SUPPLEMENTAL MATERIAL

Table of Contents

Supplemental Methods and Results

Samples, Genotype Quality Control and Imputation

We analyzed individual-level data from five cohorts that study incident CVD among participants without a previous history of CVD (Framingham Heart Study, Cardiovascular Health Study, Atherosclerosis Risk in Communities Study, Multi-Ethnic Study of Atherosclerosis, and Women's Health Initiative). Raw genotype and phenotype data were downloaded from dbGAP (accession numbers phs000007.v29.p11, phs000287.v6.p1, phs000209.v13.p3, phs000280.v4.p1, phs000200.v11.p3, phs000888.v1.p1). For each of the five cohorts, extensive genotype quality control procedures were followed according to the current standard for genomic studies[10].

Specifically, to avoid genotyping errors affecting our study findings, filtering was performed at the individual level and at the variant level as follows. For each cohort we performed individual level filtering (excluding individuals with assigned and discrepancy between reported and genotype-inferred sex, extreme deviations from heterozygosity, or missingness). For every group of individuals that were related (identity by descent (IBD) >0.125) we randomly selected one. In addition, for each study, variant level filtering was performed to exclude variants that had significant deviations from Hardy Weinberg equilibrium (Hardy-Weinberg p-value$<1*10^{-4}$), a minor allele frequency (MAF) less than 0.01 and missing call rate > 0.05[10]. For studies that analyzed their populations with different genotyping arrays, we also excluded variants that had significant deviation in MAF between the different arrays. For individuals that were genotyped in more than one genotyping array, we selected the array that had the most extensive genotyping. Since genotyping array and race group can affect imputation, we further broke down each cohort into substudies such that each substudy includes only individuals of the same race group (European, African, East Asian or Admixed American) and same genotyping array. We subsequently proceeded with imputing and analyzing each substudy separately for every cohort.

We imputed each study to the 1000 Genomes phase 3 reference panel using Minimac3[30] after pre-phasing with Eagle[31] on the Michigan Imputation Server. Prior to imputation, we lifted all single nucleotide polymorphisms (SNPs) to the hg19 human genome build using the UCSC liftOver tool, aligned all SNPs to the positive strand and filtered out SNPs whose minor allele frequencies deviated by >0.2 compared to the reference panel's MAF and A/T or G/C SNPs with MAF>0.4 as those are prone to strand alignment errors. After imputation, we excluded all imputed SNPs with imputation r squared (INFO score) < 0.7, SNPs with MAF <0.01 and SNPs with Hardy-Weinberg p-value <1e-4.

Exposure and Outcomes

We evaluated the effect of DBP on a series of outcomes. Diastolic BP was measured during each patient's first (baseline) clinic visit. If the first clinic visit BP was not available for a given participant, we used BP measured at a subsequent visit in which BP data were recorded. BP definition in each study is provided in Online Table 1. Individuals with extreme outlier values of DBP (DBP outside the range of mean ± 3 standard deviations) were excluded from further analyses over concern that these recordings were likely to be inaccurate. As a sensitivity analysis, we also performed Winsorization of the outlier DBP values in which DBP measures outside the range of mean ± 3 standard deviations were set to the closest DBP value of the included DBP range and we repeated our primary outcome Mendelian randomization (MR) analysis with that approach. Similarly, individuals with SBP values outside the range of mean ± 3 standard deviations were excluded from the MR analyses pertaining to SBP. Our primary outcome of interest was MI. Individuals with prevalent MI prior to enrollment or incomplete data of time to incident MI were also excluded from our primary analyses. Secondary CVD outcomes analyzed included coronary artery disease events, coronary revascularization, ischemic stroke, heart failure, fatal coronary disease, and all-cause mortality. All outcomes were evaluated in a time-to-event approach. More complete definitions of the primary and secondary outcomes are provided in Online Table 1.

Observational association evaluations

As baseline evaluation, we assessed the relationship between DBP and CVD outcomes using traditional analyses of the observational data. We applied a proportional hazards time-to-event model to assess the influence of DBP or systolic BP (SBP) on each of our primary and secondary outcomes controlling for age, sex and a dummy categorical variable for each individual cohort. Statistical significance was calculated based on the Wald test. In addition, we evaluated the potential non-linear relationship between either DBP or SBP with each outcome using a multivariable meta-analytic approach[11]. Specifically, for each cohort we fitted a second-degree fractional polynomial on the exposure (DBP or SBP) and performed proportional hazards regression for MI on the exposure, controlling for age

and sex. We then ran a multivariate meta-analysis on the estimates of the fractional polynomial model and plotted the hazard ratio of the outcome per unit change in the exposure (BP), centering at the mean exposure across studies.

Instrumental Variable

We generated a polygenic risk score (PRS) of DBP to use as an instrumental variable in our MR analysis. To generate the score, we used summary statistics from a blood pressure GWAS of 750,000 individuals[32] , which is to our knowledge the largest conducted to-date, in an effort to get the most accurate effect size estimates. We thresholded those summary statistics to select only SNPs that pass the genome-wide significance threshold (p-value < $5*10^{-8}$) for association with DBP and performed linkage disequilibrium (LD) based clumping at an LD threshold of $r^2<0.001$ for all variants within a 1Mb window and a secondary clumping with a threshold $r^2<0.1$ for all variants within the same chromosome using reference LD from the 1000 Genomes project. The sequential clumping approach was performed to ensure that only independent variants will be included in the PRS while also avoiding the exclusion of distant variants in very low LD ($r^2<0.1$) with each other (which is likely to represent noise rather than true dependencies between variants). This process resulted in 718 SNPs selected for inclusion in the DBP PRS instrument. Although recent literature suggests that use of a higher p-value threshold can result in higher power for prediction of the exposure[33], we avoided that approach over concern for introducing weak instrument bias in our MR estimates[34]. After selection of the candidate SNPs for MR analysis, we calculated a risk score for DBP in each of our genotyped individuals as a weighted sum of the effect size estimates from the 750,000 GWAS summary statistics and the genotype in each selected SNP: $PRS_i = \sum_{j=1}^{M} \beta_j G_{ij}$.

For 2 of the sensitivity analyses (described in more detail below), we (i) additionally calculated a PRS for SBP following the same procedure described above for DBP and (ii) we generated PRS for DBP independent of body mass index (BMI) or heart rate by excluding all variants in the original DBP PRS that were also associated with BMI or heart rate in recent large BMI[12] or heart rate[13] GWAS at a Bonferroni-adjusted p-value < 0.05. The latter sensitivity analysis was conducted to reduce the likelihood of BMI or heart rate introducing pleiotropy in any association between DBP and CVD. For our multivariate MR analysis, we also generated a combined PRS for both SBP and DBP by taking the union of all SNPs included in the DBP and SBP PRS. We added to that PRS SNPs included in a pulse pressure (PP) PRS which are expected to be relevant as they contain information on SNPs with discordant effects on SBP and DBP. We kept only one SNP of each pair that had an LD $r^2>0.001$ within 1Mb and $r^2>0.1$ within the same chromosome. For each pair that fit in these LD categories we selected the SNP that had the strongest association (lowest p-value) with DBP. This procedure resulted in the selection of 1015 SNPs for the combined PRS of both DBP and SBP.

Linear Mendelian Randomization

We performed a two-step MR analysis for the linear effect of DBP as follows. After merging together all the individual sub-studies in a common design matrix, we first regressed the DBP on the PRS via linear regression to generate an estimate of the effect of the PRS on the exposure. This estimate was the denominator of our MR estimate. We then regressed the effect of the PRS on the outcome using a proportional hazard (Cox) regression analysis for the time-to-event data to generate an estimate of the PRS on the outcome. This estimate (logarithm of the hazard ratio (HR)) was the numerator of our MR estimate. To get the MR hazard ratio of the outcome per unit increase in DBP we exponentiated the ratio of the numerator over the denominator. 95% confidence intervals (CI) of that estimate were generated using a first order Taylor series approximation (standard error (SE) of MR = $SE_{outcome}/beta_{exposure}$). Both regressions included as covariates the age, sex, and 2 genotype principal components (PCs) for each substudy (to account for population stratification within each substudy). As we naturally expected some variability in the estimates between individual substudies due to ancestry and genotype array differences, we included an additional binary dummy covariate to control for each individual's membership in each substudy. Missing values for each principal component (corresponding to values in substudies other than the substudy in which the PCs were obtained), were set to out-of-range values.

We performed a number of sensitivity analyses. In the first analysis, we independently validated our MR estimate between DBP and MI by performing two-sample MR using the summary statistics of the aforementioned blood pressure GWAS and summary statistics of a GWAS on MI from the UK BioBank. In addition to the inverse-variance weighted two-sample MR estimates, we calculated additional estimates of MR-median and MR-Egger that largely account for pleiotropic effects[35,36]. In the second analysis, we aimed to evaluate whether anti-hypertensive medication use could be influencing our MR or observational estimates. Although in general medication use is not considered a

major confounder in MR studies as medications are affecting the outcome only via the exposure, we additionally performed a sensitivity analysis in a subset of individuals (n=15,806) that had available medication use data and were not taking anti-hypertensive medications to demonstrate that phenomenon. In the third analysis, we specifically focused on the role of SBP and BMI as potential sources of pleiotropy in the associations between DBP and outcomes. Since the majority of DBP variants were not associated with BMI, we were able to evaluate the independent role of DBP in outcomes by generating a PRS independent of BMI following the procedure described in the previous section. For SBP this was not feasible as the vast majority of variants that were included in the DBP PRS were also associated with SBP. Indeed, only 146 (20%) of the 718 DBP PRS variants had a Bonferroni adjusted p-value of association with SBP > 0.05 and those variants had on average a lower strength of association with DBP and not enough power to perform MR. Therefore, we instead followed a multivariable MR approach to examine the independent influence of SBP and DBP on cardiovascular outcomes as described previously[37]. Specifically, we first generated a combined PRS of SBP and DBP and PP as detailed in the previous section. We then performed a multivariate multiple linear regression (using SBP and DBP as dependent variables, each SNP of the PRS that was genotyped or adequately imputed in >95% of our participants (828 total SNPs) as the regressors and age, sex, genotype PCs, and the dummy variable for each substudy as covariates) and obtained fitted SBP and DBP estimates based on this model. We subsequently performed a proportional hazards regression of time-to-incidence of each of our primary and secondary outcomes on the fitted values of SBP and DBP controlling for the same covariates. To increase confidence in those estimates, we supplemented this analysis with a summary-level multivariable MR as implemented in the MVMR package[38]. For that, we used the effect sizes and standard errors from a large-scale GWAS[32] for SBP and DBP for each SNP included in the PRS and their corresponding effect sizes (log(Odds Ratio)) and standard errors for their association with MI[39], coronary artery disease[40], heart failure[41] and ischemic stroke[42] based on different large GWAS.

Non-linear Mendelian Randomization

Mendelian randomization analyses to assess for a potential non-linear U-shape effects of DBP on the outcomes were performed using state-of-the-art methodologies[14,15]. In brief, there are three options to assess a non-linear MR association. The first is to evaluate for non-linearities in the relationship between the instrument variable and the outcome. That falls short in most situations, including our study, by the fact that generally the range of the PRS only accounts for a relatively narrow range of the exposure. The second approach is to separate the population in quantiles of the exposure and assess the MR relationship between the exposure and the outcome in each quantile. The main limitation of this method is that, because the PRS is associated with the exposure, conditioning on quantiles of the exposure can induce false associations between confounders and the PRS (a process called moralization). To overcome these limitations, Burgess et al.[14] proposed the strategy of conditioning on quantiles of instrument variable-free exposure which provides the benefit of assessing the full range of the exposure without biasing the results by inducing moralization. This is the approach we followed in our study.

Specifically, we first calculated instrument variable free DBP by taking the residuals of the regression of DBP on the PRS (i.e., IV-free-DBP=DBP-β*PRS). We then split the population into centiles of the instrument variable free DBP and generated linear MR estimates (localized average causal effect estimates) of the effect of DBP on MI in every centile separately- following the procedure described above in the linear MR section. Lastly, we fitted a fractional polynomial model in a meta-analysis of the MR estimates in each centile and evaluated the shape of the best fit among all second-degree fractional polynomials. We plotted the best fit fractional polynomial selected by the model to visualize the relationship between DBP and cardiovascular outcomes and assessed two different p-values for non-linearity: A p-value for the degree of the fractional polynomial, and a linear trend p-value based on meta-regression. This analysis was performed based on the nlmr R package[43] created by the group that introduced the IV-free non-linear MR methodology, modified to allow for a proportional hazards model fit in the relationship between the PRS and the outcome. In addition, as a further sensitivity analysis, we checked for the direction of MR association between DBP and cardiovascular events on a level of instrument variable-free DBP of < 70 mmHg, which is the point below which prior observational studies report a negative association between DBP and MI[5].

Finally, we performed several sensitivity analyses to assess how the shape of the association between DBP and outcomes may change in response to different potential pleiotropic factors and confounders. Specifically, following the fractional polynomial approach described above, we visualized the shape of the association between DBP and outcomes separately in males vs. females. We additionally evaluated the shape of the aforementioned associations using the DBP PRS that is independent of BMI. Lastly, since it proved difficult to disentangle the effects of SBP and

DBP in multivariate MR, we performed an additional analysis in which we generated LACE estimates of DBP and SBP on MI using IV-free exposures in which we remove the genetic effects of both SBP and DBP. Specifically, we first generate quantiles of IV-free DBP after residualizing the PRS for both DBP and SBP. We then calculate LACE MR estimates of DBP and MI using the DBP PRS and subsequently repeat the same process for SBP. This approach is expected to adjust for the genetic linear effects of both SBP and DBP when generating the quantiles of IV-free exposure, thereby allowing us to interrogate non-linearity in MR estimates when adjusting for the linear effects of both SBP and DBP.

Genotype quality control was performed using PLINK versions 1.9 and 2.0, while all statistical analyses were conducted using R version 3.5.1, we considered a p-value (2-sided) of <0.05 to be statistically significant but we also report the actual p-values in all associations.

Simulation study for power assessment

To evaluate our power for detecting non-linear relationships between DBP and the outcomes of interest in our study we performed a simulation study as follows. Using the DBP data in our cohorts we simulated event times by randomly drawing times from a Weibull distribution with a fixed shape parameter of 2 and assuming a piecewise linear log hazard ratio (fixed at 0.05 (translated into an HR of 1.05) for individuals with DBP>= 70mmHg and negative (HR<1) for individuals with DBP <70 mmHg). We then drew right censoring times following an exponential distribution and selected the minimum of pairwise times between the two distributions to determine the censoring time and status. The shape parameter of the Weibull distribution and the lambda parameter of the exponential were set to values that lead to average censoring times of ~10-15 years and average rate of events of ~7-11% (similar to our actual data that had average follow up of 16 years and event rate of 7.8%). We varied the value of the negative log hazard ratios in a range of -0.5 to -0.005 (corresponding to HR 0.6-0.995) and performed 1000 random replicates for each value. For each value of the negative log hazard ratio we evaluated the MR association between DBP and the random outcome among individuals with IV-free DBP < 70mmHg and the shape of the relationship between DBP and the outcome using the non-linear fractional polynomial MR method described above. We then plotted the fraction of replicates in which the estimate for the MR association between DBP and random outcome among individuals with IV-free DBP < 70mmHg was negative, the fraction of replicates in which that estimate was negative and statistically significant and the fraction of replicates in which the non-linear MR method showed evidence of non-linearity between DBP and the random outcome at a p-value of < 0.1.

As a simulation, this analysis should be viewed with caution as there are no guarantees that actual survival data would follow a Weibull distribution, that the log hazard would behave as piecewise linear or that the hazard would change sign exactly at DBP of 70 mmHg. However, the results show that if those assumptions hold, our log hazard MR estimate among individuals with IV-free DBP <70 mmHg would have been negative over 85% of the time even with an actual value of the HR in DBP < 70 mmHg being as high as 0.995.

Evaluation of the impact of sample overlap between cohorts on MR estimates

Although we chose to use different samples for the generation of our instrumental variables and MR estimation throughout our analyses, we should note that there is still some overlap between the cohorts (6.5% overlap between the BP GWAS and our primary study cohorts and 57% overlap between the BP GWAS and the UK Biobank used in linear MR sensitivity analysis). Burgess et al. have established via extensive simulation studies that although MR estimates are asymptotically unbiased even in a single-sample MR study, there can be substantial bias when samples are limited in size which depends on the population F-statistic and the degree of overlap between the studies[16]. To evaluate how this bias could affect our results, we estimated the population F-statistic according to the calculation used in the aforementioned study ($F = \frac{N-K-1}{K} \frac{R^2}{1-R^2}$, where N is the sample size, K the number of variants comprising the instrument and R^2 the estimate of the percent variance of the exposure explained by the instrumental variable). For the above calculation, we used the R^2 estimate from the genome-wide significant SNPs reported in the BP GWAS (4.46%), N the number of individuals in the GWAS (757,601) and K the number of SNPs in our PRS (718). This results in an F statistic estimate of 50.8 with a lower bound of its 95% confidence interval at 49.9 (see Burgess et al. Web Appendix A3 for calculation of the confidence interval[16]). Hence, given the strength of this instrument and for the degree of overlap between our cohorts, considerable bias would not be expected.

Supplemental Tables

Table I. Definitions of the main outcomes in the different cohorts

Cohort	**BP Definition**
ARIC	Average of 2nd and 3rd of 3 manual sitting blood pressure readings at rest
CHS	Supine manual blood pressure at rest
Framingham	Average of two manual sitting blood pressure readings at rest
MESA	Average of 2nd and 3rd of 3 automated blood pressure readings after 5 minutes in the sitting position
WHI	Average of two manual sitting blood pressure readings at rest
Cohort	**Myocardial Infarction Definition**
ARIC	Two of the following: Cardiac chest pain, Diagnostic ECG pattern, Cardiac enzymes Events were adjudicated by clinicians.
CHS	Two of the following: Cardiac chest pain, Diagnostic ECG pattern, Cardiac enzymes Events were adjudicated by the investigators until June 2015.
Framingham	Two of the following: Symptoms indicative of ischemia, increase in biomarkers of myocardial necrosis, serial changes in ECG suggestive of an infarction Events were adjudicated by the investigators.
MESA	In most cases, definite or probable MI required either abnormal cardiac biomarkers (2 times upper limits of normal) regardless of pain or ECG findings; evolving Q waves regardless of pain or biomarker findings; or a combination of chest pain, and ST-T evolution or new LBBB, and biomarker levels 1–2 times upper limits of normal. Events were adjudicated by the investigators.
WHI	Adjudicated based on elements of the medical history, results of cardiac enzyme/troponin determination, and ECG readings Events were adjudicated by the investigators.
	Fatal Coronary Disease Definition
ARIC	Either definite fatal MI or death with: No known non-atherosclerotic or non-cardiac atherosclerotic process or event that was probably lethal, or presence of one or both of the following findings: history of chest pain within 72 hours of death, history of ever having had chronic ischemic heart disease such as coronary insufficiency or angina pectoris. Events were adjudicated by clinicians
CHS	Fatal CHD is classified into: (1) definite fatal MI (MI within 4 weeks); (2) definite fatal CHD (no known non-atherosclerotic cause and either chest pain within 72 hours of death or a history of chronic ischemic heart disease in the absence of valvular heart disease or non-ischemic cardiomyopathy); (3) possible fatal CHD (no known non-atherosclerotic cause and death certificate consistent with CHD as the underlying cause) Events were adjudicated by the investigators until June 2015.
Framingham	Either sudden death (less than one hour from symptom onset) without an obvious alternative cause or non-sudden death if the available information implies that the cause of death was probably CHD. Events were adjudicated by the investigators.
MESA	Definite fatal CHD required a documented MI within the previous 28 days, chest pain within the 72 hours before death, or a history of CHD, and required the absence of a known non-atherosclerotic or non-cardiac cause of death. If the definite fatal CHD criteria were not met, possible fatal CHD could be assigned with an underlying cause of death consistent with fatal CHD and required the absence of a known non-atherosclerotic or non-cardiac cause of death. Events were adjudicated by the investigators.
WHI	Death consistent with Coronary disease as the cause plus any of: pre-terminal hospitalization for an MI within 28 days of death, previous angina or MI and no other known potentially lethal process, death from CABG or angioplasty Events were adjudicated by the investigators.

	Coronary revascularization Definition
ARIC	Coronary artery bypass graft or percutaneous coronary angioplasty
CHS	Coronary artery bypass graft or percutaneous coronary angioplasty
Framingham	Coronary artery bypass graft or percutaneous coronary angioplasty
MESA	Coronary artery bypass graft or percutaneous coronary angioplasty
WHI	Coronary artery bypass graft or percutaneous coronary angioplasty
	Heart Failure Definition
ARIC	Hospitalization with a heart failure diagnosis according to ICD codes in any position or a death certificate with death from heart failure in any position Events were based on ICD codes without adjudication.
CHS	The participant must have both a congestive heart failure diagnosis by a physician and be under treatment with medications for congestive heart failure Events were adjudicated by the investigators until June 2015.
Framingham	A definite diagnosis of congestive heart failure requires that a minimum of two major or one major and two minor criteria* be present concurrently. The presence of other conditions capable of producing the symptoms and signs are considered in evaluating the findings. Events were adjudicated by the investigators.
MESA	Heart failure presence adjudicated by MESA investigators based on presence of symptoms and imaging findings attributable to heart failure along with a diagnosis of heart failure by a physician and medical treatment for heart failure. Events were adjudicated by the investigators.
WHI	The participant must have both a congestive heart failure diagnosis by a physician and be under treatment with medications for congestive heart failure Events were adjudicated by the investigators.
	Ischemic Stroke Definition
ARIC	Neurologic symptoms compatible with stroke and CT/MRI with no evidence of hemorrhage and lumbar puncture either not done or no evidence of hemorrhage or infection Events were adjudicated by clinicians.
CHS	CT/MRI compatible with ischemic stroke or autopsy findings compatible with ischemic stroke Events were adjudicated by the investigators until June 2015.
Framingham	Sudden onset of a focal neurologic deficit lasting longer than 24 hours, in the absence of intracranial hemorrhage, known hypercoagulable state or other disease processes causing focal neurologic deficits Events were adjudicated by the investigators.
MESA	Rapid onset of a documented focal neurologic deficit lasting 24 hours or until death, or if < 24 hours, there was a clinically relevant lesion on brain imaging. For ischemic stroke, imaging should exclude hemorrhage or SAH. Patients with focal neurologic deficits secondary to brain trauma, tumor, infection, or other non-vascular cause were excluded. Events were adjudicated by the investigators.
WHI	Focal neurologic deficit without CT or MRI scan, LP, or evidence of blood OR CT or MRI scan with mottled cerebral pattern or showing decreased density in a location compatible with reported symptoms and signs OR Surgical or autopsy evidence of ischemic infarction (cerebral thrombosis or cerebral embolism) Events were adjudicated by the investigators.

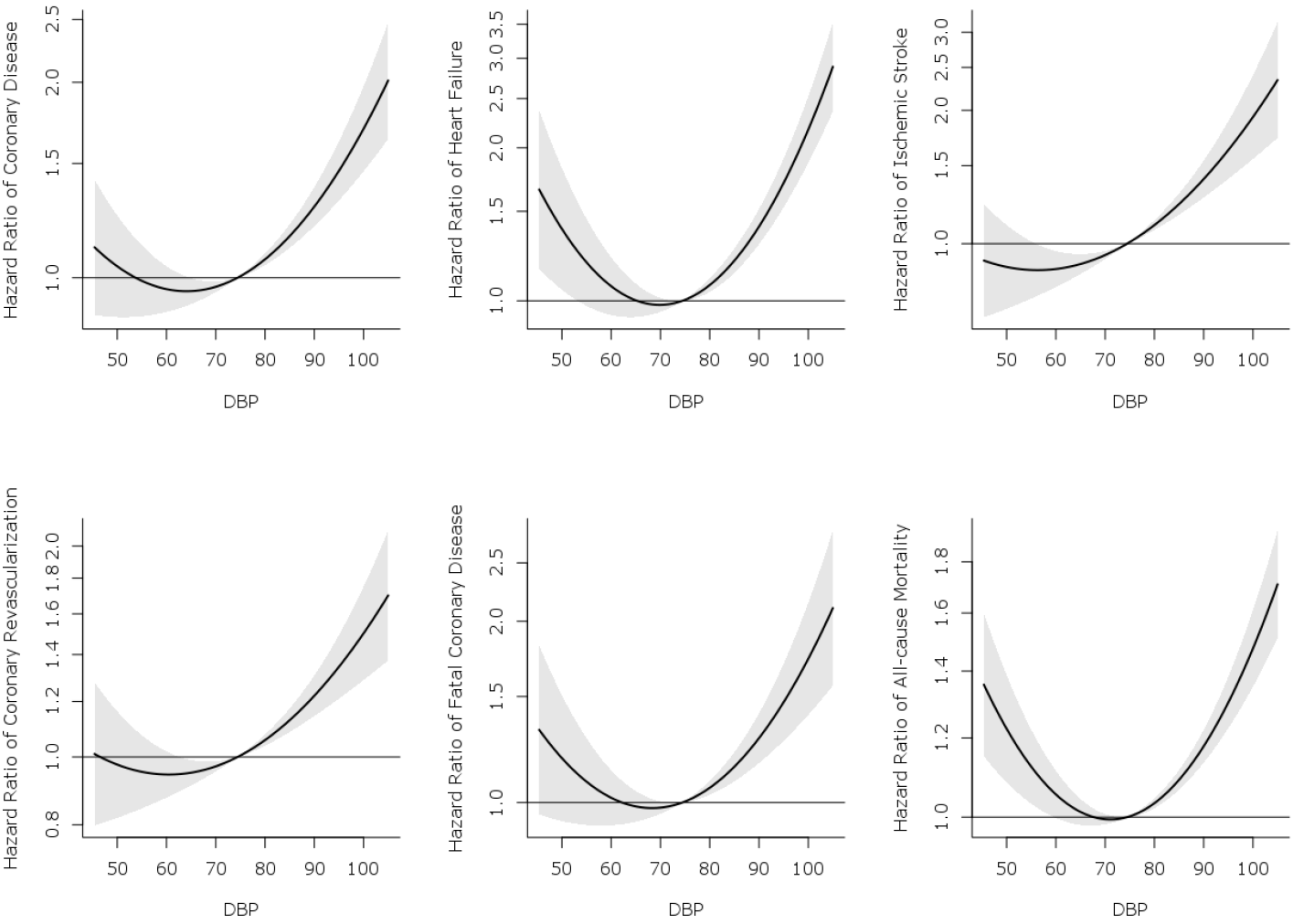
Table II. Break-down of demographic and outcome information among the different cohorts

	ARIC		CHS		Framingham		MESA		WHI	
Parameter	**Median**	**Q1-Q3**	**Median**	**Q1-Q3**	**Median**	**Q1-Q3**	**Median**	**Q1-Q3**	**Median**	**Q1-Q3**
Age at enrollment (years)	54	45-64*	71	65-100*	36	0-100*	62	45-84*	64	50-79*
Female sex (%)	55		61		55		52		100	
European ancestry (%)	77		83		100		41		41	
BMI (Kg/m^2)	26.7	23.9-30.2	26.0	23.5-28.9	26.2	23.7-29.2	27.5	24.5-31.1	28.9	25.4-33.3
SBP (mmHg)	118	108-130	136	124-150	120	110-130	124	111-140	129	118-140
DBP (mmHg)	73	66-80	72	64-78	78	70-84	72	65-79	77	70-82
MI during follow-up (%)	9.5		12.7		10.7		3		5.8	
Age at MI during follow-up (years)	67	61-73	79	76-83	71	60-80	72	64-79	75	69-81
HF during follow-up (%)	16.1		19.7		9.3		3.5		3.1	
Age at HF (years)	71	65-77	81	77-85	82	73-88	75	68-81	72	66-77
Fatal CHD (%)	2.9		9.0		3.3		1.2		3.6	
Death during follow-up (%)	28.4		42.1		32.0		10.8		22.9	
Age at death (years)	72	67-78	82	78-87	84	76-90	78	71-84	79	72-85
Ischemic stroke (%)	5.2		11.1		6.4		2.1		5.1	
Age at ischemic stroke (years)	70	64-75	80	77-84	78	70-85	74	65-81	76	70-82
Coronary revascularization (%)	14.2		8.8		9.5		4.5		7.2	
Age at coronary revascularization (years)	67	62-73	77	73-81	70	61-77	70	63-76	73	68-78

*The range represents the eligible age for cohort enrollment as opposed to the interquartile interval.

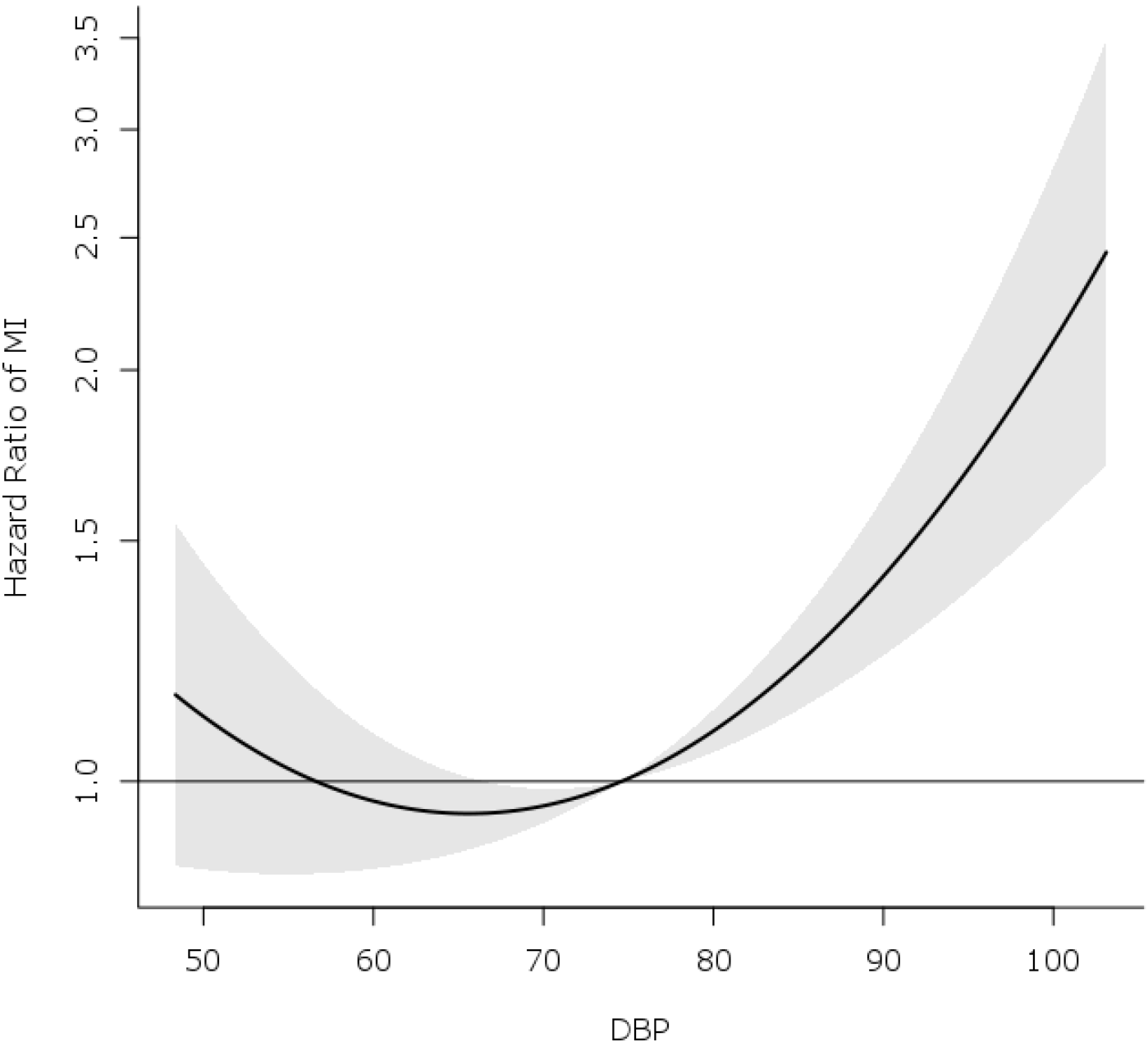
Supplemental Figures and Figure Legends

Figure I.



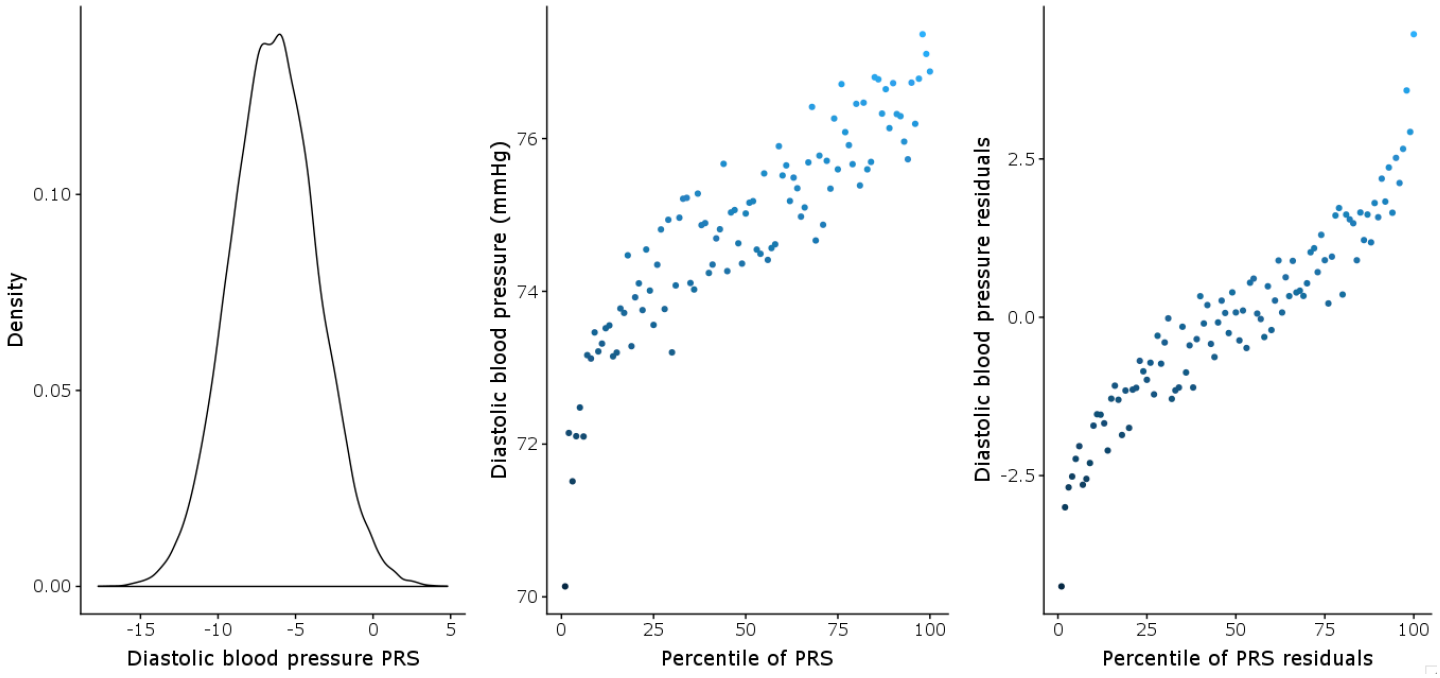
Shape of the association between diastolic blood pressure and our secondary outcomes based on observational data.

Figure II.



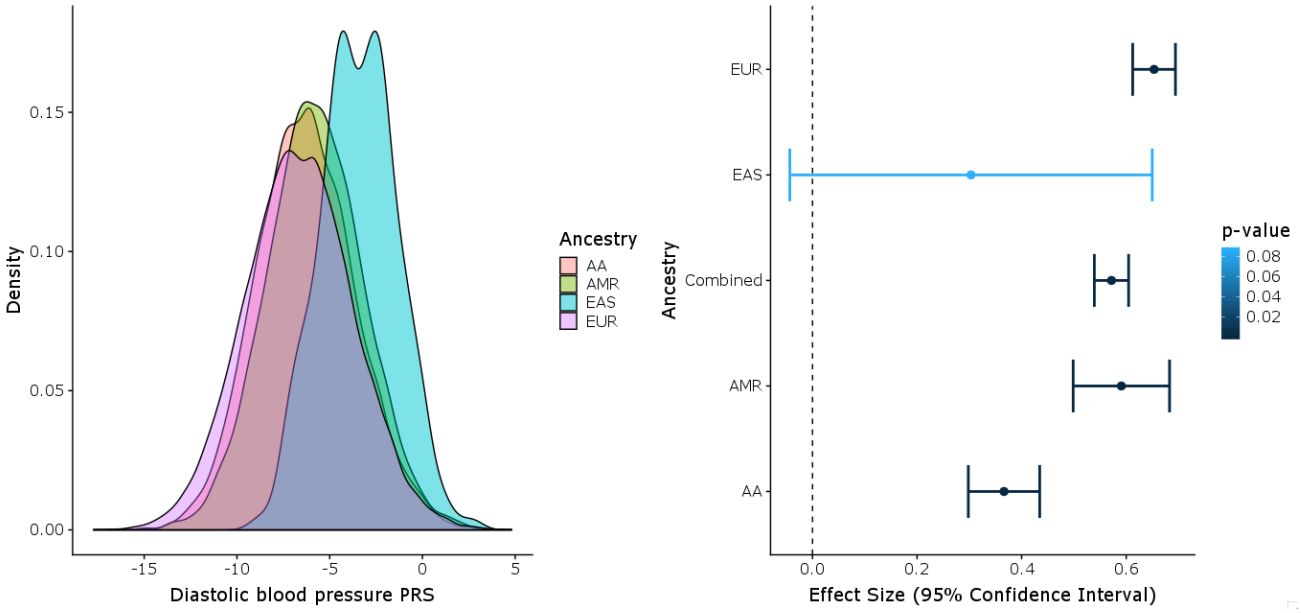
Shape of the association between DBP and MI based on observational data among participants that were on no BP meds.

Figure III.



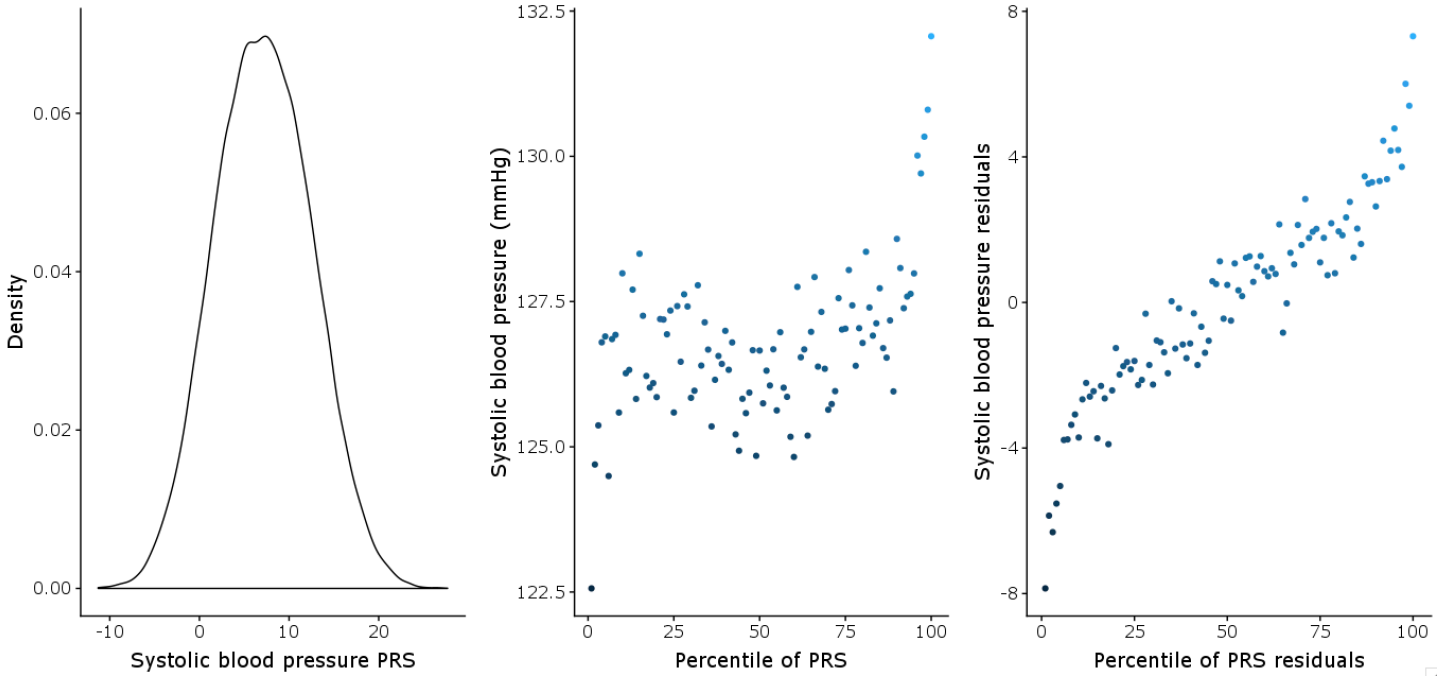
Validation of polygenic risk score for diastolic blood pressure. A. Density plot of the polygenic risk score among the individuals included in our study. B. Scatter plot of average diastolic blood pressure in each percentile of the polygenic risk score. C. Scatter plot of average diastolic blood pressure residuals in each percentile of the residuals of the PRS (residuals for both PRS and DBP are obtained by regressing out age, sex, substudy and 2 PCs for each substudy from each of the two parameters). The result of plot C is a straight line suggesting strong evidence of association between the risk score and actual blood pressure measurements which is confirmed by linear regression.

Figure IV.



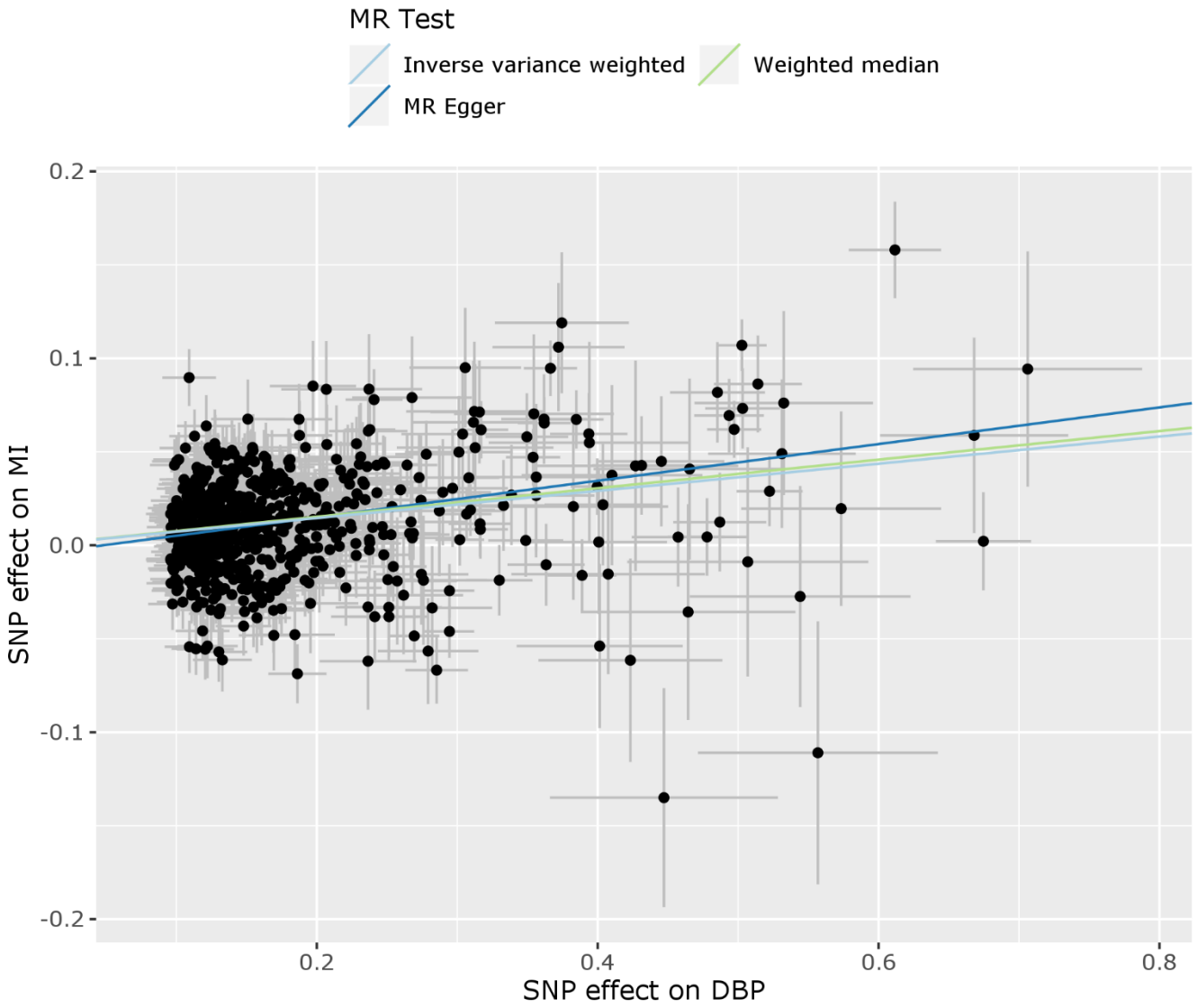
Validation of the diastolic blood pressure polygenic risk score across different ethnic groups. We see significant association between the polygenic score and diastolic blood pressure in most groups with some variability in the effect size. AA: African ancestry, AMR: Admixed American ancestry, EAS: East Asian ancestry, EUR: European ancestry.

Figure V.



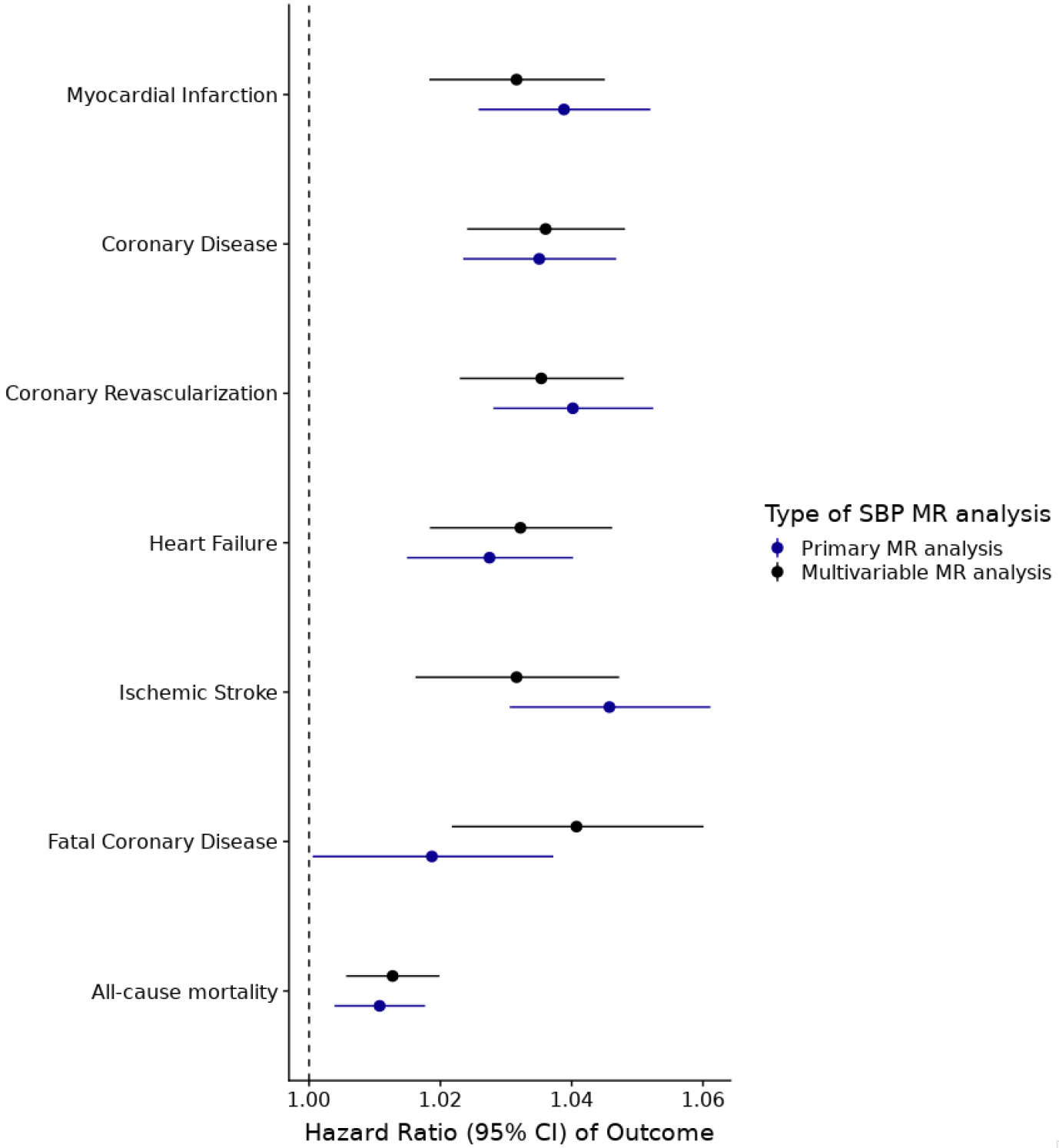
Validation of polygenic risk score for systolic blood pressure. A. Density plot of the polygenic risk score among the individuals included in our study. B. Scatter plot of average systolic blood pressure in each percentile of the polygenic risk score. C. Scatter plot of average systolic blood pressure residuals in each percentile of the residuals of the PRS (residuals for both PRS and SBP are obtained by regressing out age, sex, substudy and 2 PCs for each substudy from each of the two parameters). The result of plot C suggests strong evidence of association between the risk score and actual blood pressure measurements which is confirmed by linear regression (beta=0.612, $p<0.0001$).

Figure VI.



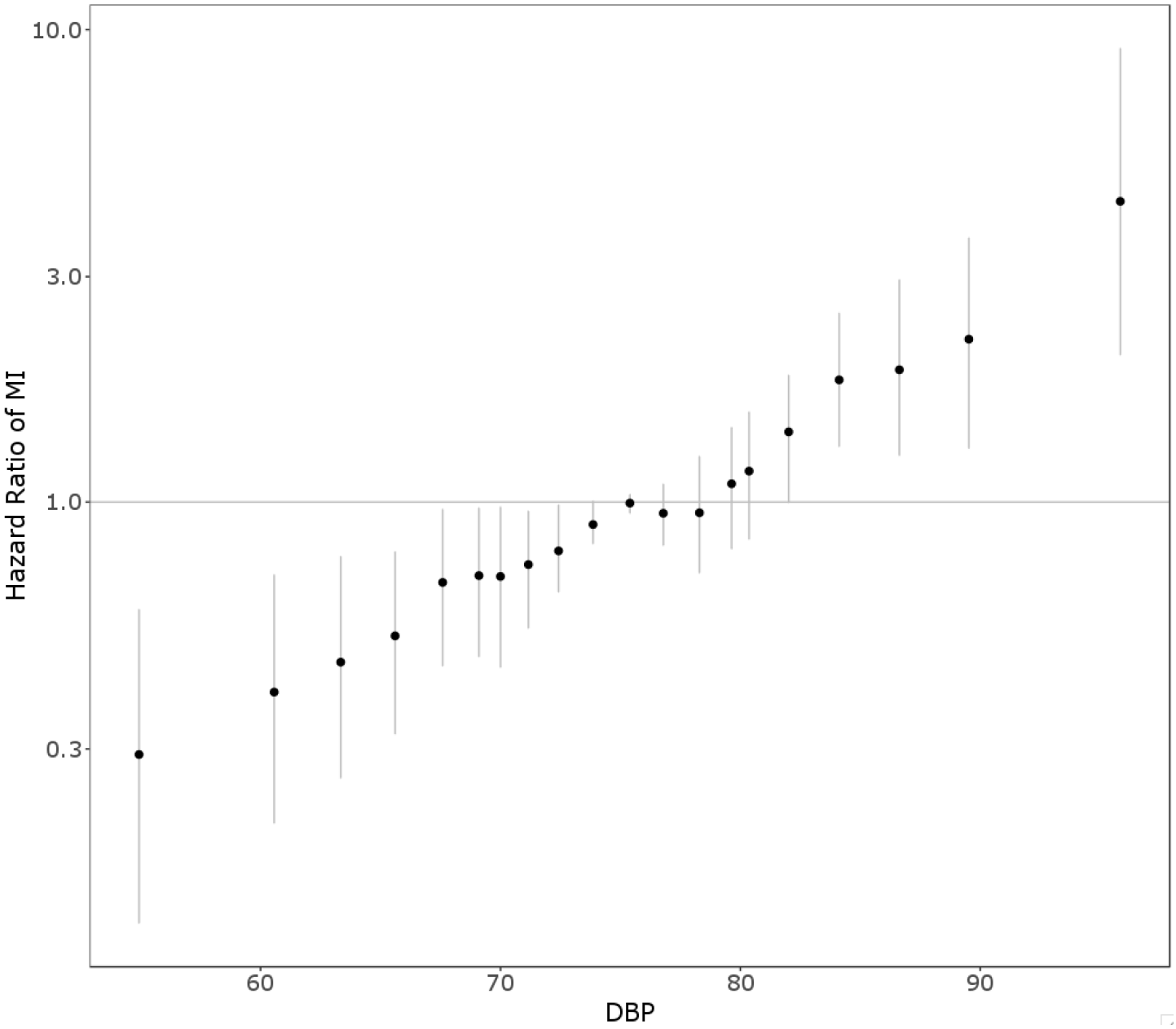
Two-sample summary statistics-based Mendelian Randomization of the effect of diastolic blood pressure on myocardial infarction. The plot presents the effect size and confidence intervals for each SNP presented in the model on diastolic blood pressure (exposure) and myocardial infarction (outcome). The slope of the plotted lines represent the mendelian randomization associations with different approaches, all of which are statistically significant.

Figure VII.



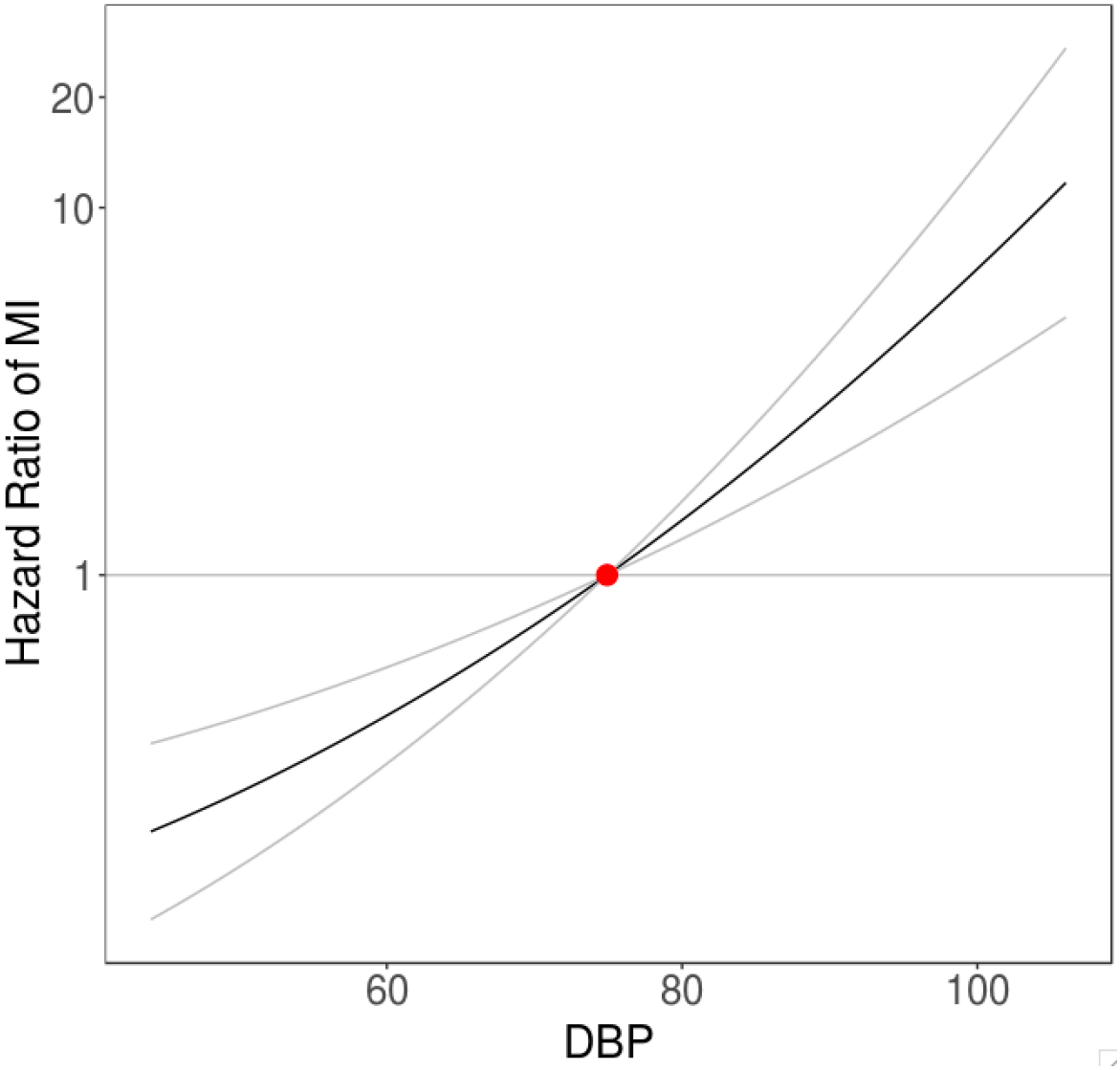
Linear association between blood pressure and cardiovascular outcomes based on MR data. The figure shows a Forest plot of the association between systolic blood pressure (in mmHg) and a series of cardiovascular events analyzed based on linear MR in the primary analysis (dark blue color), and multivariable MR analysis of the effect of systolic blood pressure after controlling for diastolic BP (black color).

Figure VIII.



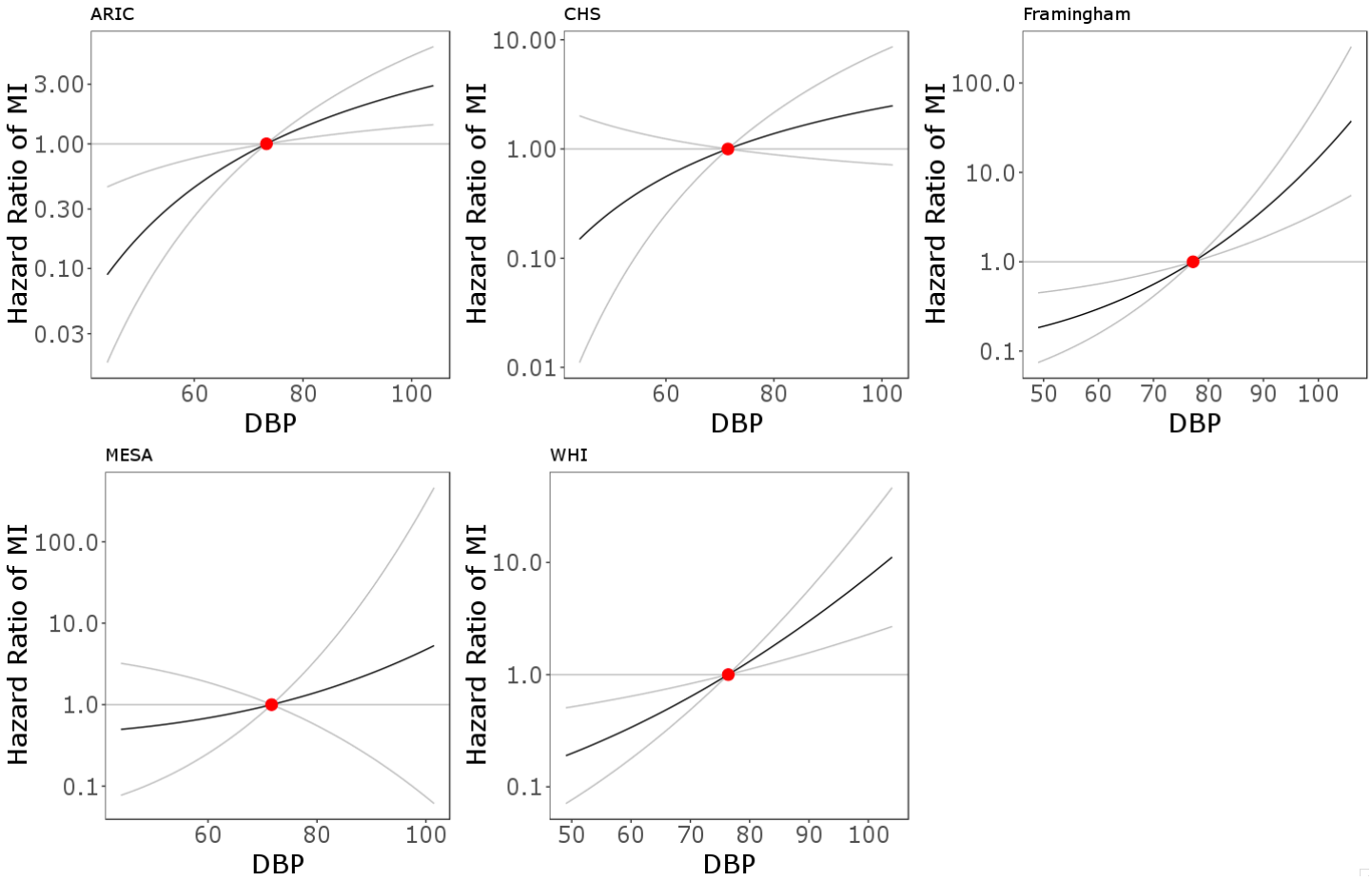
Shape of the association between DBP and MI using the piecewise non-linear Mendelian Randomization method. The exposure-outcome relationship is estimated as piecewise linear function with each stratum contributing a line segment whose gradient is the LACE estimate for that stratum. The function is constrained to be continuous, so that each line segment begins where the previous segment finished.

Figure IX.



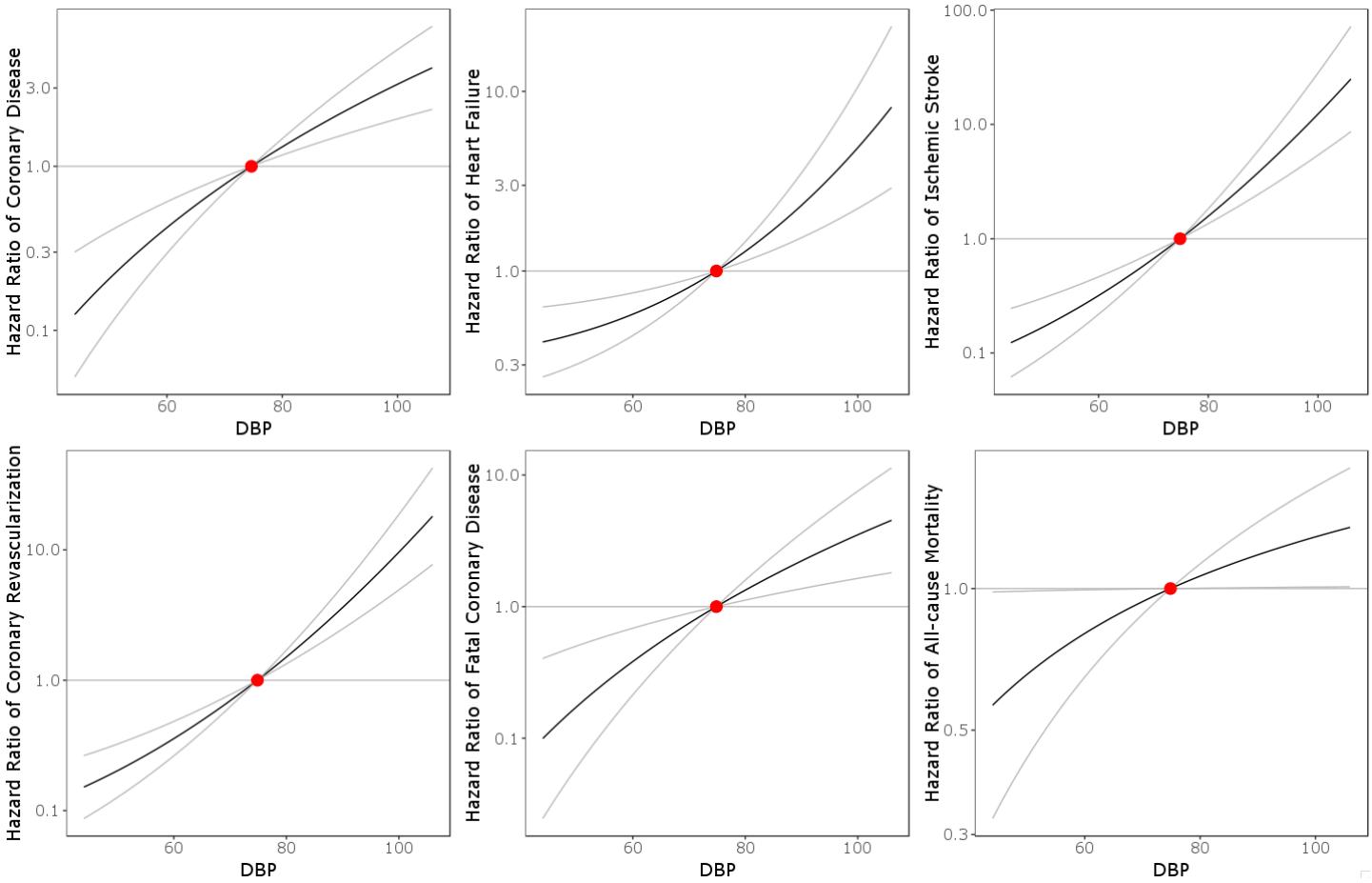
Shape of the association between DBP and MI in a sensitivity analysis using Winsorized DBP values for DBP outliers.

Figure X.



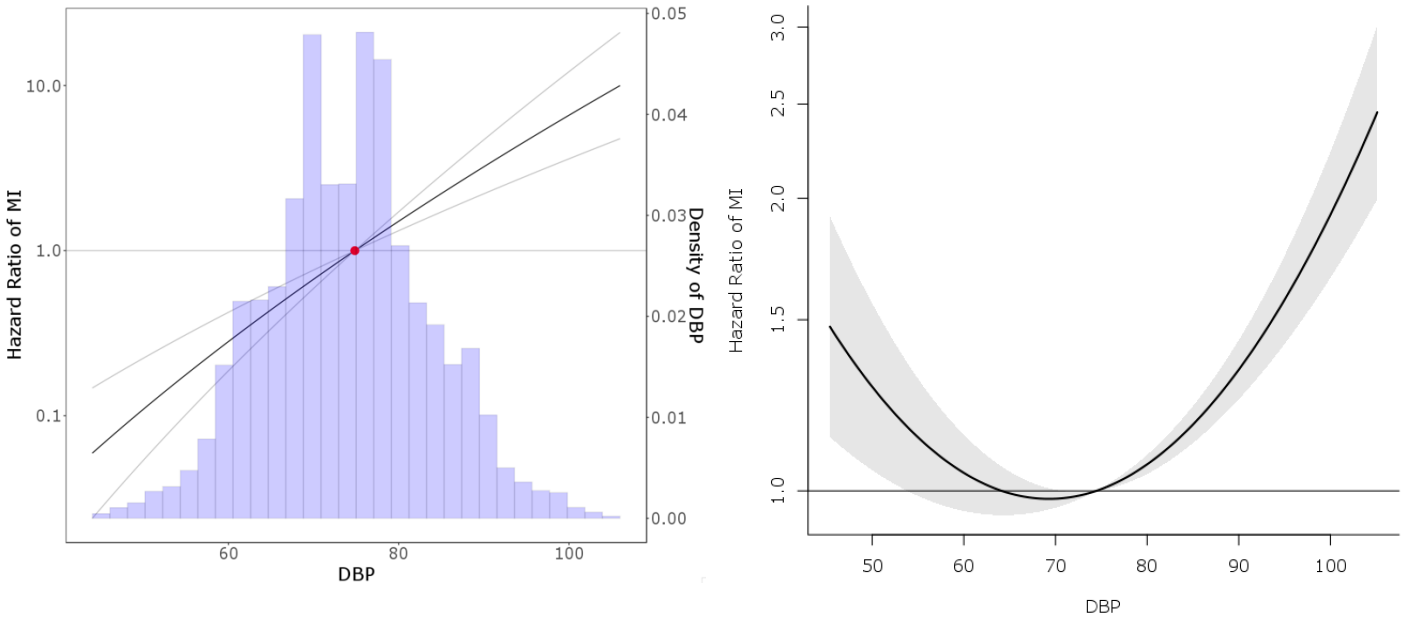
Shape of the association between diastolic blood pressure and MI based on non-linear MR in the individual cohorts

Figure XI.



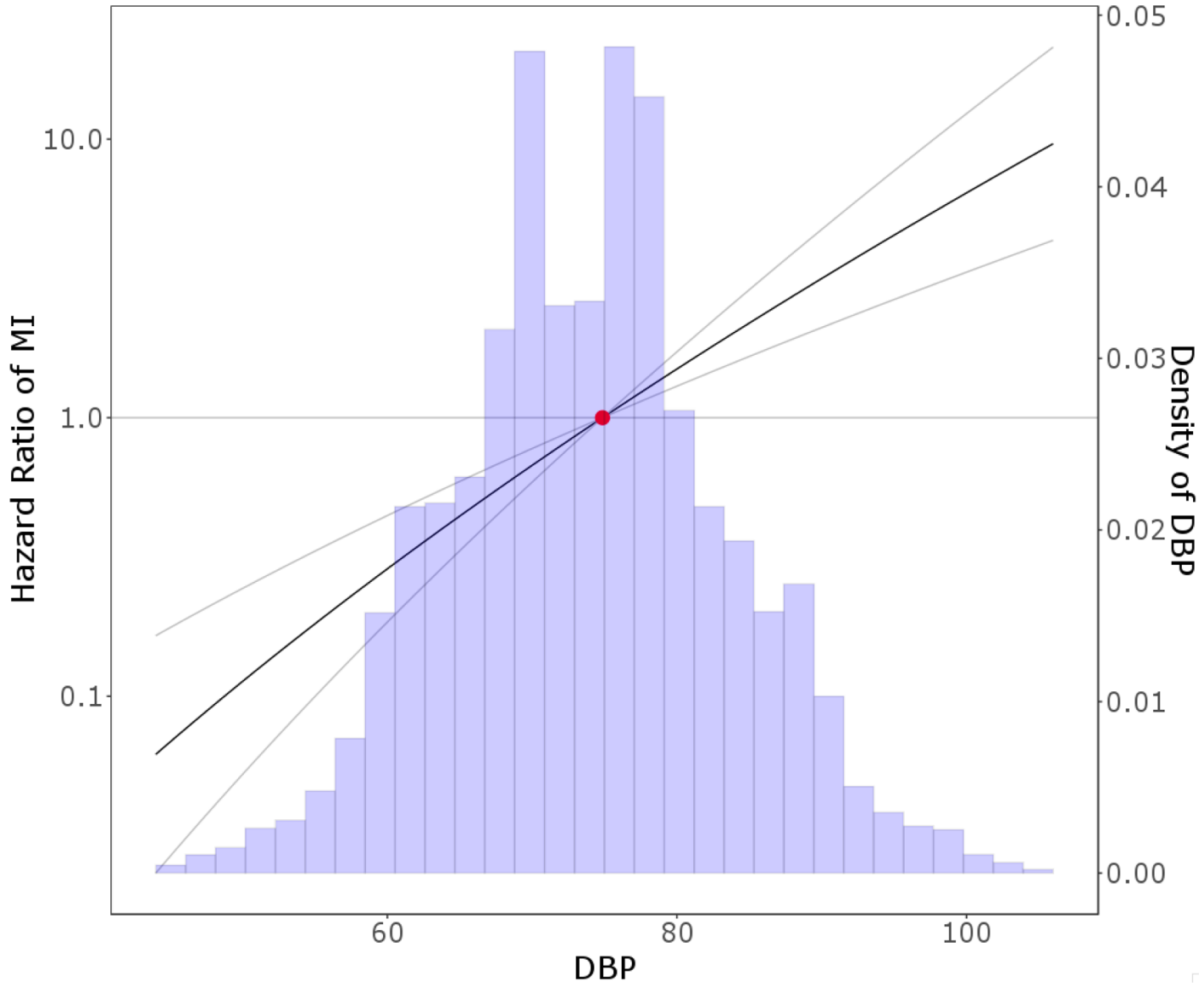
Shape of the association between diastolic blood pressure and our secondary outcomes based on non-linear MR.

Figure XII.



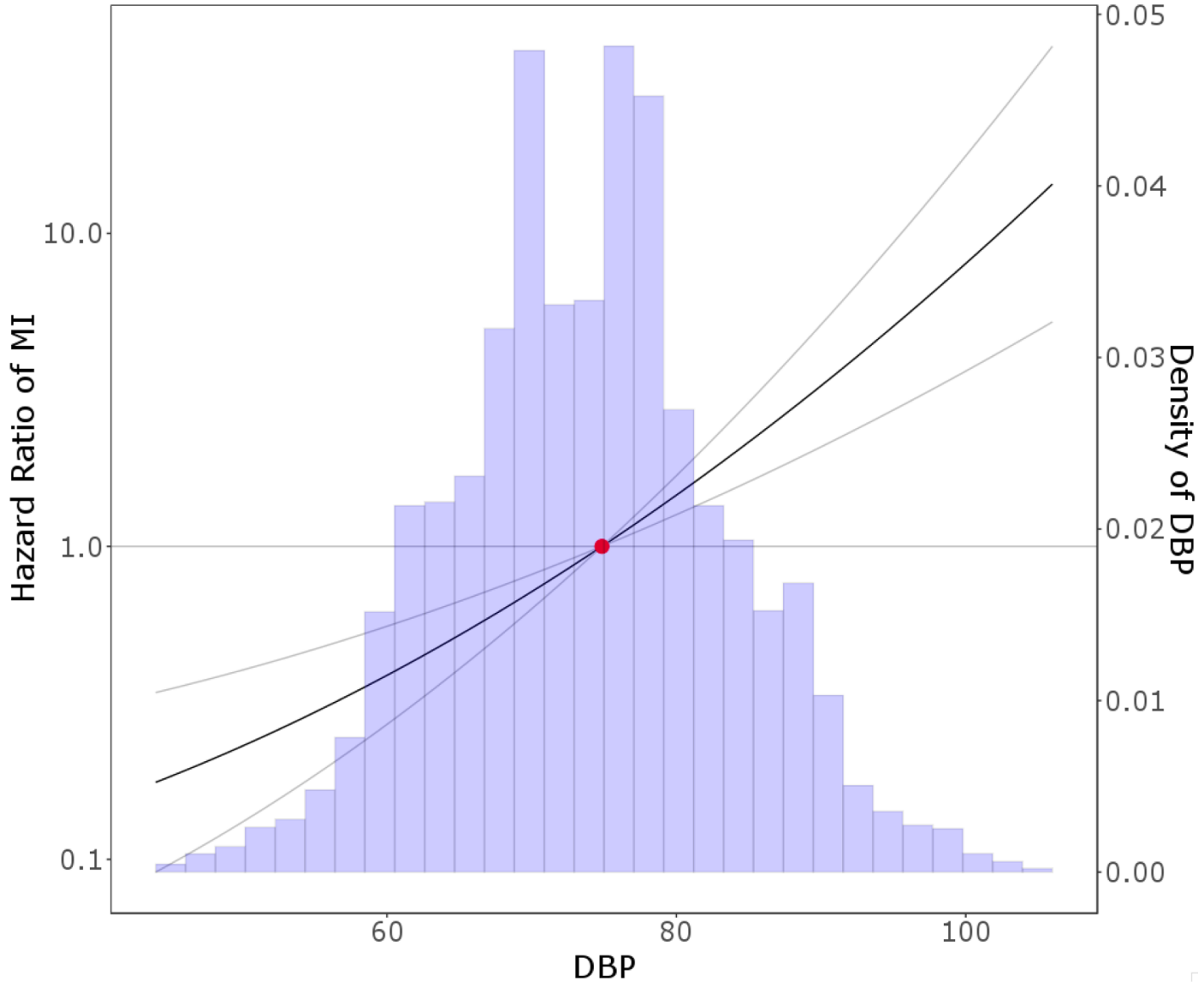
Shape of the relationship between diastolic blood pressure and MI based on MR (left figure) and observational data (right figure) in a sensitivity analysis including individuals with prevalent MI.

Figure XIII.



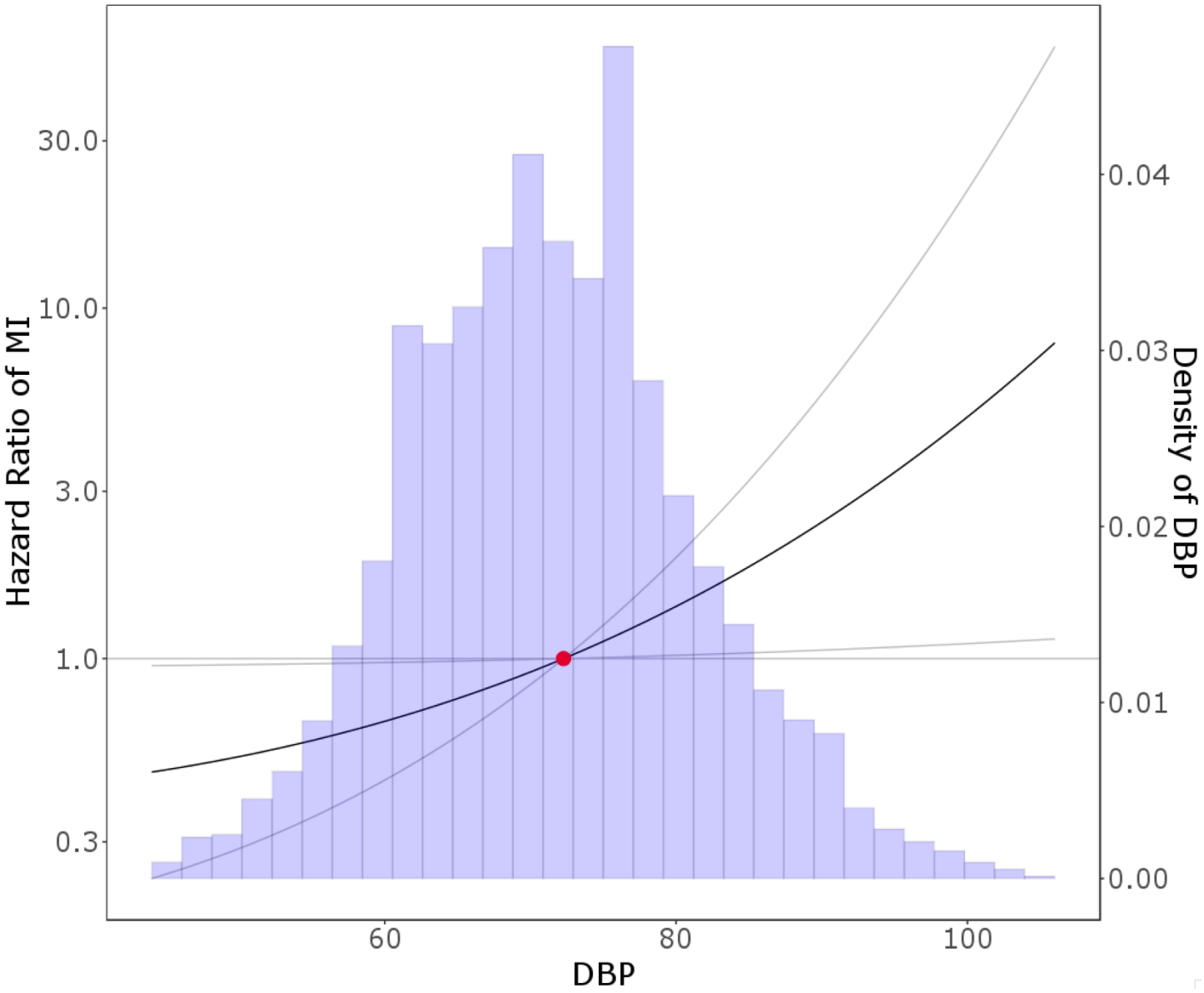
Shape of the relationship between diastolic blood pressure and MI in a sensitivity analysis using the polygenic risk score for DBP that is independent of BMI.

Figure XIV.



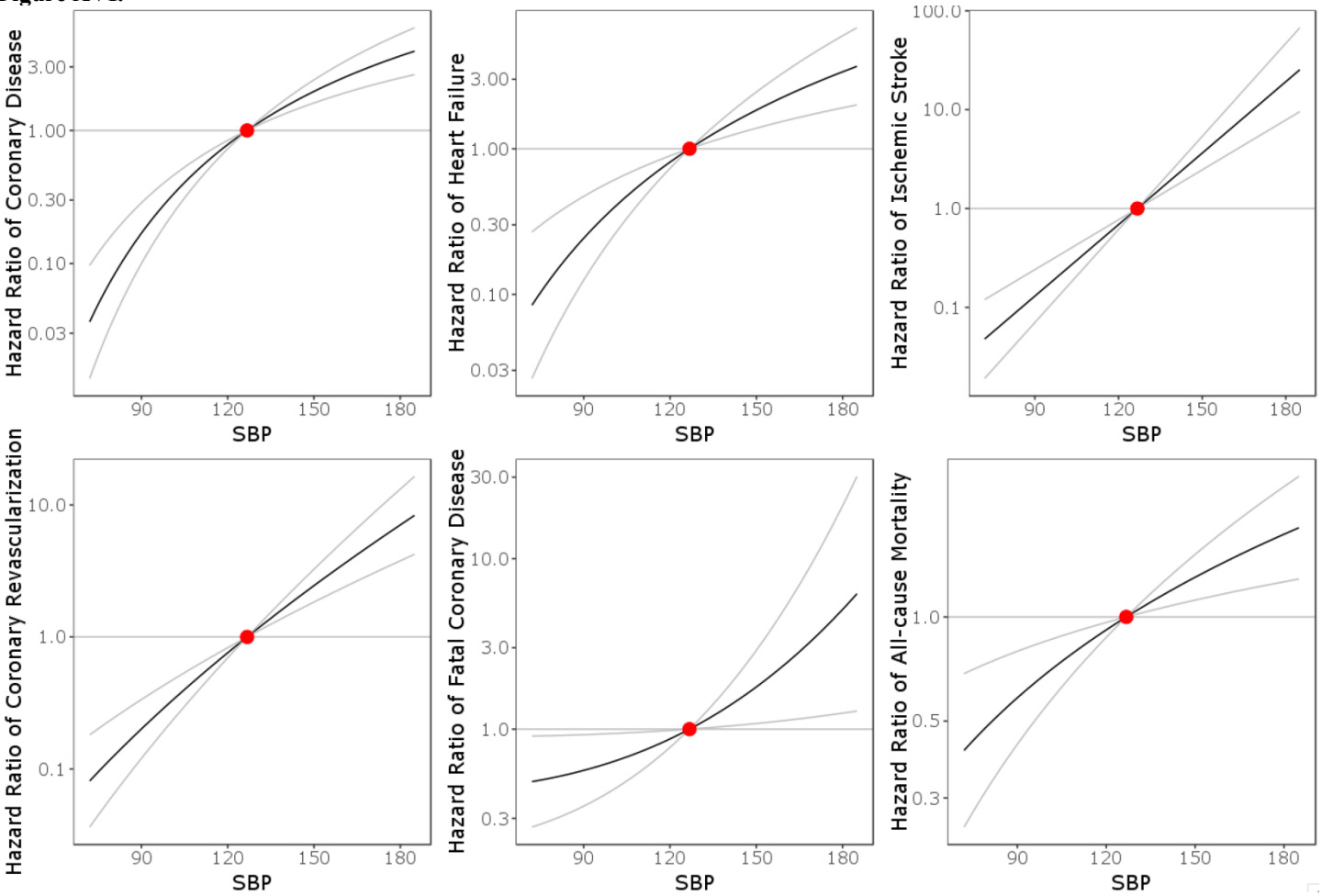
Shape of the relationship between diastolic blood pressure and MI in a sensitivity analysis using the polygenic risk score for DBP that is independent of heart rate.

Figure XV.



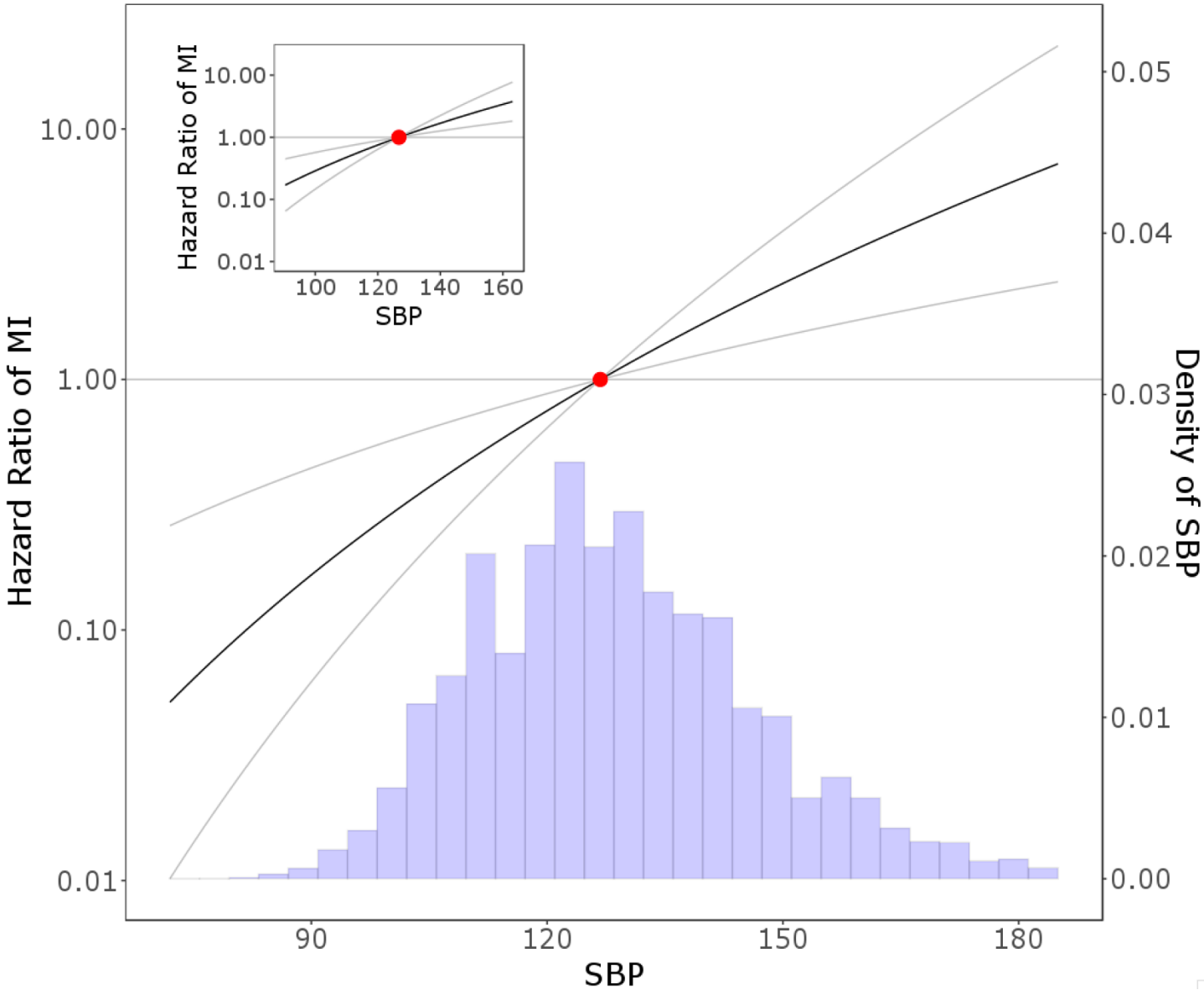
Shape of the relationship between diastolic blood pressure and MI in a sensitivity analysis in a subset of participants that have not been on blood pressure medications.

Figure XVI.



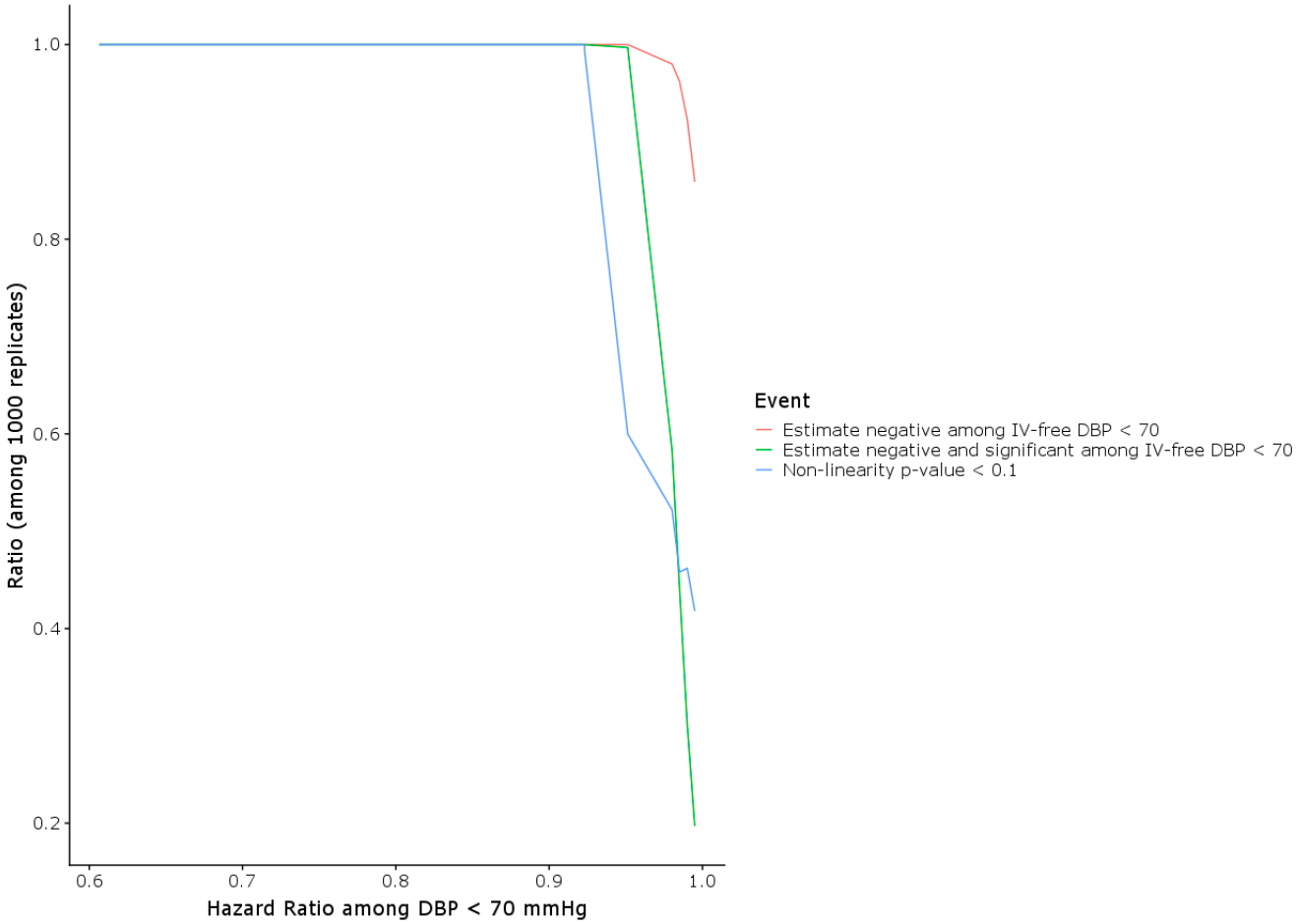
Shape of the relationship between systolic blood pressure and our secondary outcomes based on non-linear MR

Figure XVII.



Shape of the relationship between SBP and MI in an analysis of IV-free exposure quantiles accounting for both SBP and DBP linear genetic effects.

Figure XVIII.



Simulation analysis for power of non-linear effect detection