012-2015. European Journal of Preventive Cardiology. 2023;:zwad223.

30. Li Y, Sun X-L, Qiu H, Qin J, Li C-S, Yu X-Z, et al. Long-term outcomes and independent predictors of mortality in patients presenting to emergency departments with acute heart failure in Beijing: a multicenter cohort study with a 5-year follow-up. Chinese Medical Journal. 2021;134:1803–11.

31. Wang H, Chai K, Du M, Wang S, Cai J-P, Li Y, et al. Prevalence and Incidence of Heart Failure Among Urban Patients in China: A National Population-Based Analysis. Circulation: Heart Failure. 2021;14:e008406.

32. Leschke M, Wädlich A. Rechtsherzinsuffizienz und Cor pulmonale. Internist. 2007;48:948–60.

33. Orea-Tejeda A, Bozada-Gutiérrez K, Pineda-Juárez J, González-Islas D, Santellano-Juárez B, Keirns-Davies C, et al. Right Heart Failure as a Risk for Stroke in Patients with Chronic Obstructive Pulmonary Disease: A Case-Control Study. J Stroke Cerebrovasc Dis. 2017;26:2988–93.

34. Orea-Tejeda A, Navarrete-Peñaloza AG, Verdeja-Vendrell L, Jiménez-Cepeda A, González-Islas DG, Hernández-Zenteno R, et al. Right heart failure as a risk factor for severe exacerbation in patients with chronic obstructive pulmonary disease: Prospective cohort study. Clin Respir J. 2018;12:2635–41.

35. Andersson C, Hansen PW, Steffensen IE, Andreasen C, Weeke PE, Køber L, et al. Mortality associated with cardiovascular drugs in patients with chronic obstructive pulmonary disease and right-sided heart failure - A danish nationwide registry-based study. Eur J Intern Med. 2019;63:56–61.

36. Xu Y, Liu Z, Qi J, Wang L, Zhou M, Yin P. Pulmonary Heart Disease Associated Mortality — China, 2014–2021. China CDC Wkly. 2023;5:779–84.

37. Lip GYH, Hall JE. Comprehensive Hypertension. Mosby Elsevier; 2007.

38. Curran J. The Yellow Emperor’s Classic of Internal Medicine. BMJ. 2008;336:777.

39. Aird WC. Discovery of the cardiovascular system: from Galen to William Harvey. J Thromb Haemost. 2011;9 Suppl 1:118–29.

40. O’Brien E, Fitzgerald D. The history of blood pressure measurement. J Hum Hypertens. 1994;8:73–84.

41. Booth J. A short history of blood pressure measurement. Proc R Soc Med. 1977;70:793–9.

42. Mancia(Chairperson) G, Kreutz(Co-Chair) R, Brunström M, Burnier M, Grassi G, Januszewicz A, et al. 2023 ESH Guidelines for the management of arterial hypertension The Task Force for the management of arterial hypertension of the European Society of Hypertension Endorsed by the International Society of Hypertension (ISH) and the European Renal Association (ERA). Journal of Hypertension. :10.1097/HJH.0000000000003480.

43. 2018 Chinese Guidelines for Prevention and Treatment of Hypertension—A report of the Revision Committee of Chinese Guidelines for Prevention and Treatment of Hypertension. J Geriatr Cardiol. 2019;16:182–241.

44. Campbell NRC, Paccot Burnens M, Whelton PK, Angell SY, Jaffe MG, Cohn J, et al. 2021 World Health Organization guideline on pharmacological treatment of hypertension: Policy implications for the region of the Americas. Lancet Reg Health Am. 2022;9:100219.

45. Whelton PK, Carey RM, Aronow WS, Casey DE, 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: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension. 2018;71:1269–324.

46. Stanaway JD, Afshin A, Gakidou E, Lim SS, Abate D, Abate KH, et al. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392:1923–94.

47. Zhou B, Carrillo-Larco RM, Danaei G, Riley LM, Paciorek CJ, Stevens GA, et al. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. The Lancet. 2021;398:957–80.

48. Abarca-Gómez L, Abdeen ZA, Hamid ZA, Abu-Rmeileh NM, Acosta-Cazares B, Acuin C, et al. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. The Lancet. 2017;390:2627–42.

49. Zhang M, Wang L, Wu J, Huang Z, Zhao Z, Zhang X, et al. Data Resource Profile: China Chronic Disease and Risk Factor Surveillance (CCDRFS). International Journal of Epidemiology. 2022;51:e1–8.

50. Chobufo MD, Gayam V, Soluny J, Rahman EU, Enoru S, Foryoung JB, et al. Prevalence and control rates of hypertension in the USA: 2017–2018. Int J Cardiol Hypertens. 2020;6:100044.

51. Hisamatsu T, Miura K. Epidemiology and control of hypertension in Japan: a comparison with Western countries. J Hum Hypertens. 2021;:1–8.

52. Rahimi K, Bidel Z, Nazarzadeh M, Copland E, Canoy D, Ramakrishnan R, et al. Pharmacological blood pressure lowering for primary and secondary prevention of cardiovascular disease across different levels of blood pressure: an individual participant-level data meta-analysis. The Lancet. 2021;397:1625–36.

53. Kenchaiah S, Narula J, Vasan RS. Risk factors for heart failure. Medical Clinics of North America. 2004;88:1145–72.

54. Eryd SA, Gudbjörnsdottir S, Manhem K, Rosengren A, Svensson A-M, Miftaraj M, et al. Blood pressure and complications in individuals with type 2 diabetes and no previous cardiovascular disease: national population based cohort study. BMJ. 2016;354:i4070.

55. Zhao W, Katzmarzyk PT, Horswell R, Li W, Wang Y, Johnson J, et al. Blood pressure and heart failure risk among diabetic patients. Int J Cardiol. 2014;176:125–32.

56. Schnabel E, Nowak D, Brasche S, Wichmann H-E, Heinrich J. Association between lung function, hypertension and blood pressure medication. Respiratory Medicine. 2011;105:727–33.

57. Lacey B, Lewington S, Clarke R, Kong XL, Chen Y, Guo Y, et al. Age-specific association between blood pressure and vascular and non-vascular chronic diseases in 0·5 million adults in China: a prospective cohort study. The Lancet Global Health. 2018;6:e641–9.

58. Sattar N, Preiss D. Reverse Causality in Cardiovascular Epidemiological Research: More Common Than Imagined? Circulation. 2017;135:2369–72.

59. Imaizumi Y, Eguchi K, Kario K. Lung Disease and Hypertension. Pulse. 2014;2:103–12.

60. Singla P, Bardoloi A, Parkash AA. Metabolic effects of obesity: A review. World J Diabetes. 2010;1:76–88.

61. Hu FB, Hu F, editors. Measurements of Adiposity and Body Composition. In: Obesity Epidemiology. Oxford University Press; 2008. p. 0.

62. Bacon S. Central Adiposity. In: Gellman MD, Turner JR, editors. Encyclopedia of Behavioral Medicine. New York, NY: Springer; 2013. p. 368–9.

63. Powell-Wiley TM, Poirier P, Burke LE, Després J-P, Gordon-Larsen P, Lavie CJ, et al. Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation. 2021;143:e984–1010.

64. Collaboration TERF. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. The Lancet. 2011;377:1085–95.

65. Canoy D, Boekholdt SM, Wareham N, Luben R, Welch A, Bingham S, et al. Body Fat Distribution and Risk of Coronary Heart Disease in Men and Women in the European Prospective Investigation Into Cancer and Nutrition in Norfolk Cohort. Circulation. 2007;116:2933–43.

66. Rimm EB, Stampfer MJ, Giovannucci E, Ascherio A, Spiegelman D, Colditz GA, et al. Body size and fat distribution as predictors of coronary heart disease among middle-aged and older US men. Am J Epidemiol. 1995;141:1117–27.

67. Lear SA, Humphries KH, Kohli S, Birmingham CL. The use of BMI and waist circumference as surrogates of body fat differs by ethnicity. Obesity (Silver Spring). 2007;15:2817–24.

68. Williams R, Periasamy M. Genetic and Environmental Factors Contributing to Visceral Adiposity in Asian Populations. Endocrinol Metab (Seoul). 2020;35:681–95.

69. Lear SA, Humphries KH, Kohli S, Chockalingam A, Frohlich JJ, Birmingham CL. Visceral adipose tissue accumulation differs according to ethnic background: results of the Multicultural Community Health Assessment Trial (M-CHAT)2. The American Journal of Clinical Nutrition. 2007;86:353–9.

70. Yoon K-H, Lee J-H, Kim J-W, Cho JH, Choi Y-H, Ko S-H, et al. Epidemic obesity and type 2 diabetes in Asia. The Lancet. 2006;368:1681–8.

71. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants. The Lancet. 2016;387:1377–96.

72. Murray CJL, Aravkin AY, Zheng P, Abbafati C, Abbas KM, Abbasi-Kangevari M, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020;396:1223–49.

73. Pan X-F, Wang L, Pan A. Epidemiology and determinants of obesity in China. The Lancet Diabetes & Endocrinology. 2021;9:373–92.

74. Wang L, Zhou B, Zhao Z, Yang L, Zhang M, Jiang Y, et al. Body-mass index and obesity in urban and rural China: findings from consecutive nationally representative surveys during 2004–18. The Lancet. 2021;398:53–63.

75. Oguntade AS, Islam N, Malouf R, Taylor H, Jin D, Lewington S, et al. Body Composition and Risk of Incident Heart Failure in 1 Million Adults: A Systematic Review and Dose–Response Meta-Analysis of Prospective Cohort Studies. Journal of the American Heart Association. 2023;12:e029062.

76. Aune Dagfinn, Sen Abhijit, Norat Teresa, Janszky Imre, Romundstad Pål, Tonstad Serena, et al. Body Mass Index, Abdominal Fatness, and Heart Failure Incidence and Mortality. Circulation. 2016;133:639–49.

77. Kim MS, Kim WJ, Khera AV, Kim JY, Yon DK, Lee SW, et al. Association between adiposity and cardiovascular outcomes: an umbrella review and meta-analysis of observational and Mendelian randomization studies. European Heart Journal. 2021;42:3388–403.

78. Deng H, Duan X, Huang J, Zheng M, Lao M, Weng F, et al. Association of adiposity with risk of obstructive sleep apnea: a population-based study. BMC Public Health. 2023;23:1835.

79. Asia Pacific Cohort Studies Collaboration, Huxley RR, Barzi F, Woo J, Giles G, Lam TH, et al. A comparison of risk factors for mortality from heart failure in Asian and non-Asian populations: An overview of individual participant data from 32 prospective cohorts from the Asia-Pacific Region. BMC Cardiovasc Disord. 2014;14:61.

80. Zhang X, Chen H, Gu K, Chen J, Jiang X. Association of Body Mass Index with Risk of Chronic Obstructive Pulmonary Disease: A Systematic Review and Meta-Analysis. COPD: Journal of Chronic Obstructive Pulmonary Disease. 2021;18:101–13.

81. Collaboration PS. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. The Lancet. 2009;373:1083–96.

82. Li J, Zhu L, Wei Y, Lv J, Guo Y, Bian Z, et al. Association between adiposity measures and COPD risk in Chinese adults. Eur Respir J. 2020;55:1901899.

83. Floud S, Simpson RF, Balkwill A, Brown A, Goodill A, Gallacher J, et al. Body mass index, diet, physical inactivity, and the incidence of dementia in 1 million UK women. Neurology. 2020;94:e123–32.

84. Brenowitz WD. Invited Commentary: Body Mass Index and Risk of Dementia—Potential Explanations for Life-Course Differences in Risk Estimates and Future Research Directions. American Journal of Epidemiology. 2021;190:2511–4.

85. Kivimäki M, Luukkonen R, Batty GD, Ferrie JE, Pentti J, Nyberg ST, et al. Body mass index and risk of dementia: Analysis of individual-level data from 1.3 million individuals. Alzheimers Dement. 2018;14:601–9.

86. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. The Lancet. 2002;360:1903–13.

87. Lee DS, Massaro JM, Wang TJ, Kannel WB, Benjamin EJ, Kenchaiah S, et al. Antecedent Blood Pressure, Body Mass Index, and the Risk of Incident Heart Failure in Later Life. Hypertension. 2007;50:869–76.

88. Randolph TC, Greiner MA, Egwim C, Hernandez AF, Thomas KL, Curtis LH, et al. Associations Between Blood Pressure and Outcomes Among Blacks in the Jackson Heart Study. JAHA. 2016;5:e003928.

89. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1·25 million people. The Lancet. 2014;383:1899–911.

90. Magnussen C, Niiranen TJ, Ojeda FM, Gianfagna F, Blankenberg S, Vartiainen E, et al. Sex-Specific Epidemiology of Heart Failure Risk and Mortality in Europe: Results From the BiomarCaRE Consortium. JACC: Heart Failure. 2019;7:204–13.

91. Choi Y-J, Kim S-H, Kang S-H, Yoon C-H, Lee H-Y, Youn T-J, et al. Reconsidering the cut-off diastolic blood pressure for predicting cardiovascular events: a nationwide population-based study from Korea. European Heart Journal. 2019;40:724–31.

92. Borné Y, Engström G, Essén B, Hedblad B. Immigrant status and increased risk of heart failure: the role of hypertension and life-style risk factors. BMC Cardiovascular Disorders. 2012;12:20.

93. Butler J, Kalogeropoulos AP, Georgiopoulou VV, Bibbins-Domingo K, Najjar SS, Sutton-Tyrrell KC, et al. Systolic blood pressure and incident heart failure in the elderly. The Cardiovascular Health Study and the Health, Ageing and Body Composition Study. Heart. 2011;97:1304–11.

94. Baffour PK, Jahangiry L, Jain S, Sen A, Aune D. Blood pressure, hypertension and the risk of heart failure: A systematic review and meta-analysis of cohort studies. European Journal of Preventive Cardiology. 2023;:zwad344.

95. Huxley RR, Barzi F, Woo J, Giles G, Lam TH, Rahimi K, et al. A comparison of risk factors for mortality from heart failure in Asian and non-Asian populations: An overview of individual participant data from 32 prospective cohorts from the Asia-Pacific Region. BMC Cardiovascular Disorders. 2014;14:61.

96. Cui R, Iso H, Tanabe N, Watanabe Y, Tamakoshi A, Jacc Study Group. Association Between Weight Change Since 20 Years of Age With Mortality From Myocardial Infarction and Chronic Heart Failure in the Japan Collaborative Cohort (JACC) Study. Circ J. 2014;78:649–55.

97. Borné Y, Hedblad B, Essén B, Engström G. Anthropometric measures in relation to risk of heart failure hospitalization: a Swedish population-based cohort study. European Journal of Public Health. 2014;24:215–20.

98. Van Lieshout MAW, Verwoert GC, Mattace-Raso FUS, Zillikens MC, Sijbrands EJ, Deckers JW, et al. Measures of body composition and risk of heart failure in the elderly: The Rotterdam study. J Nutr Health Aging. 2011;15:393–7.

99. Hu G, Jousilahti P, Antikainen R, Katzmarzyk PT, Tuomilehto J. Joint effects of physical activity, body mass index, waist circumference, and waist-to-hip ratio on the risk of heart failure. Circulation. 2010;121:237–44.

100. Loehr LR, Rosamond WD, Poole C, Marie McNeill A, Chang PP, Folsom AR, et al. Association of Multiple Anthropometrics of Overweight and Obesity With Incident Heart Failure: The Atherosclerosis Risk in Communities Study. Circ: Heart Failure. 2009;2:18–24.

101. Levitan EB, Yang AZ, Wolk A, Mittleman MA. Adiposity and Incidence of Heart Failure Hospitalization and Mortality: A Population-Based Prospective Study. Circ: Heart Failure. 2009;2:202–8.

102. Björck L, Lundberg C, Schaufelberger M, Lissner L, Adiels M, Rosengren A. Body mass index in women aged 18 to 45 and subsequent risk of heart failure. Eur J Prev Cardiolog. 2020;27:1165–74.

103. Janszky I, Romundstad P, Laugsand LE, Vatten LJ, Mukamal KJ, Mørkedal B. Weight and weight change and risk of acute myocardial infarction and heart failure – the HUNT Study. J Intern Med. 2016;280:312–22.

104. Xing Z, Xiao B, Hu X, Chai X. Relationship Between Regional Adiposity Distribution and Incident Heart Failure in General Populations without Cardiovascular Disease. The American Journal of Medicine. 2023;136:277-283.e2.

105. Kokkinos P, Faselis C, Franklin B, Lavie CJ, Sidossis L, Moore H, et al. Cardiorespiratory fitness, body mass index and heart failure incidence. Eur J Heart Fail. 2019;21:436–44.

106. Björck L, Novak M, Schaufelberger M, Giang KW, Rosengren A. Body weight in midlife and long-term risk of developing heart failure-a 35-year follow-up of the primary prevention study in Gothenburg, Sweden. BMC Cardiovasc Disord. 2015;15:19.

107. Djoussé L, Bartz TM, Ix JH, Zieman SJ, Delaney JA, Mukamal KJ, et al. Adiposity and Incident Heart Failure in Older Adults: the Cardiovascular Health Study. Obesity. 2012;20:1936–41.

108. Ndumele CE, Matsushita K, Lazo M, Bello N, Blumenthal RS, Gerstenblith G, et al. Obesity and Subtypes of Incident Cardiovascular Disease. JAHA. 2016;5:e003921.

109. Kenchaiah S, Sesso HD, Gaziano JM. Body Mass Index and Vigorous Physical Activity and the Risk of Heart Failure Among Men. Circulation. 2009;119:44–52.

110. Thrainsdottir IS, Aspelund T, Gudnason V, Malmberg K, Sigurdsson G, Thorgeirsson G, et al. Increasing glucose levels and BMI predict future heart failure Experience from the Reykjavík Study. European J of Heart Fail. 2007;9:1051–7.

111. Murphy NF, MacIntyre K, Stewart S, Hart CL, Hole D, McMurray JJV. Long-term cardiovascular consequences of obesity: 20-year follow-up of more than 15 000 middle-aged men and women (the Renfrew–Paisley study). European Heart Journal. 2006;27:96–106.

112. Kenchaiah S, Evans JC, Levy D, Wilson PWF, Benjamin EJ, Larson MG, et al. Obesity and the Risk of Heart Failure. New England Journal of Medicine. 2002;347:305–13.

113. Wilhelmsen L, Rosengren A, Eriksson H, Lappas G. Heart failure in the general population of men – morbidity, risk factors and prognosis. Journal of Internal Medicine. 2001;249:253–61.

114. Kenchaiah S, Narula J, Vasan RS. Risk factors for heart failure. Medical Clinics of North America. 2004;88:1145–72.

115. Geissbühler M, Hincapié CA, Aghlmandi S, Zwahlen M, Jüni P, da Costa BR. Most published meta-regression analyses based on aggregate data suffer from methodological pitfalls: a meta-epidemiological study. BMC Medical Research Methodology. 2021;21:123.

116. Richardson M, Garner P, Donegan S. Interpretation of subgroup analyses in systematic reviews: A tutorial. Clinical Epidemiology and Global Health. 2019;7:192–8.

117. Chen Z, Lee L, Chen J, Collins R, Wu F, Guo Y, et al. Cohort Profile: The Kadoorie Study of Chronic Disease in China (KSCDC). International Journal of Epidemiology. 2005;34:1243–9.

118. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. International Journal of Epidemiology. 2011;40:1652–66.

119. Lewington S, Li L, Sherliker P, Guo Y, Millwood I, Bian Z, et al. Seasonal variation in blood pressure and its relationship with outdoor temperature in 10 diverse regions of China: the China Kadoorie Biobank. Journal of Hypertension. 2012;30:1383.

120. Lewington S. Lexis Expansion- Age-at-Risk Adjustment for Survival Analysis. 2013.

121. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc. 2011;43:1575–81.

122. Du H, Bennett D, Li L, Whitlock G, Guo Y, Collins R, et al. Physical activity and sedentary leisure time and their associations with BMI, waist circumference, and percentage body fat in 0.5 million adults: the China Kadoorie Biobank study. Am J Clin Nutr. 2013;97:487–96.

123. Du H, Li L, Whitlock G, Bennett D, Guo Y, Bian Z, et al. Patterns and socio-demographic correlates of domain-specific physical activities and their associations with adiposity in the China Kadoorie Biobank study. BMC Public Health. 2014;14:826.

124. Survey data — China Kadoorie Biobank (CKB). https://www.ckbiobank.org/study-resources/survey-data. Accessed 5 Oct 2022.

125. Walters RG, Millwood IY, Lin K, Schmidt Valle D, McDonnell P, Hacker A, et al. Genotyping and population characteristics of the China Kadoorie Biobank. Cell Genomics. 2023;3:100361.

126. Yang G, Murray C, Zhang Z. Exploring adult mortality in China: levels, patterns and causes. 1991.

127. The BMJ. Statistics at Square One | 11. Correlation and regression. The BMJ: leading general medical journal. Research. Education. Comment. 2020. https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression. Accessed 20 Feb 2024.

128. Frost C, Thompson SG. Correcting for Regression Dilution Bias: Comparison of Methods for a Single Predictor Variable. Journal of the Royal Statistical Society Series A (Statistics in Society). 2000;163:173–89.

129. MacMahon S, Peto R, Collins R, Godwin J, MacMahon S, Cutler J, et al. Blood pressure, stroke, and coronary heart disease: Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. The Lancet. 1990;335:765–74.

130. Clarke R, Shipley M, Lewington S, Youngman L, Collins R, Marmot M, et al. Underestimation of Risk Associations Due to Regression Dilution in Long-term Follow-up of Prospective Studies. American Journal of Epidemiology. 1999;150:341–53.

131. Fang L, Gao P, Bao H, Tang X, Wang B, Feng Y, et al. Chronic obstructive pulmonary disease in China: a nationwide prevalence study. Lancet Respir Med. 2018;6:421–30.

132. Kirkwood BR, Sterne JAC, Kirkwood BR. Essential medical statistics. 2nd ed. Malden, Mass: Blackwell Science; 2003.

133. Kirkwood BR, Sterne JAC, Kirkwood BR. Survival analysis: displaying and comparing survival patterns. In: Essential medical statistics. 2nd ed. Malden, Mass: Blackwell Science; 2003.

134. Tseng C-H. The age- and sex-specific incidence and medical expenses of heart failure hospitalization in 2005 in Taiwan: a study using data from the National Health Insurance. J Am Geriatr Soc. 2010;58:611–3.

135. Zhang Y, Zhang J, Butler J, Yang X, Xie P, Guo D, et al. Contemporary Epidemiology, Management, and Outcomes of Patients Hospitalized for Heart Failure in China: Results From the China Heart Failure (China-HF) Registry. Journal of Cardiac Failure. 2017;23:868–75.

136. Conrad N, Judge A, Tran J, Mohseni H, Hedgecott D, Crespillo AP, et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. The Lancet. 2018;391:572–80.

137. Yu Z, Nissinen A, Vartiainen E, Song G, Guo Z, Zheng G, et al. Associations between socioeconomic status and cardiovascular risk factors in an urban population in China. Bull World Health Organ. 2000;78:1296–305.

138. Parenica J, Spinar J, Vitovec J, Widimsky P, Linhart A, Fedorco M, et al. Long-term survival following acute heart failure: The Acute Heart Failure Database Main registry (AHEAD Main). European Journal of Internal Medicine. 2013;24:151–60.

139. Lim YMF, Ong SM, Koudstaal S, Hwong WY, Liew HB, Rajadurai J, et al. Trends for Readmission and Mortality After Heart Failure Hospitalisation in Malaysia, 2007 to 2016. Glob Heart. 17:20.

140. Pandey A, Patel KV, Liang L, DeVore AD, Matsouaka R, Bhatt DL, et al. Association of Hospital Performance Based on 30-Day Risk-Standardized Mortality Rate With Long-term Survival After Heart Failure Hospitalization. JAMA Cardiol. 2018;3:489–97.

141. Lam CSP, Teng T-HK, Tay WT, Anand I, Zhang S, Shimizu W, et al. Regional and ethnic differences among patients with heart failure in Asia: the Asian sudden cardiac death in heart failure registry. European Heart Journal. 2016;37:3141–53.

142. Axson EL, Sundaram V, Bloom CI, Bottle A, Cowie MR, Quint JK. Temporal Trends in the Incidence of Heart Failure among Patients with Chronic Obstructive Pulmonary Disease and Its Association with Mortality. Annals ATS. 2020;17:939–48.

143. Chen Y, Wright N, Guo Y, Turnbull I, Kartsonaki C, Yang L, et al. Mortality and recurrent vascular events after first incident stroke: a 9-year community-based study of 0·5 million Chinese adults. The Lancet Global Health. 2020;8:e580–90.

144. Chun M, Qin H, Turnbull I, Sansome S, Gilbert S, Hacker A, et al. Heterogeneity in diagnosis and prognosis of ischaemic stroke subtypes: 9-year follow-up of 22000 cases in Chinese adults. International Journal of Stroke. 2023;:17474930231162265.

145. Axson EL, Ragutheeswaran K, Sundaram V, Bloom CI, Bottle A, Cowie MR, et al. Hospitalisation and mortality in patients with comorbid COPD and heart failure: a systematic review and meta-analysis. Respiratory Research. 2020;21:54.

146. Elston DM. Survivorship bias. Journal of the American Academy of Dermatology. 2021. https://doi.org/10.1016/j.jaad.2021.06.845.

147. Lipsitch M, Donnelly CA, Fraser C, Blake IM, Cori A, Dorigatti I, et al. Potential Biases in Estimating Absolute and Relative Case-Fatality Risks during Outbreaks. PLOS Neglected Tropical Diseases. 2015;9:e0003846.

148. Pasqualetti F, Barberis A, Zanotti S, Montemurro N, De Salvo GL, Soffietti R, et al. The impact of survivorship bias in glioblastoma research. Critical Reviews in Oncology/Hematology. 2023;188:104065.

149. Zhang M, Shi Y, Zhou B, Huang Z, Zhao Z, Li C, et al. Prevalence, awareness, treatment, and control of hypertension in China, 2004-18: findings from six rounds of a national survey. BMJ. 2023;380:e071952.

150. 8. The Chi squared tests | The BMJ. The BMJ | The BMJ: leading general medical journal. Research. Education. Comment. 2020. https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/8-chi-squared-tests. Accessed 7 May 2024.

151. Basto M, Abreu T, Gonçalves R, Pereira JM. Confounding and Effect Modification in Biostatistics: An Overview. Universal Journal of Applied Mathematics. 2023;11:8–15.

152. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. Stat Med. 2005;24:2911–35.

153. Choi J, Dekkers OM, Cessie S le. A comparison of different methods for handling measurements affected by medication use. 2022;:2022.04.23.22273899.

154. Crainiceanu CM, Ruppert D. Likelihood ratio tests for goodness-of-fit of a nonlinear regression model. Journal of Multivariate Analysis. 2004;91:35–52.

155. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. Stat Med. 1991;10:1025–35.

156. Easton D, Peto J. RE: "PRESENTING STATISTICAL UNCERTAINTY IN TRENDS AND DOSE-RESPONSE RELATIONS." American Journal of Epidemiology. 2000;152:393–6.

157. Plummer M. Improved estimates of floating absolute risk. Statistics in Medicine. 2004;23:93–104.

158. Shao J. Ordinary and weighted least-squares estimators. Canadian Journal of Statistics. 1990;18:327–36.

159. McNamee R. Regression modelling and other methods to control confounding. Occupational and Environmental Medicine. 2005;62:500–6.

160. GRAYSON DA. CONFOUNDING CONFOUNDING1. American Journal of Epidemiology. 1987;126:546–53.

161. Collett D. Modelling Survival Data in Medical Research. 3rd edition. New York: Chapman and Hall/CRC; 2015.

162. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York, NY: Springer; 2000.

163. R Core Team. R: A language and environment for statistical computing. 2022.

164. Viechtbauer W. Conducting Meta-Analyses in *R* with the **metafor** Package. J Stat Soft. 2010;36.

165. Wickham H. ggplot2. Cham: Springer International Publishing; 2016.

166. Arnold M. Jasper: Jasper makes plots. 2022.

167. Fang M, Wang D, Tang O, McEvoy JW, Echouffo-Tcheugui JB, Christenson RH, et al. Subclinical Cardiovascular Disease in US Adults With and Without Diabetes. Journal of the American Heart Association. 2023;12:e029083.

168. Bertoni AG, Kramer H, Watson K, Post WS. Diabetes and Clinical and Subclinical CVD. Global Heart. 2016;11:337–42.

169. Chillarón JJ, Roux JAF-L, Benaiges D, Pedro-Botet J. Subclinical cardiovascular disease in type 2 diabetes mellitus: To screen or not to screen. World J Clin Cases. 2014;2:415–21.

170. Oh GC, Cho H-J. Blood pressure and heart failure. Clinical Hypertension. 2020;26:1.

171. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51:1173–82.

172. Lin DY, Fleming TR, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. Stat Med. 1997;16:1515–27.

173. Pillay P, Lewington S, Taylor H, Lacey B, Carter J. Adiposity, Body Fat Distribution, and Risk of Major Stroke Types Among Adults in the United Kingdom. JAMA Network Open. 2022;5:e2246613.

174. Ndumele CE, Coresh J, Lazo M, Hoogeveen RC, Blumenthal RS, Folsom AR, et al. Obesity, Subclinical Myocardial Injury, and Incident Heart Failure. JACC: Heart Failure. 2014;2:600–7.

175. Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. Journal of the American Statistical Association. 1996;91:444–55.

176. Greenland S. An introduction to instrumental variables for epidemiologists. International Journal of Epidemiology. 2000;29:722–9.

177. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. Nat Rev Methods Primers. 2022;2:1–21.

178. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations: update for summer 2023. Wellcome Open Res. 2023;4:186.

179. Mackay TFC, Anholt RRH. Gregor Mendel's legacy in quantitative genetics. PLoS Biol. 2022;20:e3001692.

180. Monaghan F, Corcos A. On the origins of the Mendelian laws. Journal of Heredity. 1984;75:67–9.

181. Smith GD. Capitalizing on Mendelian randomization to assess the effects of treatments. J R Soc Med. 2007;100:432–5.

182. Stenseth NChr, Andersson L, Hoekstra HE. Gregor Johann Mendel and the development of modern evolutionary biology. Proc Natl Acad Sci U S A. 2022;119:e2201327119.

183. Sobczyk MK, Zheng J, Davey Smith G, Gaunt TR. Systematic comparison of Mendelian randomisation studies and randomised controlled trials using electronic databases. BMJ Open. 2023;13:e072087.

184. de Leeuw C, Savage J, Bucur IG, Heskes T, Posthuma D. Understanding the assumptions underlying Mendelian randomization. Eur J Hum Genet. 2022;30:653–60.

185. Davies NM, Holmes MV, Smith GD. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362.

186. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27:1133–63.

187. Bound J, Jaeger DA, Baker RM. Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak. Journal of the American Statistical Association. 1995;90:443–50.

188. Burgess S, Thompson SG, CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. International Journal of Epidemiology. 2011;40:755–64.

189. Wan EYF, Fung WT, Schooling CM, Au Yeung SL, Kwok MK, Yu EYT, et al. Blood Pressure and Risk of Cardiovascular Disease in UK Biobank. Hypertension. 2021;77:367–75.

190. Lian J, Shi X, Jia X, Fan J, Wang Y, Zhao Y, et al. Genetically predicted blood pressure, antihypertensive drugs and risk of heart failure: a Mendelian randomization study. Journal of Hypertension. 2023;41:44.

191. Chan II, Kwok MK, Schooling CM. The total and direct effects of systolic and diastolic blood pressure on cardiovascular disease and longevity using Mendelian randomisation. Sci Rep. 2021;11:21799.

192. Clarke R, Wright N, Walters R, Gan W, Guo Y, Millwood IY, et al. Genetically Predicted Differences in Systolic Blood Pressure and Risk of Cardiovascular and Noncardiovascular Diseases: A Mendelian Randomization Study in Chinese Adults. Hypertension. 2023;80:566–76.

193. Ciofani JL, Han D, Allahwala UK, Woolf B, Gill D, Bhindi R. Lipids, Blood Pressure, and Diabetes Mellitus on Risk of Cardiovascular Diseases in East Asians: A Mendelian Randomization Study. The American Journal of Cardiology. 2023;205:329–37.

194. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet. 2018;50:1412–25.

195. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nat Commun. 2020;11:163.

196. Takeuchi F, Akiyama M, Matoba N, Katsuya T, Nakatochi M, Tabara Y, et al. Interethnic analyses of blood pressure loci in populations of East Asian and European descent. Nat Commun. 2018;9:5052.

197. Shifman S, Kuypers J, Kokoris M, Yakir B, Darvasi A. Linkage disequilibrium patterns of the human genome across populations. Human Molecular Genetics. 2003;12:771–6.

198. Evans DM, Cardon LR. A Comparison of Linkage Disequilibrium Patterns and Estimated Population Recombination Rates across Multiple Populations. Am J Hum Genet. 2005;76:681–7.

199. R: Perform LD clumping on SNP data. https://search.r-project.org/CRAN/refmans/ieugwasr/html/ld_clump.html. Accessed 29 Apr 2024.

200. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. Nature. 2015;526:68–74.

201. Swerdlow DI, Kuchenbaecker KB, Shah S, Sofat R, Holmes MV, White J, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. Int J Epidemiol. 2016;45:1600–16.

202. Wu J, Kraja AT, Oberman A, Lewis CE, Ellison RC, Arnett DK, et al. A Summary of the Effects of Antihypertensive Medications on Measured Blood Pressure. American Journal of Hypertension. 2005;18:935–42.

203. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet. 2009;41:666–76.

204. Brion M-JA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. International Journal of Epidemiology. 2013;42:1497–501.

205. Sillanpää MJ. Overview of techniques to account for confounding due to population stratification and cryptic relatedness in genomic data association analyses. Heredity. 2011;106:511–9.

206. Abdellaoui A, Dolan CV, Verweij KJH, Nivard MG. Gene–environment correlations across geographic regions affect genome-wide association studies. Nat Genet. 2022;54:1345–54.

207. Burgess S, Swanson SA, Labrecque JA. Are Mendelian randomization investigations immune from bias due to reverse causation? Eur J Epidemiol. 2021;36:253–7.

208. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med. 2017;36:1783–802.

209. Burgess S, Butterworth A, Thompson SG. Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. Genet Epidemiol. 2013;37:658–65.

210. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res. 2017;26:2333–55.

211. Bowden J, Del Greco M F, Minelli C, Zhao Q, Lawlor DA, Sheehan NA, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. International Journal of Epidemiology. 2019;48:728–42.

212. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. International Journal of Epidemiology. 2013;42:1134–44.

213. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44:512–25.

214. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. The Annals of Statistics. 2020;48:1742–69.

215. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol. 2017;46:1985–98.

216. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40:304–14.

217. Xue H, Shen X, Pan W. Constrained maximum likelihood-based Mendelian randomization robust to both correlated and uncorrelated pleiotropic effects. Am J Hum Genet. 2021;108:1251–69.

218. Hägg S, Fall T, Ploner A, Mägi R, Fischer K, Draisma HH, et al. Adiposity as a cause of cardiovascular disease: a Mendelian randomization study. International Journal of Epidemiology. 2015;44:578–86.

219. Fall T, Hägg S, Mägi R, Ploner A, Fischer K, Horikoshi M, et al. The Role of Adiposity in Cardiometabolic Traits: A Mendelian Randomization Analysis. PLOS Medicine. 2013;10:e1001474.

220. Larsson SC, Burgess S. Causal role of high body mass index in multiple chronic diseases: a systematic review and meta-analysis of Mendelian randomization studies. BMC Med. 2021;19:320.

221. Hong Y, Gao Z, Wei H, Wei Y, Qiu Z, Xiao J, et al. Bi-directional association of body size and composition with heart failure: A Mendelian randomization study. International Journal of Cardiology. 2024;:132069.

222. Fairhurst-Hunter Z, Lin K, Millwood IY, Pozarickij A, Chen T-T, Torres JM, et al. Trans-ancestry meta-analysis improves performance of genetic scores for multiple adiposity-related traits in East Asian populations. 2022;:2022.07.05.22277254.

223. Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018;50:693–8.

224. Benn M, Marott SCW, Tybjærg-Hansen A, Nordestgaard BG. Obesity increases heart failure incidence and mortality: observational and Mendelian randomization studies totalling over 1 million individuals. Cardiovascular Research. 2022;118:3576–85.

225. Lind L, Ingelsson M, Sundstrom J, Ärnlöv J. Impact of risk factors for major cardiovascular diseases: a comparison of life-time observational and Mendelian randomisation findings. Open Heart. 2021;8:e001735.

226. Chen Y-H, Yin M-Q, Fan L-H, Jiang X-C, Xu H-F, Zhu X-Y, et al. Causal relationship between nutritional assessment phenotypes and heart failure: A Mendelian randomization study. Heliyon. 2024;10.

227. Li G, Zhang H, Jiang J. Genetic associations of childhood and adult BMI on chronic heart failure and ischemic stroke: A Mendelian randomization. IJC Heart & Vasculature. 2024;52:101425.

228. Ettehad D, Emdin CA, Kiran A, Anderson SG, Callender T, Emberson J, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. The Lancet. 2016;387:957–67.

229. Canoy D, Copland E, Nazarzadeh M, Ramakrishnan R, Pinho-Gomes A-C, Salam A, et al. Antihypertensive drug effects on long-term blood pressure: an individual-level data meta-analysis of randomised clinical trials. Heart. 2022;108:1281–9.

230. Yamaguchi NU, Bernardino EG, Ferreira MEC, de Lima BP, Pascotini MR, Yamaguchi MU. Sustainable development goals: a bibliometric analysis of literature reviews. Environ Sci Pollut Res Int. 2023;30:5502–15.

231. Ferreira JP, Kraus S, Mitchell S, Perel P, Piñeiro D, Chioncel O, et al. World Heart Federation Roadmap for Heart Failure. Global Heart. 2019;14:197–214.

232. Nisar A, Yin J, Yiping N, Lanting H, Zhang J, Wang D, et al. Making therapies culturally relevant: translation, cultural adaptation and field-testing of the Thinking Healthy Programme for perinatal depression in China. BMC Pregnancy and Childbirth. 2020;20:368.

233. Joo JY, Liu MF. Effectiveness of Culturally Tailored Interventions for Chronic Illnesses among Ethnic Minorities. West J Nurs Res. 2021;43:73–84.

234. Surendra K, Nürnberg S, Bremer JP, Knorr MS, Ückert F, Wenzel JP, et al. Pragmatic screening for heart failure in the general population using an electrocardiogram-based neural network. ESC Heart Failure. 2023;10:975–84.

235. Cosentino F, Grant PJ, Aboyans V, Bailey CJ, Ceriello A, Delgado V, et al. 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD: The Task Force for diabetes, pre-diabetes, and cardiovascular diseases of the European Society of Cardiology (ESC) and the European Association for the Study of Diabetes (EASD). European Heart Journal. 2020;41:255–323.

236. Du X, Patel A, Anderson CS, Dong J, Ma C. Epidemiology of Cardiovascular Disease in China and Opportunities for Improvement. Journal of the American College of Cardiology. 2019;73:3135–47.

237. Bei Y, Yang T, Xiao J. Cardiovascular medicine in China: what can we do to achieve the Healthy China 2030 plan? BMC Medicine. 2018;16:132.

238. Sun Y, Wang Z, Zhang J, Cai J, Shi F, Dong N. Current Status of and Opinions on Heart Transplantation in China. Current Medical Science. 2021;41.

239. Yin R, Yin L, Li L, Silva-Nash J, Tan J, Pan Z, et al. Hypertension in China: burdens, guidelines and policy responses: a state-of-the-art review. J Hum Hypertens. 2022;36:126–34.

240. Piepoli MF, Adamo M, Barison A, Bestetti RB, Biegus J, Böhm M, et al. Preventing heart failure: a position paper of the Heart Failure Association in collaboration with the European Association of Preventive Cardiology. European Journal of Preventive Cardiology. 2022;29:275–300.

241. A Randomized Trial of Intensive versus Standard Blood-Pressure Control. N Engl J Med. 2015;373:2103–16.

242. Qin X, Li Y, Sun N, He M, Tang G, Yin D, et al. Impact of Achieved Blood Pressure on First Stroke in Uncomplicated Grade 1 Hypertension. Journal of the American Heart Association. 2017;6:e005247.

243. Li X, Zhang J, Xing Z, Liu Q, Zhou S, Xiao Y. Intensive blood pressure control for patients aged over 60: A meta-analysis of the SPRINT, STEP, and ACCORD BP randomized controlled trials. Maturitas. 2023;172:52–9.

244. Bussa R, Nudy M, Ahmed M, Bussa J, Sturts A, Filippone E, et al. A systematic review and meta-analysis of the effect of intensive blood pressure control on nonfatal stroke. Journal of the American College of Cardiology. 2024;83 13_Supplement:1839–1839.

245. Wang L, Zhou B, Zhao Z, Yang L, Zhang M, Jiang Y, et al. Body-mass index and obesity in urban and rural China: findings from consecutive nationally representative surveys during 2004–18. The Lancet. 2021;398:53–63.

246. Jiang Y, Kong LZ, Li LM. [Implementing the strategy of “Healthy China” and strengthening the setting-up of National Demonstration Areas, for Comprehensive Prevention and Control of Non-communicable Diseases]. Zhonghua Liu Xing Bing Xue Za Zhi. 2018;39:391–3.

247. Zhang J, Astell-Burt T, Seo D-C, Feng X, Kong L, Zhao W, et al. Multilevel evaluation of “China Healthy Lifestyles for All”, a nationwide initiative to promote lower intakes of salt and edible oil. Prev Med. 2014;67:210–5.

248. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ. Selection bias due to loss to follow up in cohort studies. Epidemiology. 2016;27:91–7.

249. Lindsted KD, Fraser GE, Steinkohl M, Beeson WL. Healthy volunteer effect in a cohort study: temporal resolution in the Adventist Health Study. J Clin Epidemiol. 1996;49:783–90.

250. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332:1080.

251. Royston P, Sauerbrei W. Building multivariable regression models with continuous covariates in clinical epidemiology--with an emphasis on fractional polynomials. Methods Inf Med. 2005;44:561–71.

252. Royston P, Sauerbrei W. Multivariable Model - Building: A Pragmatic Approach to Regression Anaylsis based on Fractional Polynomials for Modelling Continuous Variables. John Wiley & Sons; 2008.

253. Silva I dos S, Cancer IA for R on. Cancer Epidemiology: Principles and Methods. IARC; 1999.

254. Wade KH, Carslake D, Sattar N, Davey Smith G, Timpson NJ. BMI and Mortality in UK Biobank: Revised Estimates Using Mendelian Randomization. Obesity. 2018;26:1796–806.

255. Tian H, Mason AM, Liu C, Burgess S. Relaxing parametric assumptions for non-linear Mendelian randomization using a doubly-ranked stratification method. PLOS Genetics. 2023;19:e1010823.

256. Hernán MA, Robins JM. Instruments for Causal Inference: An Epidemiologist’s Dream? Epidemiology. 2006;17:360–72.

257. Swanson SA, Hernán MA. Commentary: How to Report Instrumental Variable Analyses (Suggestions Welcome). Epidemiology. 2013;24:370–4.

258. Wade KH, Hamilton FW, Carslake D, Sattar N, Davey Smith G, Timpson NJ. Challenges in undertaking nonlinear Mendelian randomization. Obesity. 2023;31:2887–90.

259. Butler-Laporte G, Richards JB. Mendelian randomisation and vitamin D: the importance of model assumptions. The Lancet Diabetes & Endocrinology. 2023;11:14–5.

260. Hamilton FW, Hughes DA, Spiller W, Tilling K, Davey Smith G. Non-linear Mendelian randomization: detection of biases using negative controls with a focus on BMI, Vitamin D and LDL cholesterol. Eur J Epidemiol. 2024. https://doi.org/10.1007/s10654-024-01113-9.

261. Yang H, Negishi K, Otahal P, Marwick TH. Clinical prediction of incident heart failure risk: a systematic review and meta-analysis. Open Heart. 2015;2:e000222.

262. Li S, Ding H, Li Q, Zeng X, Zhang Y, Lai C, et al. Association between plasma proteome and pulmonary heart disease: A two-stage Mendelian randomization analysis. Clin Respir J. 2024;18:e13775.

263. Suizu F, Fukuta Y, Ueda K, Iwasaki T, Tokumitsu H, Hosoya H. Characterization of Ca2+/calmodulin-dependent protein kinase I as a myosin II regulatory light chain kinase in vitro and in vivo. Biochem J. 2002;367 Pt 2:335–45.

264. Sun BB, Suhre K, Gibson BW. Promises and Challenges of populational Proteomics in Health and Disease. Molecular & Cellular Proteomics. 2024;23.

Appendix

Appendix 1: Genetic variants used as genetic instruments for systolic blood pressure[194].

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
1	rs2076328	1	1687482	G	T	SBP	0.5108	0.334	0.0311	7.96E-27
2	rs2493292	1	3328659	T	C	SBP	0.1446	0.4078	0.0438	1.35E-20
3	rs2252865	1	8422676	T	C	SBP	0.3485	0.2022	0.0317	1.75E-10
4	rs3820068	1	15798197	A	G	SBP	0.8038	0.2935	0.0387	3.31E-14
5	rs6686889	1	25030470	T	C	SBP	0.2531	0.205	0.0348	3.89E-09
6	rs3737801	1	27960832	C	G	SBP	0.9224	0.2984	0.0503	3.03E-09
7	rs11210029	1	41865293	G	A	SBP	0.3678	0.1484	0.0254	5.16E-09
8	rs4926923	1	48109225	T	C	SBP	0.9114	0.3199	0.0537	2.56E-09
9	rs2404715	1	57008778	C	T	SBP	0.9075	0.305	0.0523	5.49E-09
10	rs60199046	1	59663341	A	G	SBP	0.711	0.2722	0.0332	2.41E-16
11	rs10923038	1	88651771	A	C	SBP	0.6179	0.1625	0.0257	2.41E-10
12	rs7514579	1	94051350	A	C	SBP	0.7712	0.1969	0.0293	1.73E-11
13	rs17396055	1	94730954	G	A	SBP	0.6683	0.1727	0.0264	5.73E-11
14	rs17030613	1	1.13E+08	C	A	SBP	0.2102	0.4053	0.0371	8.36E-28
15	rs11585169	1	1.51E+08	A	T	SBP	0.5773	0.1796	0.0308	5.34E-09
16	rs1043069	1	1.81E+08	T	G	SBP	0.6156	0.1995	0.0255	5.19E-15
17	rs4651224	1	1.85E+08	T	C	SBP	0.4474	0.1944	0.025	7.29E-15
18	rs12042924	1	1.97E+08	C	T	SBP	0.4716	0.1615	0.0247	6.26E-11
19	rs33996239	1	2.03E+08	C	T	SBP	0.9399	0.3538	0.0554	1.71E-10
20	rs2761436	1	2.08E+08	T	C	SBP	0.5368	0.1839	0.0302	1.15E-09
21	rs9431431	1	2.21E+08	G	A	SBP	0.706	0.2107	0.0331	1.97E-10
22	rs2004776	1	2.31E+08	T	C	SBP	0.2397	0.3789	0.0354	8.37E-27
23	rs4926499	1	2.49E+08	C	G	SBP	0.8263	0.2954	0.0374	2.91E-15
24	rs67720684	2	18975439	A	C	SBP	0.2385	0.2039	0.0291	2.26E-12

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
25	rs9678851	2	27887034	C	A	SBP	0.4325	0.1722	0.0307	1.99E-08
26	rs4952611	2	40567743	C	T	SBP	0.4205	0.211	0.0315	1.97E-11
27	rs35590893	2	43716933	G	A	SBP	0.7227	0.2152	0.0275	4.65E-15
28	rs2920899	2	55279681	T	G	SBP	0.7864	0.1851	0.0304	1.17E-09
29	rs925484	2	60611437	G	C	SBP	0.4002	0.1696	0.0308	3.66E-08
30	rs13014371	2	64217786	C	T	SBP	0.4295	0.1709	0.0305	1.99E-08
31	rs2300481	2	66782467	T	C	SBP	0.3865	0.1949	0.0251	8.01E-15
32	rs6731373	2	68503044	A	G	SBP	0.3492	0.1913	0.0326	4.18E-09
33	rs11689667	2	85491365	T	C	SBP	0.5447	0.1983	0.0304	6.66E-11
34	rs72847885	2	86326717	A	G	SBP	0.663	0.2183	0.026	4.79E-17
35	rs150194832	2	1.06E+08	G	C	SBP	0.9066	0.2901	0.0523	2.88E-08
36	rs62158170	2	1.14E+08	A	G	SBP	0.7834	0.2369	0.0368	1.21E-10
37	rs4954192	2	1.36E+08	C	T	SBP	0.609	0.1796	0.0312	8.62E-09
38	rs1438896	2	1.46E+08	T	C	SBP	0.2982	0.208	0.0329	2.48E-10
39	rs79523138	2	1.61E+08	G	A	SBP	0.1193	0.2629	0.0398	3.82E-11
40	rs55732192	2	1.62E+08	G	T	SBP	0.9053	0.2399	0.0424	1.50E-08
41	rs11694601	2	1.75E+08	G	A	SBP	0.4032	0.1545	0.026	2.89E-09
42	rs1837164	2	1.79E+08	A	T	SBP	0.3697	0.1497	0.0253	3.13E-09
43	rs16823124	2	1.83E+08	A	G	SBP	0.3066	0.236	0.0326	4.59E-13
44	rs296797	2	2.01E+08	T	C	SBP	0.405	0.1647	0.0249	3.48E-11
45	rs1469760	2	2.04E+08	C	T	SBP	0.4165	0.2485	0.0307	5.87E-16
46	rs55780018	2	2.09E+08	C	T	SBP	0.4566	0.2784	0.031	2.73E-19
47	rs1047891	2	2.12E+08	C	A	SBP	0.6807	0.1959	0.0263	1.03E-13
48	rs1063281	2	2.19E+08	C	T	SBP	0.3969	0.2585	0.0313	1.40E-16
49	rs1044822	2	2.31E+08	C	T	SBP	0.8518	0.2275	0.0348	6.15E-11
50	rs347591	3	11290122	T	G	SBP	0.6625	0.3181	0.032	2.87E-23
51	rs11128722	3	14958126	G	A	SBP	0.4304	0.2865	0.0309	1.93E-20
52	rs13082711	3	27537909	C	T	SBP	0.2383	0.3087	0.0354	2.78E-18
53	rs267517	3	37539090	G	A	SBP	0.4084	0.261	0.0313	7.18E-17

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
54	rs6797587	3	48197614	G	A	SBP	0.6721	0.3271	0.0322	3.40E-24
55	rs9810888	3	53635595	G	T	SBP	0.5016	0.172	0.0304	1.60E-08
56	rs4499560	3	70920485	T	A	SBP	0.6829	0.2009	0.0265	3.69E-14
57	rs9857362	3	74710462	A	C	SBP	0.5291	0.151	0.0247	9.71E-10
58	rs1375564	3	85656311	T	C	SBP	0.6395	0.2343	0.0256	5.49E-20
59	rs1882289	3	1.14E+08	G	A	SBP	0.1154	0.2159	0.0379	1.26E-08
60	rs6438857	3	1.25E+08	T	C	SBP	0.5774	0.1153	0.0175	4.44E-11
61	rs6783086	3	1.34E+08	T	C	SBP	0.3954	0.3048	0.0308	5.13E-23
62	rs2306374	3	1.38E+08	C	T	SBP	0.162	0.2825	0.0411	6.11E-12
63	rs73158427	3	1.54E+08	A	T	SBP	0.1618	0.2804	0.0409	7.24E-12
64	rs143112823	3	1.55E+08	G	A	SBP	0.9124	0.4171	0.0557	7.16E-14
65	rs419076	3	1.69E+08	T	C	SBP	0.4732	0.4049	0.0301	3.00E-41
66	rs262986	3	1.83E+08	G	A	SBP	0.5296	0.1856	0.0248	7.79E-14
67	rs12374077	3	1.85E+08	C	G	SBP	0.3443	0.2143	0.0317	1.48E-11
68	rs55829085	4	2165493	C	A	SBP	0.0454	0.4309	0.074	5.73E-09
69	rs2610990	4	18008232	G	A	SBP	0.7359	0.2573	0.028	3.82E-20
70	rs2291435	4	38387395	C	T	SBP	0.4665	0.262	0.0303	5.31E-18
71	rs12511987	4	46595623	G	T	SBP	0.1774	0.2153	0.0323	2.71E-11
72	rs10008637	4	77414144	T	C	SBP	0.5405	0.1879	0.0244	1.40E-14
73	rs6823199	4	83925895	T	C	SBP	0.7438	0.2094	0.0348	1.72E-09
74	rs1347345	4	95938386	G	A	SBP	0.3821	0.1792	0.0253	1.45E-12
75	rs13112725	4	1.07E+08	C	G	SBP	0.7631	0.4137	0.0358	6.81E-31
76	rs6825911	4	1.11E+08	C	T	SBP	0.2085	0.3066	0.0375	3.05E-16
77	rs7439567	4	1.38E+08	T	C	SBP	0.4106	0.2355	0.0251	7.35E-21
78	rs10305838	4	1.48E+08	C	T	SBP	0.1408	0.2659	0.0433	7.94E-10
79	rs13139571	4	1.57E+08	C	A	SBP	0.7634	0.3299	0.0355	1.33E-20
80	rs17035181	4	1.58E+08	T	G	SBP	0.8552	0.2949	0.0348	2.17E-17
81	rs869396	4	1.7E+08	C	A	SBP	0.5341	0.2115	0.0305	4.12E-12
82	rs4957026	5	361148	A	G	SBP	0.3399	0.1446	0.026	2.57E-08

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
83	rs10069690	5	1279790	T	C	SBP	0.2582	0.2735	0.029	3.78E-21
84	rs1173771	5	32815028	G	A	SBP	0.6013	0.6321	0.0307	5.92E-94
85	rs1694068	5	53283630	A	T	SBP	0.6139	0.2339	0.0251	1.36E-20
86	rs3121685	5	65662133	C	T	SBP	0.5162	0.1475	0.0246	1.92E-09
87	rs246973	5	68007803	T	C	SBP	0.2882	0.2244	0.0278	7.25E-16
88	rs10057188	5	77837789	G	A	SBP	0.5416	0.2275	0.0307	1.19E-13
89	rs709668	5	96174186	G	A	SBP	0.7997	0.2534	0.0306	1.24E-16
90	rs10077885	5	1.14E+08	C	A	SBP	0.4916	0.259	0.0308	4.43E-17
91	rs9885577	5	1.21E+08	T	C	SBP	0.3709	0.2309	0.0321	6.53E-13
92	rs6595838	5	1.28E+08	A	G	SBP	0.2985	0.3229	0.0331	1.54E-22
93	rs702395	5	1.4E+08	T	C	SBP	0.4369	0.191	0.0246	8.25E-15
94	rs1650911	5	1.42E+08	C	G	SBP	0.7647	0.1787	0.0293	1.09E-09
95	rs9687065	5	1.48E+08	A	G	SBP	0.8099	0.3356	0.0387	4.62E-18
96	rs11953630	5	1.58E+08	C	T	SBP	0.6341	0.4529	0.032	1.55E-45
97	rs72812846	5	1.73E+08	T	A	SBP	0.722	0.2579	0.0344	6.91E-14
98	rs2745599	6	1613686	A	G	SBP	0.552	0.1908	0.0253	4.17E-14
99	rs9392172	6	7723962	G	C	SBP	0.4643	0.1969	0.0302	7.31E-11
100	rs1630736	6	12295987	C	T	SBP	0.535	0.1706	0.0309	3.52E-08
101	rs9368222	6	20686996	A	C	SBP	0.2688	0.2301	0.0273	3.85E-17
102	rs409558	6	31708147	T	C	SBP	0.8487	0.3844	0.0431	4.66E-19
103	rs115245297	6	34244132	C	T	SBP	0.0445	0.4453	0.0769	7.06E-09
104	rs78648104	6	50683009	C	T	SBP	0.0925	0.4287	0.0541	2.37E-15
105	rs10943605	6	79655477	A	G	SBP	0.4887	0.239	0.0302	2.72E-15
106	rs35410524	6	96885405	T	C	SBP	0.1885	0.3368	0.0387	3.19E-18
107	rs9486916	6	1.09E+08	T	C	SBP	0.1979	0.2657	0.0385	5.42E-12
108	rs2693560	6	1.18E+08	G	A	SBP	0.634	0.206	0.0316	6.99E-11
109	rs9372498	6	1.19E+08	A	T	SBP	0.0812	0.3967	0.0556	9.43E-13
110	rs7763294	6	1.4E+08	G	T	SBP	0.6837	0.1804	0.0263	6.85E-12
111	rs17080102	6	1.51E+08	G	C	SBP	0.9306	0.8085	0.0594	3.52E-42

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
112	rs449789	6	1.6E+08	C	G	SBP	0.139	0.3233	0.0438	1.50E-13
113	rs9456648	6	1.62E+08	C	T	SBP	0.6773	0.1771	0.0322	3.76E-08
114	rs73030266	6	1.66E+08	A	T	SBP	0.9331	0.4393	0.062	1.42E-12
115	rs6959688	7	1966831	G	A	SBP	0.4019	0.2002	0.0253	2.40E-15
116	rs73049928	7	4669949	G	A	SBP	0.1939	0.2382	0.0392	1.20E-09
117	rs2107595	7	19049388	A	G	SBP	0.159	0.4176	0.0415	7.37E-24
118	rs2069833	7	22767664	C	T	SBP	0.4254	0.1988	0.0306	8.34E-11
119	rs12979	7	24738164	C	G	SBP	0.8698	0.2025	0.0364	2.60E-08
120	rs7777128	7	27337113	C	G	SBP	0.0798	0.5158	0.0557	2.06E-20
121	rs10233127	7	30933453	A	T	SBP	0.1076	0.3025	0.0409	1.48E-13
122	rs76206723	7	40447971	G	A	SBP	0.893	0.4037	0.0491	2.02E-16
123	rs11977526	7	46008110	G	A	SBP	0.5991	0.3213	0.0312	6.62E-25
124	rs6593297	7	56122058	A	T	SBP	0.3029	0.1581	0.0274	8.38E-09
125	rs6963105	7	75097488	G	A	SBP	0.5612	0.1779	0.0263	1.39E-11
126	rs848445	7	77572461	C	T	SBP	0.7149	0.2004	0.0279	6.57E-13
127	rs10245696	7	90449362	A	C	SBP	0.4003	0.1695	0.0308	3.59E-08
128	rs1947228	7	96461649	C	T	SBP	0.582	0.1756	0.0308	1.19E-08
129	rs4728142	7	1.29E+08	G	A	SBP	0.5561	0.1814	0.0305	2.59E-09
130	rs13238550	7	1.31E+08	A	G	SBP	0.3964	0.2572	0.0309	7.80E-17
131	rs7810028	7	1.39E+08	C	G	SBP	0.8068	0.2835	0.0382	1.24E-13
132	rs12703989	7	1.4E+08	A	G	SBP	0.4991	0.1521	0.025	1.26E-09
133	rs73727605	7	1.49E+08	A	G	SBP	0.0663	0.3616	0.0623	6.60E-09
134	rs11771693	7	1.5E+08	A	G	SBP	0.6755	0.1673	0.0265	2.72E-10
135	rs10224002	7	1.51E+08	G	A	SBP	0.2859	0.3672	0.0337	1.31E-27
136	rs4875958	8	1721090	A	G	SBP	0.7119	0.1762	0.0271	7.77E-11
137	rs1986971	8	10268736	A	G	SBP	0.7064	0.2377	0.0274	3.93E-18
138	rs62503324	8	23400615	T	C	SBP	0.2393	0.2638	0.0356	1.26E-13
139	rs6557876	8	25900675	C	T	SBP	0.7497	0.4156	0.0349	1.11E-32
140	rs1906672	8	38130025	A	G	SBP	0.2319	0.2715	0.0291	1.03E-20

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
141	rs4873492	8	51947549	T	C	SBP	0.1724	0.3012	0.0332	1.09E-19
142	rs6996733	8	60535824	T	C	SBP	0.8488	0.1882	0.0341	3.34E-08
143	rs2354862	8	64501744	A	C	SBP	0.6407	0.1921	0.0257	7.18E-14
144	rs1449544	8	76591880	A	C	SBP	0.5436	0.2117	0.0303	2.57E-12
145	rs72688070	8	81393697	C	T	SBP	0.8314	0.241	0.0326	1.56E-13
146	rs56345595	8	82814156	A	G	SBP	0.5853	0.2066	0.0308	1.91E-11
147	rs4582532	8	95969257	G	A	SBP	0.4941	0.2003	0.0303	3.64E-11
148	rs2978098	8	1.02E+08	A	C	SBP	0.5468	0.2227	0.0307	3.78E-13
149	rs35783704	8	1.06E+08	G	A	SBP	0.8958	0.4619	0.0507	8.81E-20
150	rs28499085	8	1.1E+08	A	G	SBP	0.7253	0.2096	0.0339	6.42E-10
151	rs2071518	8	1.2E+08	T	C	SBP	0.2611	0.2757	0.0343	8.76E-16
152	rs894344	8	1.36E+08	G	A	SBP	0.4074	0.2136	0.0306	3.11E-12
153	rs34591516	8	1.42E+08	T	C	SBP	0.0494	0.4862	0.071	7.65E-12
154	rs4129585	8	1.43E+08	A	C	SBP	0.4413	0.1749	0.0246	1.06E-12
155	rs28558845	9	4334791	G	C	SBP	0.8447	0.218	0.0339	1.22E-10
156	rs1332813	9	9350706	T	C	SBP	0.3514	0.1965	0.0259	3.58E-14
157	rs4553000	9	34223553	C	T	SBP	0.4859	0.2035	0.03	1.09E-11
158	rs76452347	9	35906471	C	T	SBP	0.795	0.2974	0.0397	7.13E-14
159	rs7045409	9	95201540	T	A	SBP	0.6331	0.1823	0.0255	9.15E-13
160	rs13290326	9	1.17E+08	C	T	SBP	0.4986	0.2122	0.03	1.52E-12
161	rs10818775	9	1.26E+08	C	T	SBP	0.8787	0.302	0.0462	6.23E-11
162	rs72765298	9	1.28E+08	C	T	SBP	0.1249	0.3675	0.0462	1.75E-15
163	rs7023828	9	1.28E+08	C	T	SBP	0.582	0.2436	0.0249	1.22E-22
164	rs184457	9	1.32E+08	G	A	SBP	0.692	0.1659	0.0272	1.09E-09
165	rs11145807	9	1.4E+08	A	G	SBP	0.4057	0.2135	0.0322	3.54E-11
166	rs56352451	10	5804865	T	C	SBP	0.1324	0.1992	0.0358	2.54E-08
167	rs4373814	10	18419972	C	G	SBP	0.4277	0.3266	0.0306	1.20E-26
168	rs1813353	10	18707448	T	C	SBP	0.6625	0.5198	0.0319	8.15E-60
169	rs10826995	10	32082658	C	T	SBP	0.2839	0.1936	0.0335	7.55E-09

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
170	rs10761530	10	62390726	T	C	SBP	0.4989	0.2004	0.03	2.54E-11
171	rs1530440	10	63524591	C	T	SBP	0.8125	0.5669	0.0386	7.36E-49
172	rs75270364	10	74751579	C	T	SBP	0.0623	0.3646	0.0517	1.83E-12
173	rs10887914	10	82215288	T	C	SBP	0.4625	0.2007	0.0302	3.18E-11
174	rs77413490	10	89681688	T	G	SBP	0.0424	0.473	0.0626	4.11E-14
175	rs11187142	10	94468685	T	C	SBP	0.1047	0.2378	0.0396	1.89E-09
176	rs11191548	10	1.05E+08	T	C	SBP	0.9188	1.0983	0.0553	1.16E-87
177	rs34872471	10	1.15E+08	C	T	SBP	0.2924	0.2201	0.0333	3.83E-11
178	rs2782980	10	1.16E+08	C	T	SBP	0.7159	0.4249	0.0338	3.58E-36
179	rs11592107	10	1.23E+08	A	G	SBP	0.3096	0.2306	0.0264	2.77E-18
180	rs7096563	10	1.34E+08	A	G	SBP	0.3548	0.1943	0.0259	6.76E-14
181	rs1133400	10	1.34E+08	G	A	SBP	0.214	0.2607	0.0303	7.04E-18
182	rs661348	11	1905292	C	T	SBP	0.4253	0.4451	0.0316	5.23E-45
183	rs10743086	11	8774923	G	A	SBP	0.7928	0.1832	0.0303	1.57E-09
184	rs360153	11	9762274	C	T	SBP	0.5834	0.3445	0.0306	1.73E-29
185	rs10766533	11	19224677	A	T	SBP	0.7108	0.1665	0.0272	8.82E-10
186	rs11030119	11	27728102	G	A	SBP	0.6987	0.2218	0.033	1.79E-11
187	rs11031051	11	30355707	C	A	SBP	0.3099	0.1899	0.0264	5.80E-13
188	rs919045	11	31111810	T	C	SBP	0.6281	0.1704	0.0312	4.64E-08
189	rs7103648	11	47461783	G	A	SBP	0.3871	0.3862	0.031	1.27E-35
190	rs2688716	11	54835623	G	T	SBP	0.7106	0.2331	0.0281	1.17E-16
191	rs751984	11	61278246	T	C	SBP	0.8824	0.4408	0.0479	3.60E-20
192	rs3741378	11	65408937	C	T	SBP	0.8648	0.4087	0.0446	4.83E-20
193	rs67976715	11	68023742	C	G	SBP	0.2342	0.1971	0.0295	2.30E-11
194	rs504217	11	72006086	T	C	SBP	0.0738	0.3855	0.0583	3.90E-11
195	rs7927515	11	76125330	A	C	SBP	0.3459	0.2271	0.0319	1.05E-12
196	rs2289125	11	89224453	C	A	SBP	0.7795	0.2592	0.0375	4.90E-12
197	rs633185	11	1.01E+08	C	G	SBP	0.7127	0.6373	0.0334	5.60E-81
198	rs12362593	11	1.12E+08	G	C	SBP	0.271	0.2004	0.0347	7.91E-09

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
199	rs17119370	11	1.16E+08	A	T	SBP	0.6926	0.2207	0.033	2.16E-11
200	rs1076485	11	1.17E+08	T	C	SBP	0.1304	0.3126	0.0365	1.06E-17
201	rs78998485	12	434755	G	C	SBP	0.2557	0.2036	0.0279	3.18E-13
202	rs7132012	12	8832203	A	G	SBP	0.6748	0.2	0.0321	4.76E-10
203	rs2024385	12	12888438	T	A	SBP	0.576	0.2326	0.0256	1.05E-19
204	rs12579720	12	20173764	G	C	SBP	0.7566	0.258	0.0352	2.46E-13
205	rs7976167	12	24210599	T	C	SBP	0.6889	0.1638	0.0266	7.21E-10
206	rs11168245	12	48204499	C	G	SBP	0.7655	0.2784	0.0357	5.83E-15
207	rs2261608	12	48721634	A	T	SBP	0.3511	0.1723	0.0313	3.65E-08
208	rs73099903	12	53440779	T	C	SBP	0.0824	0.4694	0.0557	3.43E-17
209	rs4143175	12	67782397	T	C	SBP	0.2409	0.1802	0.0286	3.14E-10
210	rs7963801	12	79685226	C	T	SBP	0.5779	0.2027	0.0256	2.29E-15
211	rs17249754	12	90060586	G	A	SBP	0.8317	0.8446	0.0403	1.25E-97
212	rs10778174	12	1.03E+08	G	A	SBP	0.7521	0.2236	0.0282	2.15E-15
213	rs11112548	12	1.06E+08	A	T	SBP	0.9556	0.4608	0.0623	1.44E-13
214	rs12184466	12	1.11E+08	T	C	SBP	0.1991	0.2992	0.0434	5.30E-12
215	rs10850411	12	1.15E+08	T	C	SBP	0.6969	0.2908	0.0326	4.58E-19
216	rs35444	12	1.16E+08	A	G	SBP	0.6138	0.4368	0.031	3.47E-45
217	rs117206641	12	1.33E+08	T	C	SBP	0.1108	0.2157	0.0357	1.48E-09
218	rs606950	13	22298923	A	G	SBP	0.6163	0.2324	0.0253	3.77E-20
219	rs9508495	13	30146201	C	T	SBP	0.2435	0.3557	0.0353	6.34E-24
220	rs4274337	13	41967193	G	A	SBP	0.8303	0.22	0.0328	1.96E-11
221	rs73187288	13	42738672	C	A	SBP	0.1091	0.2558	0.04	1.57E-10
222	rs912434	13	47189928	T	G	SBP	0.7594	0.2151	0.0287	6.27E-14
223	rs9526707	13	51489186	G	A	SBP	0.6784	0.1939	0.0264	1.90E-13
224	rs7988232	13	79808655	A	G	SBP	0.4113	0.148	0.0249	2.99E-09
225	rs9549328	13	1.14E+08	T	C	SBP	0.2305	0.2915	0.0362	7.91E-16
226	rs452036	14	23865885	G	A	SBP	0.6454	0.2218	0.0315	2.02E-12
227	rs17115145	14	30122409	T	C	SBP	0.3967	0.1569	0.025	3.74E-10

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
228	rs8904	14	35871217	A	G	SBP	0.3678	0.3061	0.0314	1.71E-22
229	rs9888615	14	53377540	C	T	SBP	0.709	0.274	0.0332	1.46E-16
230	rs57786342	14	69260028	A	G	SBP	0.2059	0.2317	0.0374	5.63E-10
231	rs11623535	14	72462381	A	G	SBP	0.7439	0.2022	0.0277	3.02E-13
232	rs8013933	14	94465789	T	C	SBP	0.6913	0.2027	0.0327	5.48E-10
233	rs1475130	14	1E+08	C	T	SBP	0.654	0.1977	0.0318	4.83E-10
234	rs8014182	14	1.04E+08	C	T	SBP	0.8676	0.3397	0.0361	4.99E-21
235	rs11629850	15	40317075	A	G	SBP	0.5287	0.1995	0.0244	2.86E-16
236	rs2925345	15	41311799	T	C	SBP	0.4678	0.274	0.0301	9.14E-20
237	rs3098186	15	50810621	C	T	SBP	0.4844	0.2422	0.0303	1.41E-15
238	rs1378942	15	75077367	C	A	SBP	0.3324	0.5076	0.032	9.37E-57
239	rs3743157	15	85680532	A	C	SBP	0.1669	0.2602	0.0325	1.27E-15
240	rs28611491	15	90641809	T	C	SBP	0.0814	0.3343	0.0576	6.54E-09
241	rs34756251	15	1E+08	C	T	SBP	0.829	0.2497	0.0325	1.42E-14
242	rs28590346	16	2080653	T	A	SBP	0.3403	0.2709	0.0332	3.37E-16
243	rs12921187	16	4943019	G	T	SBP	0.5717	0.2925	0.0303	5.19E-22
244	rs3915425	16	15912544	T	C	SBP	0.6819	0.206	0.0323	1.86E-10
245	rs11639856	16	24788645	T	A	SBP	0.8066	0.3197	0.0379	3.53E-17
246	rs10468291	16	49768046	C	A	SBP	0.4312	0.1868	0.0305	9.38E-10
247	rs34941092	16	50550137	G	A	SBP	0.8502	0.2303	0.0341	1.52E-11
248	rs28633979	16	65282820	C	A	SBP	0.5697	0.1794	0.0302	2.97E-09
249	rs35261357	16	75444572	T	C	SBP	0.585	0.3502	0.0307	4.43E-30
250	rs7500448	16	83045790	A	G	SBP	0.7467	0.2227	0.035	1.98E-10
251	rs6540125	16	87993889	T	G	SBP	0.3411	0.164	0.0263	4.54E-10
252	rs12941318	17	1333598	C	T	SBP	0.5045	0.2351	0.0309	2.74E-14
253	rs4480845	17	1958609	T	C	SBP	0.3652	0.1954	0.0183	1.15E-26
254	rs4925159	17	18185510	A	G	SBP	0.4246	0.2281	0.0247	2.27E-20
255	rs1551355	17	30032420	T	C	SBP	0.2334	0.1932	0.0288	1.96E-11
256	rs9904409	17	42680402	A	G	SBP	0.0974	0.3682	0.0508	4.20E-13

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
257	rs12946454	17	43208121	T	A	SBP	0.2652	0.4125	0.0341	1.01E-33
258	rs17608766	17	45013271	C	T	SBP	0.1445	0.6903	0.0433	2.48E-57
259	rs12940887	17	47402807	T	C	SBP	0.3671	0.2597	0.0312	8.03E-17
260	rs1036902	17	58950791	C	T	SBP	0.1531	0.2273	0.0338	1.85E-11
261	rs4308	17	61559625	A	G	SBP	0.3777	0.2731	0.0314	3.11E-18
262	rs2467099	17	73949045	C	T	SBP	0.7768	0.2462	0.036	8.26E-12
263	rs9302885	17	76799898	A	G	SBP	0.4452	0.1945	0.0245	2.28E-15
264	rs112280096	17	79367409	C	A	SBP	0.6376	0.1706	0.0264	1.13E-10
265	rs34413141	18	777282	T	A	SBP	0.8178	0.3084	0.0325	2.10E-21
266	rs1154214	18	24546824	G	T	SBP	0.6037	0.1476	0.0251	3.87E-09
267	rs12958173	18	42141977	A	C	SBP	0.297	0.295	0.0329	2.97E-19
268	rs7236548	18	43097750	A	C	SBP	0.1848	0.3431	0.0388	8.51E-19
269	rs745821	18	48142854	T	G	SBP	0.7548	0.2214	0.0351	2.80E-10
270	rs34163044	18	51851616	A	C	SBP	0.4199	0.1753	0.031	1.59E-08
271	rs72930904	18	52607301	C	T	SBP	0.8403	0.2316	0.041	1.57E-08
272	rs10048404	18	54578482	C	T	SBP	0.6299	0.2277	0.027	3.10E-17
273	rs10460108	18	73034151	A	G	SBP	0.4801	0.2118	0.0294	5.67E-13
274	rs3760994	19	1435771	G	A	SBP	0.5085	0.1787	0.032	2.29E-08
275	rs740406	19	2232221	G	A	SBP	0.0599	0.5905	0.0659	3.25E-19
276	rs7248104	19	7224431	G	A	SBP	0.5876	0.2326	0.0306	2.74E-14
277	rs2304130	19	19789528	G	A	SBP	0.0843	0.3387	0.0552	8.22E-10
278	rs8105753	19	31927547	A	C	SBP	0.6315	0.2178	0.0318	6.84E-12
279	rs1821295	19	32590773	C	T	SBP	0.3013	0.2253	0.0328	6.75E-12
280	rs7412	19	45412079	C	T	SBP	0.9181	0.4339	0.0574	4.04E-14
281	rs73046792	19	49605705	G	A	SBP	0.8412	0.282	0.0359	3.90E-15
282	rs1764975	20	4101290	A	T	SBP	0.7954	0.2162	0.0321	1.56E-11
283	rs6108168	20	8626271	C	A	SBP	0.7458	0.2957	0.0344	8.71E-18
284	rs1327235	20	10969030	G	A	SBP	0.4705	0.426	0.03	7.88E-46
285	rs6141767	20	31225069	C	G	SBP	0.157	0.3107	0.0419	1.17E-13

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect allele Frequency	Beta	Standard error	P-value
286	rs6031435	20	42797358	G	A	SBP	0.4602	0.2592	0.0303	1.09E-17
287	rs6095241	20	47308798	G	A	SBP	0.5605	0.1718	0.0302	1.26E-08
288	rs35213536	20	62694319	T	G	SBP	0.2466	0.2937	0.0357	1.94E-16
289	rs1882961	21	16556367	T	C	SBP	0.3087	0.2346	0.0267	1.61E-18
290	rs11701033	21	33788341	G	C	SBP	0.1822	0.2187	0.0392	2.52E-08
291	rs12627651	21	44760603	A	G	SBP	0.2872	0.3498	0.0341	1.02E-24
292	rs9306160	21	45107562	C	T	SBP	0.5893	0.2123	0.0309	6.35E-12
293	rs12628032	22	19967980	T	C	SBP	0.3094	0.2445	0.0328	9.81E-14
294	rs9608690	22	28921347	G	A	SBP	0.9318	0.2983	0.0502	2.90E-09
295	rs737721	22	30172254	G	C	SBP	0.0522	0.4074	0.0684	2.62E-09
296	rs28578714	22	50727921	T	C	SBP	0.6062	0.1951	0.0258	4.17E-14

Appendix 2: Genetic variants used as genetic instruments for body mass index[222].

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
1	rs115866895	1	1592638	A	G	BMI	0.0196	0.00232	6.85E-17
2	rs779765342	1	23353906	C	CT	BMI	0.0186	0.00273	1.53E-11
3	rs945211	1	32191798	C	G	BMI	0.0132	0.00209	4.09E-10
4	rs4652839	1	33784929	C	G	BMI	-0.0136	0.00216	4.46E-10
5	rs112566467	1	39562627	C	T	BMI	-0.0183	0.0025	3.47E-13
6	rs12144626	1	47670525	C	T	BMI	-0.0164	0.00206	3.07E-15
7	rs1167311	1	49996959	A	G	BMI	-0.0205	0.00219	1.68E-20
8	rs630602	1	54728864	C	G	BMI	0.0124	0.00209	4.12E-09
9	rs74806710	1	57833008	A	T	BMI	0.0193	0.00334	1.08E-08
10	rs7519259	1	66434743	A	G	BMI	0.0151	0.00204	2.13E-13
11	rs2613498	1	72752939	C	T	BMI	0.0321	0.00257	3.16E-36
12	rs2568955	1	72762169	C	T	BMI	0.0133	0.00246	8.51E-08
13	1:75000417:C:CT	1	75000417	C	CT	BMI	-0.017	0.00205	1.97E-16
14	rs6696236	1	96291369	A	G	BMI	0.0154	0.00218	8.89E-13
15	rs995258	1	97431052	A	C	BMI	0.0142	0.00204	6.59E-12
16	1:98427843:C:CT	1	98427843	C	CT	BMI	0.0236	0.0031	5.70E-14
17	rs1335055	1	1.08E+08	A	G	BMI	-0.0134	0.00215	7.94E-10
18	rs76210342	1	1.08E+08	A	AT	BMI	0.0136	0.00228	3.36E-09
19	rs197374	1	1.12E+08	C	T	BMI	-0.0147	0.00209	3.31E-12
20	rs6577585	1	6715440	A	G	BMI	0.0123	0.00214	1.21E-08
21	rs2902811	1	1.51E+08	C	G	BMI	-0.0176	0.0024	3.92E-13
22	rs1778830	1	1.56E+08	A	G	BMI	0.0139	0.00212	8.72E-11
23	rs34720381	1	1.71E+08	C	T	BMI	-0.0234	0.00351	4.19E-11
24	rs544871667	1	1.74E+08	C	CT	BMI	0.0329	0.00468	3.48E-12
25	rs539515	1	1.78E+08	A	C	BMI	-0.0478	0.00251	9.73E-82
26	rs9787132	1	1.85E+08	A	T	BMI	-0.0118	0.00203	1.04E-08
27	rs815163	1	1.9E+08	C	T	BMI	-0.0155	0.00204	5.63E-14
28	1:195029618:C:CT	1	1.95E+08	C	CT	BMI	0.0161	0.0023	4.39E-12
29	1:197133836:A:AATATATTAT	1	1.97E+08	A	AATATATTAT	BMI	-0.015	0.00234	2.40E-10
30	rs2678204	1	2.02E+08	G	T	BMI	0.0241	0.00214	5.40E-29
31	rs4342901	1	2.02E+08	A	G	BMI	-0.0131	0.00215	1.56E-09
32	rs4971239	1	2.03E+08	A	G	BMI	0.0176	0.00273	1.65E-10
33	rs1507336	1	2.1E+08	C	G	BMI	-0.0195	0.00302	1.60E-10
34	rs6695101	1	2.1E+08	A	G	BMI	0.0172	0.00275	5.73E-10
35	rs33956517	1	8546815	T	TAA	BMI	-0.0113	0.00208	3.26E-08
36	1:11254006:A:AC	1	11254006	A	AC	BMI	-0.0157	0.0025	5.10E-10
37	rs10737913	1	14004623	C	T	BMI	-0.0119	0.00226	3.48E-08
38	rs17713879	2	254215	A	G	BMI	-0.0104	0.0021	9.42E-07
39	rs7561278	2	48954905	C	T	BMI	-0.0179	0.00247	7.63E-13
40	rs376151889	2	52654979	A	G	BMI	-0.012	0.00209	1.12E-08
41	rs7601895	2	55281901	C	G	BMI	0.0169	0.0022	2.97E-14
42	rs2609181	2	6161899	C	G	BMI	-0.0158	0.00227	4.44E-12
43	rs4671328	2	58935282	G	T	BMI	-0.0213	0.00205	1.01E-24
44	rs13404126	2	62848632	A	G	BMI	0.0126	0.00206	9.63E-10

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
45	rs10520327	2	81840224	A	G	BMI	-0.0126	0.00214	6.62E-09
46	2:86742320:T:TA	2	86742320	T	TA	BMI	-0.017	0.00215	5.19E-15
47	rs11692215	2	1.01E+08	C	T	BMI	0.0179	0.00237	1.81E-15
48	rs2384054	2	25156773	C	T	BMI	0.0331	0.00203	2.80E-59
49	rs3087523	2	1.06E+08	A	G	BMI	0.019	0.00308	1.01E-09
50	rs111501168	2	1.14E+08	C	T	BMI	0.0141	0.00282	4.03E-09
51	rs34847111	2	1.34E+08	C	T	BMI	-0.0137	0.00235	7.01E-09
52	rs79503448	2	1.43E+08	A	G	BMI	0.0159	0.0028	1.73E-08
53	rs453520	2	1.48E+08	C	T	BMI	0.0181	0.00205	2.43E-18
54	2:157597637:G:GA	2	1.58E+08	G	GA	BMI	0.0126	0.00209	2.59E-09
55	rs80330591	2	1.59E+08	A	G	BMI	-0.0164	0.00286	1.28E-08
56	rs2384463	2	26942256	A	G	BMI	0.0161	0.00203	7.44E-16
57	rs10803762	2	1.61E+08	A	G	BMI	0.0136	0.00218	6.93E-10
58	rs141961555	2	1.65E+08	A	AGC	BMI	-0.0173	0.00296	3.82E-09
59	rs12692738	2	1.66E+08	C	T	BMI	0.0141	0.00244	6.95E-09
60	rs34234296	2	1.75E+08	A	G	BMI	-0.0158	0.0021	9.72E-14
61	rs9630986	2	1.82E+08	C	G	BMI	0.0169	0.00215	7.81E-15
62	rs1064213	2	1.99E+08	A	G	BMI	0.0141	0.00203	5.28E-12
63	rs2351774	2	2.04E+08	A	G	BMI	-0.013	0.00205	2.53E-10
64	rs4482463	2	2.05E+08	A	C	BMI	-0.0346	0.00384	3.83E-19
65	rs11889536	2	2.2E+08	A	G	BMI	0.0197	0.00283	5.34E-12
66	rs6543913	2	35437986	C	G	BMI	0.0142	0.00204	7.12E-12
67	rs4625852	2	2.37E+08	A	G	BMI	-0.0156	0.00251	7.75E-10
68	rs7568228	2	2.37E+08	C	G	BMI	-0.0125	0.00203	1.25E-09
69	rs11677279	2	37021090	C	G	BMI	0.0125	0.00213	6.18E-09
70	rs7589518	2	37995422	A	T	BMI	0.0126	0.00214	6.15E-09
71	rs13062931	3	38637160	C	T	BMI	-0.0112	0.00207	3.63E-08
72	3:41418131:T:TA	3	41418131	T	TA	BMI	0.0149	0.00251	3.77E-09
73	rs5848679	3	44044401	G	GTT	BMI	0.0127	0.00204	7.78E-10
74	rs72906474	3	47817007	G	T	BMI	0.0174	0.00208	1.46E-16
75	3:49959570:C:CA	3	49959570	C	CA	BMI	0.0292	0.00206	2.66E-45
76	3:50785304:A:AG	3	50785304	A	AG	BMI	0.0173	0.00298	8.88E-09
77	rs2253795	3	53778875	G	T	BMI	0.0149	0.0024	8.68E-10
78	rs1564907	3	56193500	C	T	BMI	-0.0182	0.00328	3.84E-08
79	rs815715	3	61264084	C	G	BMI	0.0175	0.00206	4.71E-17
80	rs557951	3	62713263	G	T	BMI	0.0138	0.0022	4.40E-10
81	rs571211727	3	77665524	G	GT	BMI	0.0127	0.00205	9.71E-10
82	rs6419734	3	78458928	C	T	BMI	-0.0164	0.00291	2.12E-08
83	rs11915747	3	85699040	C	G	BMI	0.019	0.00212	7.09E-19
84	rs770811933	3	85905981	G	GC	BMI	0.0202	0.00243	1.47E-16
85	rs56398417	3	88024986	C	T	BMI	0.0154	0.00219	4.42E-13
86	rs1454687	3	94038085	C	G	BMI	0.0226	0.00203	2.51E-28
87	rs17402168	3	1.01E+08	C	G	BMI	-0.0183	0.00322	2.01E-08
88	rs1436348	3	1.05E+08	A	G	BMI	-0.0152	0.00206	2.37E-13
89	rs1561073	3	11642114	A	T	BMI	-0.0131	0.00232	2.39E-08
90	rs2016469	3	1.08E+08	A	G	BMI	0.0131	0.00212	9.18E-10

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
91	rs553616034	3	1.18E+08	G	GAAAAAAAAAA	BMI	-0.012	0.00219	4.86E-08
92	rs12330631	3	1.23E+08	C	T	BMI	0.0129	0.00211	1.48E-09
93	rs1225004	3	1.32E+08	C	T	BMI	0.0239	0.00227	1.32E-25
94	rs10935143	3	1.35E+08	A	G	BMI	-0.0119	0.00205	2.78E-09
95	rs2035936	3	1.41E+08	G	T	BMI	-0.0365	0.00449	8.42E-16
96	rs12634936	3	1.48E+08	C	T	BMI	0.0262	0.00464	2.34E-08
97	rs355748	3	1.54E+08	G	T	BMI	-0.0165	0.0021	6.88E-15
98	rs9834519	3	1.56E+08	C	T	BMI	0.0236	0.00376	3.23E-10
99	rs62644679	3	1.58E+08	A	G	BMI	-0.0158	0.00238	5.62E-11
100	rs67481408	3	1.71E+08	T	TA	BMI	-0.0182	0.00229	4.30E-15
101	rs13099930	3	1.71E+08	A	G	BMI	0.0124	0.00216	9.78E-09
102	rs529200	3	1.73E+08	A	G	BMI	-0.0158	0.00204	1.37E-14
103	3:183513243:A:ATTTTTTTTTTTTT	3	1.84E+08	A	ATTTTTTTTTTTTT	BMI	-0.0122	0.00208	4.96E-09
104	rs11546878	3	1.84E+08	C	T	BMI	0.0171	0.00268	2.30E-10
105	rs2569993	3	12926096	C	T	BMI	0.0131	0.00218	1.50E-09
106	rs869400	3	1.86E+08	G	T	BMI	0.0296	0.00262	2.87E-29
107	rs6583310	3	1.96E+08	C	G	BMI	0.0141	0.00205	1.05E-11
108	3:15807354:C:CAA	3	15807354	C	CAA	BMI	-0.0133	0.00214	6.52E-10
109	rs1609783	3	25095911	A	G	BMI	0.014	0.00205	4.66E-12
110	rs75384631	3	35141900	A	T	BMI	0.0198	0.0032	1.05E-09
111	3:35719107:T:TA	3	35719107	T	TA	BMI	-0.0159	0.00222	1.48E-12
112	4:3142660:C:CACTT	4	3142660	C	CACTT	BMI	0.0281	0.0039	1.07E-12
113	rs35488113	4	53289490	T	TTA	BMI	0.0117	0.00204	1.30E-08
114	rs1871028	4	54576884	A	C	BMI	-0.0115	0.00205	2.64E-08
115	rs28537413	4	65785466	A	T	BMI	0.0137	0.00224	1.53E-09
116	rs11724750	4	80808820	C	T	BMI	0.0133	0.00207	1.75E-10
117	4:91258772:C:CT	4	91258772	C	CT	BMI	0.0119	0.00211	2.39E-08
118	rs66679256	4	18351898	C	T	BMI	-0.0167	0.00205	6.18E-16
119	4:96157764:A:AG	4	96157764	A	AG	BMI	0.0128	0.00213	3.02E-09
120	rs1229984	4	1E+08	C	T	BMI	0.0436	0.00681	1.58E-10
121	rs326894	4	1.13E+08	C	T	BMI	-0.013	0.00208	4.65E-10
122	rs4398538	4	1.31E+08	C	T	BMI	-0.0146	0.00213	1.19E-11
123	rs57800857	4	1.41E+08	A	C	BMI	0.0156	0.00212	3.21E-13
124	rs35390852	4	1.43E+08	A	G	BMI	0.0171	0.00309	2.62E-08
125	rs750090	4	1.53E+08	C	T	BMI	-0.0128	0.00215	9.17E-10
126	rs77978620	4	1.63E+08	C	T	BMI	-0.0152	0.00256	4.07E-09
127	rs7683836	4	1.8E+08	A	G	BMI	-0.0135	0.00205	8.72E-11
128	rs34811474	4	25408838	A	G	BMI	-0.03	0.00241	4.45E-35
129	4:28526950:A:AAAAC	4	28526950	A	AAAAC	BMI	-0.0207	0.00275	7.51E-14
130	rs4527444	4	30842780	A	G	BMI	-0.0145	0.00204	2.07E-12
131	rs6852808	4	34926958	C	T	BMI	0.0124	0.00223	3.31E-08
132	rs3209570	4	38699657	A	G	BMI	-0.0134	0.0021	2.18E-10
133	rs6867471	5	3574564	C	T	BMI	0.0133	0.00212	4.33E-10
134	rs323759	5	86759713	A	T	BMI	-0.0434	0.00658	6.94E-11
135	rs1477290	5	87988934	C	T	BMI	0.0355	0.00299	2.55E-32
136	rs159029	5	94215894	C	T	BMI	0.0135	0.00236	1.27E-08

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
137	rs6235	5	95728898	C	G	BMI	-0.0172	0.0023	1.14E-13
138	5:103942055:A:AATTTAT	5	1.04E+08	A	AATTTAT	BMI	-0.0122	0.00204	3.47E-09
139	rs199968817	5	27180374	G	GT	BMI	-0.0138	0.00226	1.27E-09
140	rs35727810	5	1.06E+08	A	G	BMI	0.0136	0.00235	8.54E-09
141	rs10623997	5	1.07E+08	T	TATAATA	BMI	0.0245	0.00246	4.69E-23
142	rs765791981	5	1.12E+08	T	TATACACAC	BMI	0.0146	0.00239	1.56E-09
143	rs17669543	5	1.14E+08	A	G	BMI	0.0161	0.00243	5.64E-11
144	rs11416364	5	1.19E+08	A	AC	BMI	0.0143	0.00205	4.60E-12
145	rs34732995	5	1.23E+08	C	CTA	BMI	0.0156	0.00204	3.31E-14
146	5:124336163:G:GA	5	1.24E+08	G	GA	BMI	0.0117	0.00207	2.48E-08
147	rs7341051	5	1.3E+08	A	G	BMI	0.0154	0.00273	2.38E-08
148	rs796338714	5	43190861	T	TA	BMI	-0.0156	0.00214	5.06E-13
149	rs2074613	5	1.4E+08	C	T	BMI	-0.0118	0.00205	8.74E-09
150	rs10063055	5	1.41E+08	C	T	BMI	-0.0135	0.00233	1.08E-08
151	rs2579040	5	1.53E+08	A	T	BMI	-0.0142	0.00227	6.28E-10
152	rs4958702	5	1.54E+08	C	T	BMI	-0.0162	0.00206	5.35E-15
153	rs245769	5	1.71E+08	C	T	BMI	-0.0188	0.00229	3.74E-16
154	rs6874700	5	50701750	A	T	BMI	0.0129	0.00211	1.37E-09
155	rs71606623	5	59364244	G	GTTAT	BMI	-0.0127	0.00224	2.01E-08
156	rs55908499	5	63020950	G	GA	BMI	-0.0155	0.00208	1.80E-13
157	rs10050620	5	63927239	C	T	BMI	0.014	0.00217	1.69E-10
158	rs2307111	5	75003678	C	T	BMI	-0.0289	0.00208	3.03E-43
159	rs9378999	6	5979069	C	T	BMI	0.0124	0.00216	1.26E-08
160	rs12203240	6	24749104	C	T	BMI	-0.0225	0.00375	2.92E-09
161	rs2066295	6	26168903	A	G	BMI	0.017	0.00242	3.73E-12
162	6:31324372:G:GA	6	31324372	G	GA	BMI	-0.0185	0.00248	1.46E-13
163	rs113344808	6	34701981	C	CATTT	BMI	-0.0289	0.00222	2.85E-38
164	rs2820232	6	35003603	G	T	BMI	-0.00807	0.00206	0.000107
165	rs17757975	6	38214150	C	T	BMI	-0.017	0.00286	4.10E-09
166	rs9471333	6	40362023	C	T	BMI	0.0233	0.00205	9.12E-30
167	rs72892910	6	50816887	G	T	BMI	-0.0407	0.0027	9.43E-51
168	rs200769207	6	64174931	C	CA	BMI	-0.0123	0.00213	1.21E-08
169	rs7381960	6	70195719	A	C	BMI	-0.0127	0.00224	2.04E-08
170	rs9344002	6	81148579	A	G	BMI	-0.0175	0.00319	1.64E-08
171	rs9294260	6	83433228	A	G	BMI	0.0135	0.00205	7.17E-11
172	rs2506942	6	93629042	A	G	BMI	0.0119	0.00205	1.03E-08
173	6:97806419:G:GGA	6	97806419	G	GGA	BMI	-0.0158	0.00215	3.75E-13
174	rs6938973	6	98421721	C	T	BMI	0.0181	0.00208	4.25E-18
175	rs10947793	6	12142817	A	G	BMI	0.0139	0.00212	9.38E-11
176	rs2253310	6	1.09E+08	C	G	BMI	-0.0182	0.00211	1.19E-17
177	rs13218383	6	1.2E+08	C	G	BMI	0.0139	0.00216	1.95E-10
178	rs2749929	6	1.32E+08	C	T	BMI	0.0146	0.0025	8.44E-09
179	6:143186572:G:GA	6	1.43E+08	G	GA	BMI	-0.0133	0.00204	1.34E-10
180	rs7749708	6	1.53E+08	C	T	BMI	-0.0145	0.00224	1.65E-10
181	rs9478496	6	1.54E+08	C	T	BMI	0.0195	0.00275	2.15E-12
182	rs9371992	6	1.57E+08	A	G	BMI	-0.0132	0.0023	1.28E-08

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
183	rs13210756	6	1.63E+08	A	T	BMI	-0.0222	0.0031	1.28E-12
184	rs6909668	6	19724989	C	T	BMI	-0.0118	0.00208	1.94E-08
185	7:1272278:C:CCCCT	7	1272278	C	CCCCT	BMI	0.0155	0.00251	9.93E-10
186	rs799449	7	44784697	C	T	BMI	-0.0159	0.00206	1.01E-14
187	rs10269783	7	49616203	A	G	BMI	0.0136	0.00209	3.99E-11
188	rs10486875	7	69819103	A	T	BMI	0.0153	0.00238	2.03E-10
189	rs10950207	7	70105168	C	T	BMI	-0.0131	0.0021	5.33E-10
190	rs236660	7	75050086	C	T	BMI	0.0246	0.00214	2.74E-30
191	rs6953561	7	76637391	A	G	BMI	-0.0275	0.00275	3.71E-24
192	rs3840590	7	77827064	G	GA	BMI	0.0162	0.00215	8.54E-14
193	rs4721089	7	1872921	C	T	BMI	-0.019	0.00248	3.17E-14
194	rs3901286	7	99107727	A	C	BMI	-0.0232	0.00284	7.20E-16
195	rs1524445	7	1.13E+08	C	T	BMI	0.0191	0.00208	8.70E-20
196	rs1840660	7	1.14E+08	A	G	BMI	0.0157	0.0021	8.11E-14
197	rs11406302	7	1.3E+08	T	TC	BMI	-0.0112	0.0022	3.79E-07
198	rs11525873	7	1.39E+08	C	T	BMI	-0.023	0.00342	3.13E-11
199	rs4722398	7	3125220	C	T	BMI	-0.0188	0.00296	3.74E-10
200	rs1805123	7	1.51E+08	G	T	BMI	-0.0165	0.00236	5.35E-12
201	rs4307239	7	24354300	A	G	BMI	-0.0115	0.00205	2.62E-08
202	rs215634	7	32369148	A	G	BMI	0.0159	0.0021	3.39E-14
203	7:39403859:C:CT	7	39403859	C	CT	BMI	-0.0173	0.0025	7.46E-12
204	rs62488401	8	2933476	A	G	BMI	-0.013	0.00233	2.92E-08
205	rs7009996	8	28143259	A	C	BMI	0.0115	0.00213	3.96E-08
206	rs375300236	8	33291129	C	CA	BMI	-0.0121	0.0022	4.85E-08
207	rs2468964	8	38333062	G	T	BMI	-0.0126	0.00208	2.16E-09
208	rs583948	8	60883426	A	G	BMI	0.0125	0.00212	5.16E-09
209	rs1915792	8	64703342	C	T	BMI	0.0133	0.00204	2.40E-11
210	8:67209546:C:CGT	8	67209546	C	CGT	BMI	-0.0157	0.00234	3.03E-11
211	rs35957544	8	73440371	G	T	BMI	0.02	0.00207	6.39E-22
212	8:74678436:G:GGGAGGAGAGGAGAAGAAAGAGAAAA	8	74678436	G	GGGAGGAGAGGAGAAGAAAGAGAAAA	BMI	0.02	0.00323	8.54E-10
213	rs2596121	8	76660225	A	G	BMI	-0.0183	0.00208	3.61E-18
214	rs35853762	8	77235251	T	TG	BMI	-0.0208	0.00206	1.52E-23
215	8:85642746:C:CAT	8	85642746	C	CAT	BMI	-0.0159	0.00241	6.73E-11
216	8:87469638:A:ATCTG	8	87469638	A	ATCTG	BMI	0.0137	0.00209	8.03E-11
217	rs1504797	8	89434405	C	T	BMI	-0.0138	0.00222	8.44E-10
218	rs4876611	8	1.17E+08	A	G	BMI	-0.0195	0.00227	1.26E-17
219	rs754601466	8	1.19E+08	T	TA	BMI	0.0232	0.00334	4.79E-12
220	8:126504383:C:CCACCAT	8	1.27E+08	C	CCACCAT	BMI	0.0142	0.00224	3.18E-10
221	rs11782074	8	1.43E+08	G	T	BMI	-0.0151	0.00213	2.26E-12
222	rs13275517	8	1.43E+08	C	T	BMI	0.0127	0.00206	1.09E-09
223	rs6530737	8	14095763	A	G	BMI	0.0152	0.00213	1.72E-12
224	rs146671801	8	15577699	G	GTA	BMI	-0.0132	0.00213	1.02E-09
225	rs59104534	8	25666169	C	T	BMI	-0.0123	0.00223	4.55E-08
226	rs2570967	9	8851463	G	T	BMI	-0.0126	0.00216	7.86E-09
227	rs13289199	9	37278064	A	T	BMI	0.0137	0.00214	1.26E-10
228	rs7038966	9	73777777	C	T	BMI	-0.013	0.00208	2.70E-10

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
229	rs10869977	9	80457276	A	G	BMI	-0.012	0.00206	8.75E-09
230	9:10125423:T:TTTATTATTATTATTATTATTA	9	10125423	T	TTTATTATTATTATTATTATTA	BMI	0.0148	0.00266	3.38E-08
231	rs3949781	9	92178472	A	T	BMI	0.0146	0.00205	1.61E-12
232	rs34696009	9	96395221	G	GCTGT	BMI	0.0166	0.0022	8.18E-14
233	9:103113656:A:AC	9	1.03E+08	A	AC	BMI	0.0178	0.00214	1.50E-16
234	rs2417998	9	1.12E+08	C	G	BMI	0.0128	0.00225	1.74E-08
235	rs149607570	9	1.2E+08	C	CGT	BMI	0.0136	0.00215	3.35E-10
236	rs6478538	9	1.25E+08	A	G	BMI	0.0125	0.00218	1.41E-08
237	rs56141370	9	1.27E+08	C	T	BMI	-0.0193	0.00317	1.79E-09
238	rs6597653	9	1.34E+08	C	G	BMI	0.013	0.00209	6.65E-10
239	rs10756713	9	15880555	A	G	BMI	0.021	0.00206	6.26E-24
240	rs10962547	9	16712247	A	T	BMI	0.0231	0.00272	3.60E-17
241	rs1934102	9	23228275	A	G	BMI	-0.0133	0.00217	1.13E-09
242	rs7874154	9	27777012	C	T	BMI	0.0145	0.00204	1.44E-12
243	rs35475612	10	16750948	C	CTT	BMI	-0.0144	0.0022	1.02E-10
244	rs10824218	10	76421216	A	T	BMI	0.0126	0.00207	1.23E-09
245	rs11001963	10	78760959	C	T	BMI	-0.0116	0.00208	3.20E-08
246	rs17399739	10	87490850	A	G	BMI	-0.0258	0.00402	1.86E-10
247	rs11189058	10	99001258	A	G	BMI	-0.0127	0.00215	5.10E-09
248	rs2439823	10	99778226	A	G	BMI	-0.0201	0.00205	2.34E-22
249	rs4075353	10	1.02E+08	A	G	BMI	-0.0107	0.00215	8.38E-07
250	rs10764256	10	20905269	C	T	BMI	0.0146	0.00243	3.00E-09
251	rs144741702	10	1.04E+08	T	TACC	BMI	0.0171	0.00282	1.77E-09
252	rs4575195	10	1.15E+08	A	C	BMI	-0.0165	0.0022	9.77E-14
253	rs10510025	10	1.19E+08	C	T	BMI	-0.0169	0.00237	1.89E-12
254	rs845084	10	1.25E+08	A	G	BMI	0.017	0.00234	6.78E-13
255	rs72828935	10	1.27E+08	C	G	BMI	0.0165	0.00227	6.62E-13
256	rs2542615	10	1.31E+08	C	T	BMI	0.0128	0.00217	4.46E-09
257	rs11012732	10	21830104	A	G	BMI	-0.0212	0.00217	3.50E-22
258	rs34479808	10	27309906	T	TA	BMI	-0.017	0.00302	2.51E-08
259	rs36021393	10	33971717	C	T	BMI	-0.0262	0.00382	1.11E-11
260	rs4595495	10	53673286	A	G	BMI	-0.0132	0.00207	1.07E-10
261	rs12761779	10	63782043	C	G	BMI	0.011	0.00217	4.57E-08
262	rs7924036	10	65191645	G	T	BMI	0.0133	0.00204	9.92E-11
263	rs7103389	11	881639	C	T	BMI	0.0146	0.00211	9.60E-12
264	rs7942037	11	30430332	C	G	BMI	-0.0183	0.00212	1.24E-17
265	rs10838117	11	43538772	A	T	BMI	-0.0128	0.00204	5.50E-10
266	rs77762353	11	43648786	A	AT	BMI	0.0238	0.00221	8.43E-27
267	rs35072015	11	45426928	C	T	BMI	-0.0144	0.00221	1.07E-10
268	rs7124681	11	47529947	A	C	BMI	0.0265	0.00207	3.07E-37
269	rs34292685	11	64049021	C	T	BMI	0.0193	0.00276	2.56E-12
270	rs10657468	11	69442884	C	CCTT	BMI	0.015	0.00207	6.77E-13
271	rs112326082	11	76479391	A	AGCCTTCCT	BMI	0.0161	0.00259	7.22E-10
272	11:84790528:A:ATT	11	84790528	A	ATT	BMI	-0.0159	0.00208	4.62E-14
273	rs61903695	11	89922417	A	G	BMI	-0.0165	0.00234	2.83E-12
274	rs11607476	11	1.15E+08	A	C	BMI	-0.0155	0.00205	7.67E-14

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
275	rs35483388	11	1.23E+08	C	T	BMI	-0.0128	0.00211	1.63E-09
276	rs7940866	11	1.31E+08	A	T	BMI	0.017	0.00205	1.70E-16
277	rs11223204	11	1.33E+08	A	G	BMI	-0.0141	0.00206	1.12E-11
278	rs12364470	11	1.35E+08	G	T	BMI	0.0194	0.00274	2.43E-12
279	rs2054117	11	11796897	G	T	BMI	0.0117	0.00205	1.75E-08
280	rs72867447	11	13301875	C	G	BMI	-0.0172	0.00206	1.15E-16
281	rs1002226	11	17405617	C	T	BMI	-0.0124	0.0021	5.30E-09
282	rs6265	11	27679916	C	T	BMI	0.0409	0.0026	8.78E-56
283	rs491711	11	28742220	A	C	BMI	0.013	0.00221	5.89E-09
284	rs55726687	12	991306	A	G	BMI	0.0255	0.00249	3.21E-24
285	rs1458156	12	41887940	C	T	BMI	-0.0148	0.00204	7.16E-13
286	rs4077093	12	51593616	G	T	BMI	-0.0138	0.00249	3.90E-08
287	rs35969387	12	53680955	C	T	BMI	-0.0189	0.00334	2.01E-08
288	rs2271189	12	56494991	A	G	BMI	-0.0177	0.00208	3.56E-17
289	rs4267103	12	60966740	C	T	BMI	0.0146	0.00263	3.76E-08
290	rs1819844	12	68205604	A	G	BMI	0.0155	0.00271	1.68E-08
291	rs10774018	12	2157925	C	G	BMI	0.0175	0.00247	2.54E-12
292	rs704061	12	89771903	C	T	BMI	0.0157	0.00205	3.14E-14
293	rs591776	12	99538895	C	T	BMI	0.0153	0.00212	9.86E-13
294	rs17608150	12	1.1E+08	C	T	BMI	-0.0215	0.00375	1.56E-08
295	rs76406609	12	1.14E+08	C	T	BMI	0.0255	0.00406	5.18E-10
296	rs752843328	12	1.25E+08	G	GAA	BMI	0.0154	0.00249	9.47E-10
297	rs12422552	12	14413931	C	G	BMI	-0.014	0.00231	1.93E-09
298	12:16458403:A:AGTTT	12	16458403	A	AGTTT	BMI	0.0179	0.00291	1.20E-09
299	rs10842240	12	24060075	C	G	BMI	0.023	0.0032	6.77E-13
300	rs9579775	13	20616557	A	C	BMI	-0.0209	0.00309	2.46E-11
301	rs755841629	13	62713381	T	TGTGACTAATTTTCTTCTATAAG	BMI	0.0122	0.00227	4.77E-08
302	rs5804250	13	66204127	C	CT	BMI	0.0124	0.00207	3.69E-09
303	rs1441264	13	79580919	A	G	BMI	0.0197	0.00212	3.05E-20
304	rs116394958	13	86477072	C	T	BMI	-0.0159	0.00229	5.72E-12
305	rs55911231	13	96983940	C	T	BMI	-0.0169	0.00207	5.94E-16
306	rs9584870	13	99245866	C	T	BMI	-0.0132	0.00215	1.43E-09
307	13:109857398:T:TA	13	1.1E+08	T	TA	BMI	-0.012	0.00205	6.53E-09
308	rs9512696	13	28012527	A	G	BMI	-0.018	0.00216	1.67E-16
309	rs9522183	13	1.12E+08	G	T	BMI	0.0151	0.00207	5.32E-13
310	rs12872889	13	28674628	C	T	BMI	0.013	0.00245	1.20E-08
311	rs75983170	13	40788838	G	T	BMI	0.0152	0.00221	4.03E-12
312	rs9536410	13	53971828	C	T	BMI	0.0107	0.00211	4.03E-07
313	rs6561766	13	54691484	A	G	BMI	0.0203	0.0032	3.69E-10
314	rs9537309	13	56533807	C	T	BMI	-0.0136	0.00238	1.67E-08
315	rs4055791	13	59266053	C	T	BMI	0.018	0.00207	6.12E-18
316	rs142720498	13	59418167	C	CAA	BMI	-0.0179	0.00242	2.49E-13
317	14:25931404:G:GTC	14	25931404	G	GTC	BMI	0.0225	0.00219	3.53E-25
318	rs217669	14	62360075	C	T	BMI	0.0173	0.00229	5.99E-14
319	rs56069292	14	62620696	A	G	BMI	-0.0114	0.00205	2.14E-08
320	rs61986330	14	73314450	A	C	BMI	-0.0151	0.00228	6.11E-11

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
321	rs8008772	14	88321884	A	T	BMI	-0.0153	0.00236	6.24E-11
322	rs575958301	14	89211507	G	GTGTTTGTT	BMI	-0.0199	0.00355	2.76E-08
323	14:91499131:G:GA	14	91499131	G	GA	BMI	0.0158	0.00213	1.94E-13
324	rs9788550	14	29681138	C	G	BMI	-0.02	0.00237	5.99E-17
325	rs6575340	14	94023972	A	G	BMI	0.0211	0.00212	7.86E-23
326	rs761594282	14	1.02E+08	C	CG	BMI	0.0126	0.00246	3.36E-07
327	rs7159203	14	1.02E+08	C	T	BMI	0.0246	0.00389	4.16E-10
328	rs3212038	14	1.04E+08	A	G	BMI	-0.0142	0.00217	1.34E-12
329	14:33298731:C:CA	14	33298731	C	CA	BMI	0.0183	0.00207	1.80E-18
330	rs534930508	14	40104783	A	AGGAGT	BMI	-0.0179	0.00316	2.03E-08
331	rs67272968	14	40845125	A	G	BMI	0.0161	0.00261	8.39E-10
332	rs11161335	15	27004095	A	T	BMI	0.0124	0.00206	2.74E-09
333	rs147384209	15	55454511	C	T	BMI	0.0419	0.00677	8.67E-10
334	15:57099411:G:GA	15	57099411	G	GA	BMI	-0.0143	0.00252	2.10E-08
335	rs1369159	15	66360842	C	T	BMI	0.0118	0.00208	1.52E-08
336	rs2241420	15	68082816	A	G	BMI	-0.0302	0.00244	1.06E-34
337	rs7164727	15	73093991	C	T	BMI	-0.0169	0.00217	1.33E-14
338	rs2870111	15	79403585	C	T	BMI	0.0163	0.00208	7.95E-15
339	rs34769775	15	80989172	C	T	BMI	0.0155	0.00223	6.41E-12
340	rs7498044	15	92573639	A	G	BMI	-0.0157	0.00249	4.85E-10
341	rs8025516	15	95271872	G	T	BMI	-0.0161	0.00214	9.37E-14
342	rs72767957	15	99240793	A	G	BMI	0.0165	0.00255	1.28E-10
343	15:42081490:G:GA	15	42081490	G	GA	BMI	-0.0163	0.00265	1.28E-09
344	rs12440603	15	46585722	C	T	BMI	-0.0145	0.00206	3.26E-12
345	rs563857745	15	52226908	C	CAAAAAA	BMI	0.0132	0.00205	6.64E-11
346	rs80243702	15	53163603	A	G	BMI	0.016	0.0028	1.83E-08
347	rs7201895	16	407723	A	G	BMI	-0.0152	0.00214	2.02E-12
348	rs146315450	16	15323336	C	CAA	BMI	0.0117	0.00228	4.25E-08
349	rs868554	16	20050466	C	G	BMI	-0.0186	0.0024	1.59E-14
350	rs4780885	16	20380004	C	G	BMI	-0.0164	0.00204	1.78E-15
351	rs34898535	16	31025641	C	T	BMI	0.0206	0.00211	1.22E-24
352	rs76488452	16	53756885	A	G	BMI	-0.0171	0.0046	0.000249
353	rs2516739	16	2097158	A	G	BMI	-0.0164	0.00247	4.23E-11
354	16:56389827:C:CA	16	56389827	C	CA	BMI	0.0155	0.00275	2.56E-08
355	16:62867929:T:TAG	16	62867929	T	TAG	BMI	-0.0127	0.00211	2.66E-09
356	rs4451971	16	64300792	A	G	BMI	0.0112	0.00209	3.55E-08
357	rs2307022	16	68381978	A	G	BMI	0.0123	0.00216	2.00E-08
358	rs10694251	16	69720803	G	GTAAA	BMI	-0.0238	0.00208	6.86E-30
359	16:70479748:C:CT	16	70479748	C	CT	BMI	-0.0222	0.00277	1.98E-15
360	rs929866	16	72038363	C	T	BMI	-0.0166	0.00218	4.25E-14
361	rs11403987	16	73605132	A	AT	BMI	0.0115	0.00205	3.11E-08
362	rs11150461	16	82448195	C	G	BMI	0.0159	0.0023	8.77E-12
363	rs8047251	16	3708567	A	G	BMI	0.0139	0.0022	4.30E-10
364	rs7206608	16	82872628	C	G	BMI	-0.0138	0.00218	4.29E-10
365	rs10083803	16	6701400	C	T	BMI	0.0137	0.00229	3.35E-09
366	rs4516268	17	1846831	A	C	BMI	-0.0246	0.00258	3.34E-21

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
367	rs59237168	17	31479035	C	T	BMI	-0.0185	0.00249	1.59E-13
368	rs349792	17	39261560	A	T	BMI	-0.0149	0.00222	4.78E-12
369	17:42285460:A:ATTTTTTTTTT	17	42285460	A	ATTTTTTTTTT	BMI	0.0144	0.00229	4.79E-10
370	rs8064802	17	44262496	C	T	BMI	-0.0136	0.0022	9.91E-10
371	rs62072006	17	52938468	A	C	BMI	-0.0172	0.0029	4.54E-09
372	rs62058023	17	55336891	C	T	BMI	0.0173	0.00302	1.24E-08
373	rs78506059	17	56107400	A	G	BMI	0.0207	0.00358	1.05E-08
374	rs55931203	17	65854602	C	T	BMI	-0.0203	0.00264	2.80E-14
375	rs2619976	17	71754545	C	T	BMI	-0.0118	0.00208	6.49E-10
376	rs1285245	17	77796889	C	G	BMI	-0.0127	0.00211	6.04E-10
377	17:7096707:T:TC	17	7096707	T	TC	BMI	0.0129	0.00209	9.67E-10
378	rs35375077	17	16030866	C	T	BMI	-0.0129	0.00206	5.23E-10
379	rs55678940	17	21251092	C	CTGTAAAGAAA	BMI	0.0198	0.00215	6.91E-20
380	rs11080090	17	27502029	A	G	BMI	-0.0155	0.00281	4.67E-08
381	18:52479513:C:CT	18	52479513	C	CT	BMI	0.0151	0.00241	5.50E-10
382	rs200519957	18	56878075	A	AG	BMI	0.0186	0.00275	2.24E-11
383	rs1893514	18	57687573	A	G	BMI	-0.0189	0.00235	1.90E-15
384	18:57850927:G:GTCT	18	57850927	G	GTCT	BMI	0.0538	0.00241	########
385	rs12454712	18	60845884	C	T	BMI	0.0134	0.0021	2.69E-10
386	rs512121	18	7548501	C	T	BMI	-0.0164	0.00261	4.92E-10
387	rs35761930	18	13182325	A	T	BMI	-0.0165	0.00265	6.57E-10
388	rs16940823	18	22137319	A	C	BMI	-0.0159	0.00264	2.64E-09
389	rs11082094	18	36165574	A	G	BMI	0.0139	0.00235	3.80E-09
390	rs559231	18	39644247	G	T	BMI	-0.0144	0.0021	1.22E-11
391	rs749887256	18	40701257	C	CTAATTA	BMI	-0.0141	0.00212	5.79E-11
392	rs72975653	19	1869152	G	T	BMI	-0.0159	0.00206	1.88E-14
393	rs60497719	19	33971746	A	G	BMI	0.0133	0.00235	1.55E-08
394	rs60235724	19	34304260	A	AT	BMI	-0.015	0.00219	1.48E-11
395	rs6857	19	45392254	C	T	BMI	0.0253	0.00271	2.17E-20
396	rs10410266	19	49650872	C	T	BMI	-0.0124	0.00211	6.89E-09
397	rs760129606	19	51812546	A	AAAAGAAAG	BMI	0.0145	0.00208	1.96E-12
398	19:12962283:G:GTTGT	19	12962283	G	GTTGT	BMI	0.0147	0.00231	2.82E-10
399	rs273505	19	18217147	C	T	BMI	0.0178	0.00207	1.60E-17
400	rs8112818	19	18812785	A	G	BMI	0.0207	0.00209	1.01E-22
401	rs113696534	19	30286037	G	GCCTGTAATC	BMI	-0.0206	0.00219	9.38E-21
402	rs6047046	20	2129907	A	G	BMI	0.0116	0.00206	2.22E-08
403	rs2425847	20	44907927	A	G	BMI	-0.0116	0.00208	3.47E-08
404	rs11086469	20	53441622	T	TA	BMI	0.0148	0.00245	2.42E-09
405	rs11472251	20	54158730	G	GTCTC	BMI	-0.0131	0.00216	1.88E-09
406	rs1884897	20	6612832	A	G	BMI	-0.0215	0.00212	7.89E-24
407	20:16561232:G:GA	20	16561232	G	GA	BMI	0.0133	0.00232	1.27E-08
408	rs150890260	20	17099926	A	G	BMI	0.0135	0.00236	1.58E-08
409	rs143850972	20	25189120	A	ATAAT	BMI	-0.015	0.00219	1.38E-11
410	rs4911382	20	32553095	C	T	BMI	-0.0146	0.00207	3.49E-12
411	rs6103254	20	41990761	C	T	BMI	-0.0192	0.0031	8.12E-10
412	rs8132491	21	40288577	A	G	BMI	-0.0153	0.00226	1.90E-11

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	Trait	Effect	Standard error	P-value
413	rs5843995	21	41418349	A	AT	BMI	-0.0131	0.00213	1.14E-09
414	rs1964927	21	42653237	A	G	BMI	-0.0136	0.00214	2.80E-10
415	rs394608	21	46581798	C	T	BMI	0.0177	0.00205	1.60E-17
416	rs11538	22	18220831	A	G	BMI	-0.0166	0.00269	1.12E-09
417	rs165722	22	19949013	C	T	BMI	-0.0113	0.00206	4.65E-08
418	rs2413485	22	38193920	C	T	BMI	0.0116	0.00211	4.57E-08
419	rs543124501	22	40677499	A	AT	BMI	0.0173	0.00221	9.13E-15
420	rs28489620	22	41804716	A	G	BMI	-0.0154	0.00226	1.58E-11
421	rs9615723	22	48386670	C	T	BMI	0.012	0.00209	1.34E-08

Appendix 3: Genetic variants used as genetic instruments for waist-to-hip ratio[222].

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	TRAIT	Effect	Standard error	P-value
1	rs2742690	1	2987268	A	C	WHR	0.0169	0.00256	6.19E-11
2	rs10889560	1	65989878	A	C	WHR	0.0267	0.00378	1.75E-12
3	rs80349432	1	72854380	A	T	WHR	0.0167	0.00252	6.13E-11
4	rs313732	1	86258603	A	G	WHR	-0.0146	0.0021	5.54E-12
5	rs2061708	1	1.03E+08	C	G	WHR	0.0163	0.0021	1.61E-14
6	rs2311050	1	9334328	C	T	WHR	-0.021	0.00243	6.46E-19
7	rs147190211	1	9348566	G	GA	WHR	0.0221	0.00311	1.96E-12
8	rs3789615	1	1.15E+08	C	T	WHR	0.0144	0.00208	6.37E-12
9	rs1289021	1	1.64E+08	C	G	WHR	0.0121	0.00208	6.74E-09
10	rs3119837	1	1.7E+08	C	T	WHR	0.0289	0.00233	1.76E-40
11	rs2001127	1	1.72E+08	C	T	WHR	0.0233	0.00208	9.83E-33
12	rs12131072	1	2E+08	C	T	WHR	0.0132	0.00229	1.31E-08
13	rs10753935	1	2.03E+08	A	G	WHR	-0.0119	0.00214	3.39E-08
14	rs60206411	1	19966039	C	T	WHR	0.0143	0.00244	6.81E-09
15	rs2821226	1	2.04E+08	A	G	WHR	0.00514	0.00208	3.54E-08
16	rs3767848	1	2.14E+08	A	G	WHR	-0.0129	0.00237	1.19E-08
17	rs3897379	1	2.2E+08	A	G	WHR	0.0199	0.00261	2.59E-25
18	rs12042959	1	2.44E+08	A	G	WHR	0.0179	0.00294	1.59E-09
19	rs2298632	1	23710475	C	T	WHR	0.0176	0.0021	8.49E-17
20	rs3768321	1	40035928	G	T	WHR	-0.0164	0.00259	4.28E-10
21	rs6751993	2	635864	A	G	WHR	-0.0269	0.00277	5.53E-22
22	2:60170649:C:CTG	2	60170649	C	CTG	WHR	0.0122	0.0021	3.99E-09
23	2:67659345:A:ATT	2	67659345	A	ATT	WHR	-0.0131	0.00218	2.57E-09
24	rs2861699	2	67855332	C	T	WHR	0.0204	0.00218	7.76E-21
25	rs59786977	2	68459272	C	T	WHR	-0.0124	0.00218	1.92E-08
26	rs6542924	2	1.01E+08	A	C	WHR	-0.0166	0.00223	1.55E-13
27	rs149862546	2	9601872	C	CTT	WHR	0.0125	0.00215	9.25E-09
28	rs75030383	2	1.12E+08	A	C	WHR	-0.0137	0.00233	6.50E-09
29	2:119462725:G:GAC	2	1.19E+08	G	GAC	WHR	0.0162	0.00267	2.09E-09
30	rs7578604	2	1.21E+08	G	T	WHR	-0.00714	0.00237	3.25E-07
31	rs4664013	2	1.61E+08	C	G	WHR	0.0147	0.00221	5.66E-11
32	rs111713556	2	1.65E+08	C	CA	WHR	0.0136	0.00253	4.59E-08
33	rs13389219	2	1.66E+08	C	T	WHR	0.0244	0.00211	6.34E-76
34	rs781693294	2	1.71E+08	C	CAT	WHR	0.0132	0.00212	6.08E-10
35	rs779390	2	13088063	C	T	WHR	-0.0187	0.00207	2.39E-19
36	rs1609303	2	1.82E+08	A	T	WHR	0.0124	0.00214	9.18E-09
37	rs10177093	2	1.88E+08	G	T	WHR	-0.0245	0.00207	9.70E-32
38	rs16846100	2	2.12E+08	A	C	WHR	0.0133	0.00234	2.00E-08
39	rs10184221	2	2.13E+08	A	G	WHR	-0.0144	0.00234	1.08E-09
40	rs208786	2	2.31E+08	A	C	WHR	-0.0126	0.00208	1.98E-09
41	rs199780151	2	2.39E+08	A	AT	WHR	0.0287	0.00452	3.30E-10
42	rs146095395	2	2.4E+08	C	T	WHR	0.0184	0.00294	5.82E-10
43	rs28445639	2	25182488	C	T	WHR	-0.0215	0.00249	1.96E-18
44	rs34552049	2	26944300	A	G	WHR	0.0112	0.00206	5.19E-09

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	TRAIT	Effect	Standard error	P-value
45	rs71393178	2	35483782	T	TTTG	WHR	0.0126	0.00218	1.05E-08
46	rs35050623	2	37063240	C	CT	WHR	-0.0116	0.00208	2.66E-08
47	rs62137421	2	43594091	C	T	WHR	-0.0111	0.0026	1.21E-08
48	rs17326656	2	48962291	G	T	WHR	-0.0174	0.00243	7.43E-15
49	rs34681404	3	9344648	C	CTTTA	WHR	0.0136	0.00244	3.48E-08
50	rs13100572	3	53237601	A	G	WHR	-0.0143	0.0021	7.07E-12
51	rs2049760	3	62482552	A	G	WHR	-0.0143	0.00214	3.98E-11
52	rs2371767	3	64718258	C	G	WHR	-0.0331	0.00231	6.26E-63
53	rs34234711	3	82709447	G	T	WHR	-0.0129	0.00214	2.58E-09
54	rs4856553	3	84972827	C	T	WHR	0.0124	0.00219	1.90E-08
55	rs4577503	3	99657922	A	G	WHR	-0.013	0.00209	8.93E-10
56	rs72628504	3	1.29E+08	A	G	WHR	-0.0323	0.00426	5.53E-18
57	rs764881249	3	1.32E+08	C	CA	WHR	0.0144	0.00231	1.26E-10
58	rs61789562	3	1.36E+08	C	T	WHR	-0.0226	0.0032	2.98E-12
59	rs139016349	3	1.38E+08	T	TTTC	WHR	-0.0126	0.00281	1.14E-10
60	rs9817452	3	1.57E+08	G	T	WHR	0.0298	0.00213	5.97E-44
61	rs998749	3	1.69E+08	A	G	WHR	0.014	0.00207	2.13E-11
62	rs7647305	3	1.86E+08	C	T	WHR	0.0148	0.00252	6.47E-09
63	rs2248422	3	1.88E+08	C	T	WHR	0.0157	0.00257	1.35E-09
64	rs7631359	3	12296751	G	T	WHR	-0.0232	0.00294	2.46E-20
65	rs11718898	3	12848822	C	T	WHR	-0.0151	0.00219	7.55E-12
66	rs9867068	3	35674248	C	G	WHR	-0.0198	0.00239	1.23E-16
67	rs10662748	3	37594504	T	TA	WHR	-0.0156	0.00226	8.46E-12
68	rs13092573	3	46988561	C	T	WHR	-0.0142	0.00216	1.38E-12
69	3:49642946:A:AAT	3	49642946	A	AAT	WHR	-0.0194	0.00214	3.30E-19
70	rs13101828	4	965720	A	G	WHR	0.0178	0.00208	1.44E-17
71	rs57826934	4	56323379	C	T	WHR	-0.0149	0.0022	1.16E-13
72	rs344837	4	77778487	C	T	WHR	0.0205	0.00373	4.93E-08
73	rs201172243	4	78661181	A	AT	WHR	-0.0166	0.00297	2.91E-08
74	rs2167750	4	89730074	C	T	WHR	-0.0208	0.00207	1.53E-38
75	rs974801	4	1.06E+08	A	G	WHR	-0.0112	0.00215	2.19E-11
76	4:120297981:G:GT	4	1.2E+08	G	GT	WHR	0.0137	0.00306	2.44E-08
77	rs308403	4	1.24E+08	C	T	WHR	0.0126	0.00224	2.38E-08
78	rs809955	4	1.41E+08	A	G	WHR	-0.0152	0.00215	2.22E-12
79	rs200457388	4	1.46E+08	C	CA	WHR	0.0126	0.00209	2.70E-09
80	rs4450871	4	4990298	A	G	WHR	0.0121	0.00208	1.47E-15
81	rs13137905	4	15375668	G	T	WHR	-0.0132	0.00234	2.39E-08
82	rs7695004	4	26303026	A	C	WHR	-0.0197	0.00223	2.70E-18
83	rs10938398	4	45186139	A	G	WHR	0.0138	0.00209	5.67E-11
84	rs13188941	5	4009651	C	G	WHR	-0.0143	0.00214	5.67E-12
85	5:87767158:G:GTA	5	87767158	G	GTA	WHR	-0.0173	0.00225	2.52E-14
86	rs13163336	5	87943710	A	C	WHR	0.0208	0.00283	3.46E-13
87	rs200312312	5	1.04E+08	C	T	WHR	0.0135	0.00227	4.36E-09
88	rs56206672	5	1.13E+08	C	T	WHR	0.0176	0.00265	3.23E-11
89	rs10477191	5	1.42E+08	A	G	WHR	-0.036	0.00504	2.93E-14
90	rs3021428	5	1.53E+08	A	C	WHR	0.012	0.00207	1.03E-08

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	TRAIT	Effect	Standard error	P-value
91	rs111330621	5	1.72E+08	A	G	WHR	0.0128	0.00227	2.07E-08
92	rs3836828	5	1.73E+08	T	TG	WHR	0.0109	0.00206	2.70E-10
93	rs17738166	5	1.73E+08	A	G	WHR	0.0121	0.00211	1.20E-08
94	rs6897617	5	1.73E+08	A	G	WHR	0.0196	0.00228	1.95E-19
95	rs244723	5	1.77E+08	A	G	WHR	0.0158	0.00209	7.50E-14
96	rs3836821	5	53305795	A	ACT	WHR	0.0164	0.00236	5.87E-12
97	rs465983	5	55812130	A	G	WHR	-0.0262	0.00243	6.77E-27
98	rs12186959	5	57567840	A	C	WHR	-0.0111	0.00228	2.11E-08
99	rs559785	5	66308227	C	T	WHR	0.0117	0.00209	2.71E-08
100	rs2112347	5	75015242	G	T	WHR	-0.016	0.00215	1.77E-13
101	rs1294415	6	6740633	A	G	WHR	0.0255	0.00213	8.13E-35
102	rs4960246	6	6782298	A	G	WHR	0.0199	0.00214	4.05E-22
103	rs9369409	6	43346462	C	G	WHR	0.0174	0.0021	1.93E-16
104	rs10456526	6	43814625	A	G	WHR	-0.0207	0.00227	2.36E-22
105	rs9370243	6	53789830	G	T	WHR	-0.021	0.00377	3.40E-08
106	rs1902066	6	81346033	C	T	WHR	0.0163	0.00209	1.09E-15
107	rs4434440	6	85419999	C	T	WHR	0.0117	0.00207	2.12E-08
108	rs34815026	6	97385640	C	T	WHR	-0.0114	0.00219	4.94E-08
109	rs7742854	6	98528143	C	T	WHR	0.0146	0.00211	5.26E-12
110	rs1268180	6	1.09E+08	C	G	WHR	0.0152	0.00222	1.17E-11
111	rs11381821	6	1.1E+08	C	CA	WHR	-0.0123	0.00211	5.83E-09
112	rs2800710	6	1.27E+08	C	T	WHR	0.0351	0.00206	6.40E-81
113	rs71562509	6	1.4E+08	G	T	WHR	0.0207	0.0021	5.18E-31
114	rs9459596	6	1.67E+08	A	G	WHR	0.0128	0.00206	7.84E-10
115	rs112266013	6	15230743	A	G	WHR	-0.0172	0.00293	5.72E-09
116	rs1276814	6	23723437	C	T	WHR	-0.0131	0.0022	3.87E-09
117	rs1265097	6	31106459	A	C	WHR	0.0313	0.00381	3.69E-16
118	rs852424	7	5567596	C	T	WHR	-0.0122	0.0022	3.31E-08
119	rs7783234	7	84621549	A	G	WHR	-0.0145	0.00221	7.16E-11
120	rs369428586	7	99134799	C	CA	WHR	-0.0202	0.00291	5.21E-12
121	rs35992811	7	1.02E+08	G	GT	WHR	0.0154	0.00218	3.51E-12
122	7:104885481:A:AACACAC	7	1.05E+08	A	AACACAC	WHR	0.016	0.00213	1.34E-13
123	rs76335902	7	1.08E+08	A	G	WHR	0.0217	0.00356	8.54E-10
124	rs200299025	7	1.21E+08	A	G	WHR	0.0147	0.00219	3.16E-11
125	rs34518086	7	1.3E+08	C	CT	WHR	0.00896	0.0021	2.64E-22
126	rs759544745	7	1.51E+08	C	CGTGTGTGAGT	WHR	-0.0134	0.0021	2.62E-10
127	rs57246313	7	25889697	A	G	WHR	0.0251	0.0026	1.81E-27
128	rs1723939	7	46028577	C	T	WHR	0.0119	0.00209	1.52E-08
129	rs78702390	7	68618770	C	G	WHR	-0.0241	0.00437	4.45E-08
130	rs55747707	7	73037366	A	G	WHR	-0.0159	0.00257	4.61E-14
131	rs77497827	7	75038408	A	G	WHR	-0.0134	0.0023	8.16E-09
132	8:19855844:G:GTATTTT	8	19855844	G	GTATTTT	WHR	-0.0142	0.00241	4.88E-09
133	rs66843503	8	60220561	C	T	WHR	0.0161	0.00282	1.61E-08
134	rs7014001	8	72414289	C	T	WHR	0.04	0.00396	1.91E-24
135	rs12679106	8	73443198	G	T	WHR	0.0142	0.00229	8.36E-10
136	rs112875651	8	1.27E+08	A	G	WHR	-0.0101	0.00214	6.22E-13

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	TRAIT	Effect	Standard error	P-value
137	8:23607534:A:AGAGT	8	23607534	A	AGAGT	WHR	-0.0188	0.00242	8.66E-23
138	rs200395159	8	25435262	G	GT	WHR	0.0276	0.00509	6.52E-08
139	rs11992444	8	25464690	G	T	WHR	-0.0189	0.00207	1.34E-19
140	rs1485741	8	25919020	C	T	WHR	-0.0195	0.00325	2.84E-09
141	rs328086	8	26325935	A	G	WHR	0.0172	0.00298	1.05E-08
142	rs36061954	8	38329650	C	T	WHR	-0.0146	0.00211	2.41E-12
143	rs1411431	9	16728721	A	C	WHR	0.0165	0.00284	8.74E-09
144	rs34081699	9	1.13E+08	A	G	WHR	-0.0111	0.00222	7.02E-09
145	rs12237685	9	1.27E+08	C	T	WHR	0.0138	0.00246	2.60E-08
146	9:134946074:A:ATATCTGCCTGGGGC	9	1.35E+08	A	ATATCTGCCTGGGGC	WHR	-0.0138	0.00222	8.80E-10
147	rs28647893	9	1.37E+08	C	T	WHR	0.0124	0.00206	3.01E-09
148	rs1412239	9	28425515	C	G	WHR	-0.0148	0.0022	2.65E-11
149	rs12001634	9	94186527	A	T	WHR	-0.012	0.00218	4.99E-08
150	rs1326775	9	95359555	C	T	WHR	-0.0183	0.00278	7.88E-11
151	rs10761252	9	96424301	A	G	WHR	0.0137	0.0022	8.16E-10
152	9:96709153:T:TG	9	96709153	T	TG	WHR	-0.0168	0.00228	2.70E-13
153	rs1800978	9	1.08E+08	C	G	WHR	0.0215	0.00315	3.63E-12
154	rs1962883	9	1.08E+08	C	T	WHR	-0.0218	0.00212	3.11E-34
155	rs10795055	10	3581221	A	G	WHR	0.0117	0.00211	2.07E-10
156	10:80916777:T:TGGAGGTGGCTCTCAGAGAAGGCAGCATGGA	10	80916777	T	TGGAGGTGGCTCTCAGAGAAGGCAGCATGGA	WHR	-0.0165	0.00214	2.56E-14
157	rs1799734	10	89690952	T	TTTATC	WHR	-0.0121	0.00217	1.57E-10
158	rs1776209	10	93810872	A	G	WHR	0.0141	0.00228	1.00E-09
159	rs17515035	10	95930225	A	G	WHR	-0.0142	0.00275	1.62E-08
160	rs2781810	10	1.16E+08	C	T	WHR	0.0119	0.00211	2.51E-08
161	rs66923674	10	4955702	G	T	WHR	0.0173	0.00304	1.57E-10
162	rs7099771	10	5654742	A	G	WHR	0.012	0.00214	3.16E-08
163	rs7894565	10	21872913	C	T	WHR	0.016	0.00224	1.95E-13
164	rs72778386	10	32230774	A	G	WHR	-0.014	0.0024	7.83E-09
165	rs1757471	10	34168090	C	T	WHR	-0.0132	0.00206	1.29E-10
166	rs2767073	10	36229012	A	G	WHR	-0.0124	0.00209	4.35E-09
167	rs7070670	10	61842645	C	T	WHR	0.0123	0.00221	3.24E-08
168	rs5791099	11	36328427	T	TA	WHR	0.0123	0.00218	9.91E-12
169	11:43610528:T:TTTATTA	11	43610528	T	TTTATTA	WHR	-0.0132	0.00212	8.86E-10
170	rs7128207	11	45230061	G	T	WHR	0.0124	0.00208	3.51E-09
171	rs2509963	11	62192931	C	T	WHR	0.016	0.00236	9.72E-12
172	rs10896012	11	65278461	C	T	WHR	0.0176	0.00251	3.97E-12
173	rs2276106	11	66105817	C	T	WHR	-0.0171	0.00229	1.63E-13
174	rs150648223	11	68942162	A	ATTT	WHR	0.0214	0.00357	3.24E-09
175	rs11042029	11	8688125	A	C	WHR	-0.0156	0.00232	3.23E-11
176	rs647248	11	89917699	A	G	WHR	0.0117	0.0021	3.16E-08
177	11:111878217:G:GAC	11	1.12E+08	G	GAC	WHR	-0.0219	0.00216	4.02E-24
178	rs782229786	11	1.13E+08	A	AT	WHR	0.0187	0.00325	1.14E-08
179	rs719802	11	1.13E+08	C	T	WHR	-0.0134	0.00212	3.53E-10
180	rs11216183	11	1.17E+08	A	C	WHR	0.0261	0.00358	6.00E-13
181	rs579682	11	1.22E+08	C	T	WHR	0.0136	0.0023	4.75E-09
182	rs34154767	11	9978117	T	TC	WHR	-0.0156	0.00233	2.81E-11

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	TRAIT	Effect	Standard error	P-value
183	rs2957658	11	10393468	A	G	WHR	-0.0116	0.00207	2.73E-08
184	rs11029441	11	26279145	C	T	WHR	0.0211	0.00383	1.60E-08
185	rs10835211	11	27701365	A	G	WHR	0.019	0.00235	1.09E-15
186	12:1006052:C:CCACACACACACA	12	1006052	C	CCACACACACACA	WHR	-0.0143	0.00227	2.76E-10
187	rs894739	12	54346869	C	T	WHR	-0.0303	0.0025	6.57E-36
188	rs11176010	12	66425335	A	G	WHR	-0.0161	0.00231	1.03E-12
189	rs7311622	12	98772975	C	T	WHR	0.012	0.00208	9.88E-09
190	rs1922432	12	1.07E+08	A	G	WHR	0.012	0.00217	2.62E-09
191	rs3764002	12	1.09E+08	C	T	WHR	0.0206	0.00235	3.01E-18
192	rs35806699	12	1.22E+08	T	TC	WHR	-0.0182	0.00309	5.94E-09
193	12:123697258:C:CA	12	1.24E+08	C	CA	WHR	0.0176	0.00263	2.97E-18
194	rs1727332	12	1.24E+08	C	T	WHR	-0.0225	0.00241	1.54E-21
195	rs11057413	12	1.24E+08	A	G	WHR	0.03	0.00219	1.45E-58
196	rs12580347	12	3388932	C	T	WHR	0.0125	0.0021	1.61E-09
197	rs10771216	12	9076740	G	T	WHR	-0.0129	0.0023	2.76E-08
198	rs150460514	12	14415382	A	G	WHR	-0.0145	0.00259	3.06E-08
199	rs718314	12	26453283	A	G	WHR	-0.0271	0.0024	2.96E-38
200	rs771716856	12	33712334	T	TG	WHR	-0.0148	0.00216	1.11E-11
201	13:31024245:A:AAAAC	13	31024245	A	AAAAC	WHR	0.0175	0.00228	3.47E-14
202	13:51193971:A:AACAC	13	51193971	A	AACAC	WHR	0.0252	0.00232	3.70E-27
203	rs7982447	13	54453811	C	T	WHR	0.0147	0.00256	1.30E-08
204	rs7985834	13	54695928	A	C	WHR	-0.0215	0.00314	1.38E-11
205	rs4772092	13	99166614	A	G	WHR	-0.0118	0.00214	2.37E-08
206	rs664532	13	1.11E+08	C	T	WHR	-0.0119	0.0021	1.90E-08
207	rs112697444	14	23748201	A	AT	WHR	-0.0123	0.00217	1.89E-08
208	rs58939796	14	25933965	G	T	WHR	0.0107	0.00226	1.57E-11
209	rs1317805	14	53947075	A	G	WHR	0.0137	0.00244	2.71E-08
210	rs2412107	14	65426216	G	T	WHR	-0.0145	0.00252	1.29E-08
211	rs4902632	14	69149428	A	T	WHR	0.0157	0.00278	2.24E-08
212	rs2526886	14	71359064	G	T	WHR	-0.0124	0.00225	4.07E-08
213	rs11159989	14	91531842	G	T	WHR	-0.0124	0.00223	2.95E-10
214	rs12441765	15	31696936	C	T	WHR	0.0146	0.00245	3.69E-09
215	rs72757416	15	92574321	A	C	WHR	0.0149	0.00254	6.34E-09
216	rs8024294	15	94023132	A	G	WHR	0.0189	0.00337	2.36E-08
217	rs2603229	15	1E+08	C	T	WHR	-0.0133	0.00241	4.04E-08
218	rs2456530	15	53091553	C	T	WHR	-0.019	0.00306	7.95E-10
219	rs1657930	15	57120989	A	G	WHR	-0.0159	0.00262	1.87E-09
220	rs3784692	15	67988133	C	T	WHR	-0.0143	0.00211	1.57E-18
221	16:69564990:G:GA	16	69564990	G	GA	WHR	-0.0197	0.00212	3.47E-20
222	rs2925979	16	81534790	C	T	WHR	-0.0242	0.00226	3.65E-45
223	rs35193388	16	4420817	G	GC	WHR	0.0166	0.00236	3.73E-12
224	rs760117	16	4941047	C	T	WHR	-0.0169	0.00304	8.37E-09
225	rs13329943	16	24733751	C	T	WHR	-0.0164	0.00234	4.41E-12
226	rs11642015	16	53802494	C	T	WHR	-0.0399	0.00211	9.44E-81
227	rs235129	16	66228955	A	G	WHR	0.0117	0.00212	4.52E-08
228	rs4073992	17	1845462	A	T	WHR	0.0126	0.00224	2.60E-08

	RSID	Chromosome	Position (GRCh37)	Effect allele	Other allele	TRAIT	Effect	Standard error	P-value
229	rs8882	17	34869155	A	G	WHR	0.0165	0.00208	4.13E-15
230	rs7215553	17	40896077	C	T	WHR	-0.00959	0.00208	1.13E-07
231	rs15689	17	46684929	A	G	WHR	0.0145	0.00239	2.01E-09
232	rs9905140	17	59496238	C	T	WHR	-0.013	0.00212	1.30E-09
233	rs35087650	17	2058349	A	ATT	WHR	-0.0156	0.00242	1.67E-10
234	rs4789168	17	73313277	A	G	WHR	0.016	0.00247	1.40E-10
235	rs9893691	17	4185624	C	T	WHR	0.0149	0.00229	1.22E-10
236	rs7222664	17	9785283	C	G	WHR	0.0126	0.0022	1.00E-08
237	rs12936587	17	17543722	A	G	WHR	-0.0145	0.00207	4.48E-12
238	rs10655994	17	18247891	G	GA	WHR	0.014	0.00231	7.16E-11
239	rs78879665	17	21280113	A	AG	WHR	0.0147	0.00217	1.98E-11
240	rs5819871	17	27878029	A	AAAG	WHR	-0.0102	0.00207	7.89E-09
241	17:31502546:C:CT	17	31502546	C	CT	WHR	0.0122	0.00219	3.65E-08
242	rs11664106	18	2846812	A	T	WHR	0.0201	0.00219	5.02E-28
243	rs1548343	18	63319714	C	G	WHR	-0.0165	0.00264	7.18E-10
244	rs571218	18	13127923	A	G	WHR	0.0118	0.00209	2.31E-08
245	18:21069873:T:TG	18	21069873	T	TG	WHR	-0.0164	0.00219	1.11E-13
246	rs8085083	18	40677365	A	G	WHR	-0.0132	0.0021	5.66E-10
247	rs627840	18	42446447	A	G	WHR	0.0186	0.00323	1.19E-08
248	rs7239114	18	45921214	A	G	WHR	0.0115	0.00209	3.17E-08
249	rs12967135	18	57849023	A	G	WHR	0.0264	0.00244	8.02E-27
250	18:57968538:A:AAAAAAG	18	57968538	A	AAAAAAG	WHR	0.0219	0.00232	6.54E-21
251	rs7249081	19	2157167	C	T	WHR	-0.0145	0.00208	6.16E-12
252	rs2903738	19	32946043	A	T	WHR	0.0142	0.00249	1.73E-08
253	19:33785832:C:CA	19	33785832	C	CA	WHR	-0.0351	0.00289	1.82E-33
254	rs12461964	19	41341229	A	G	WHR	-0.0118	0.00207	1.61E-09
255	rs429358	19	45411941	C	T	WHR	-0.0339	0.00286	4.74E-32
256	rs8103017	19	55999142	C	G	WHR	-0.0162	0.00229	2.62E-12
257	rs762433880	19	7197588	G	GGTGT	WHR	-0.0132	0.00233	3.81E-12
258	rs73504817	19	17167723	C	T	WHR	-0.0128	0.00228	2.73E-09
259	rs58726300	19	18235750	C	CT	WHR	0.0203	0.00218	1.46E-20
260	19:30303379:G:GA	19	30303379	G	GA	WHR	0.014	0.00212	5.47E-11
261	rs35679975	20	51194941	C	T	WHR	-0.0255	0.00263	6.91E-22
262	rs6090040	20	62692060	A	C	WHR	0.0103	0.00208	6.99E-12
263	rs805770	20	5668714	C	T	WHR	-0.0206	0.00212	3.33E-22
264	rs6138524	20	25150535	A	T	WHR	-0.0149	0.00256	7.86E-09
265	rs2865505	20	39203782	A	G	WHR	-0.0117	0.00223	1.00E-10
266	rs4812492	20	39938122	C	T	WHR	0.0143	0.0021	1.76E-11
267	rs6066101	20	45520152	A	G	WHR	0.0231	0.00213	1.77E-27
268	rs2823096	21	16522082	A	G	WHR	0.0158	0.00267	4.49E-09
269	rs2836171	21	39526409	A	G	WHR	-0.0141	0.00208	1.29E-12
270	rs5844832	22	29455687	C	CCTT	WHR	0.0231	0.00219	1.78E-25
271	22:38191342:C:CT	22	38191342	C	CT	WHR	0.0134	0.00226	4.14E-09
272	rs138049	22	40627059	A	C	WHR	-0.0133	0.00209	3.00E-10
273	22:43282457:G:GGTAT	22	43282457	G	GGTAT	WHR	0.0121	0.00208	1.06E-08