Supplementary Materials for

Multisite chronic pain as a causal risk factor for coronary artery disease: Findings from Mendelian randomization

Jiahao Zhu[a], Nini Wang[b], Houpu Liu[c], Han Jiang[a], Bingyue Cai[a], Dingwan Chen[d], Yingjun Li[c]

[a] School of Public Health, Hangzhou Medical College, Hangzhou, China, [b] Puyan Community Healthcare Center of Binjiang District, Hangzhou, China, [c] Department of Epidemiology and Health Statistics, School of Public Health, Hangzhou Medical College, Hangzhou, China, [d] Primary Health Research Center of Zhejiang Province, Hangzhou Medical College, Hangzhou, China

Supplementary Methods Description of the MR methods for exploring and adjusting for pleiotropy.

Supplementary Figure 1 Scatter plots of the primary analyses.

Supplementary Figure 2 “Leave-one-out” analysis plots of the primary analyses.

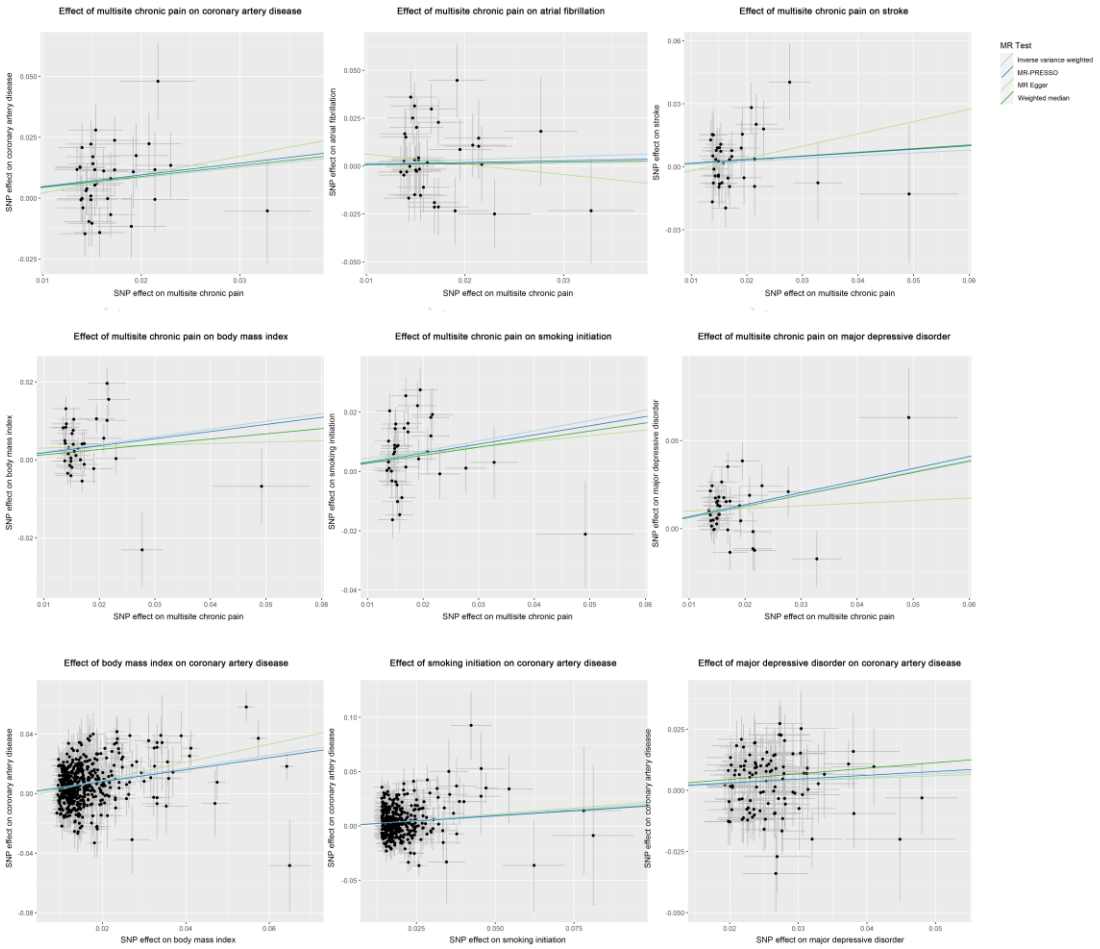
Supplementary Methods Description of the MR methods for exploring and adjusting for pleiotropy.

The weighted median estimator is defined as the median of a weighted empirical density function of the ratio estimates. This method can provide a consistent estimate of causal effect if at least 50% of instrumental variables (IVs) are valid and no single IV contributes more than 50% of the weight.

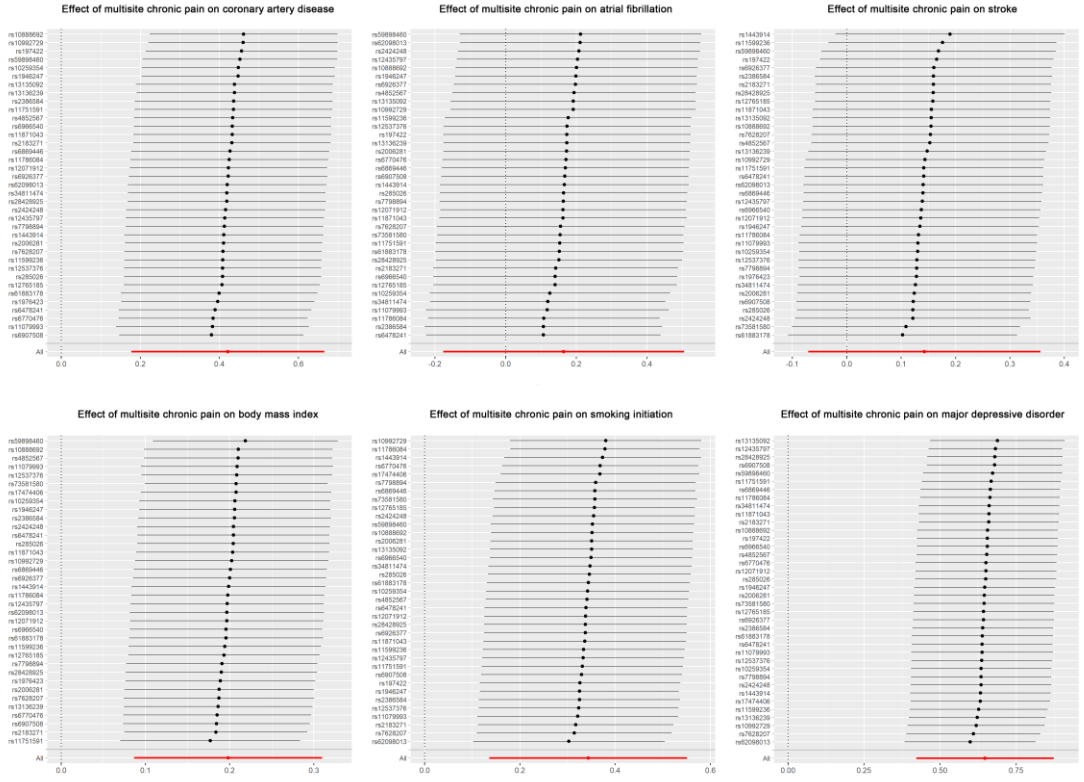
The MR-Egger regression can test for directional (unbalanced) pleiotropy and provide an estimate of the causal effect whilst taking pleiotropic effects into account. It requires the InSIDE (instrument strength independent of direct effect) assumption to hold, which means the strength of the genetic variant-exposure association should not correlate with the strength of bias due to pleiotropy.

The MR-PRESSO (Mendelian randomization pleiotropy residual sum and outlier), can detect horizontal pleiotropy, adjust for horizontal pleiotropy via outlier removal, and test the significant differences in the causal estimates before and after removal of outliers. The MR-PRESSO outlier test requires that at least 50% of the variants are valid instruments and relies on the InSIDE assumption.

Supplementary Figure 1 Scatter plots of the primary analyses.



Supplementary Figure 2 “Leave-one-out” analysis plots of the primary analyses. “Leave-one-out” analysis plots of the effects of three studied mediators on coronary artery disease were not shown, because the number of SNPs associated with each mediator is too large to be clearly presented on a plot.



STROBE-MR checklist of recommended items to address in reports of Mendelian randomization studies[1 2]

Item No.	Section	Checklist item	Page No.	Relevant text from manuscript
1	**TITLE and ABSTRACT**	Indicate Mendelian randomization (MR) as the study's design in the title and/or the abstract if that is a main purpose of the study	Title and abstract	Multisite chronic pain as a causal risk factor for coronary artery disease: Findings from Mendelian randomization.
	INTRODUCTION			
2	**Background**	Explain the scientific background and rationale for the reported study. What is the exposure? Is a potential causal relationship between exposure and outcome plausible? Justify why MR is a helpful method to address the study question	1-2	Accumulating evidence reported that chronic pain conditions are co-morbid with a wide range of adverse health outcomes such as cardiovascular diseases (CVDs). Mendelian randomization (MR) offers an alternative way to address these methodological challenges.
3	**Objectives**	State specific objectives clearly, including pre-specified causal hypotheses (if any). State that MR is a method that, under specific assumptions, intends to estimate causal effects	1-2	We hypothesized that multisite chronic pain would increase the risk of CVDs potentially through mediated effect of BMI, smoking and/or depression. MR utilizes genetic variants as instruments to estimate the causal effect of the exposure on the outcome.
	METHODS			
4	**Study design and data sources**	Present key elements of the study design early in the article. Consider including a table listing sources of data for all phases of the study. For each data source contributing to the analysis, describe the following:	3	Summary statistics for each respective phenotype of interest were obtain from genome-wide association studies (GWASs) (Table 1).
	a)	Setting: Describe the study design and the underlying population, if possible. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection, when available.		Not applicable since this is a two-sample MR study based on summary-level data.
	b)	Participants: Give the eligibility criteria, and the sources and methods of selection of participants. Report the sample size, and whether any power or sample size calculations were carried out prior to the main analysis		Not applicable.
	c)	Describe measurement, quality control and selection of genetic variants		Not applicable.
	d)	For each exposure, outcome, and other relevant variables, describe methods of assessment and diagnostic criteria for diseases		Not applicable.
	e)	Provide details of ethics committee approval and participant informed consent, if relevant	2	The relevant ethical approval and participant consent has been obtained in original research.
5	**Assumptions**	Explicitly state the three core IV assumptions for the main analysis (relevance, independence and exclusion restriction) as well assumptions for any additional or sensitivity analysis	2	See Figure 1 Legend and Supplementary Methods.
6	**Statistical methods: main analysis**	Describe statistical methods and statistics used		

		a) Describe how quantitative variables were handled in the analyses (i.e., scale, units, model)	3	Multisite chronic pain is a quantitative phenotype defined as the sum of self-reported pain lasting at least three months at seven different body areas (i.e., head, face, neck/shoulder, back, stomach/ abdomen, hip, and knee).
		b) Describe how genetic variants were handled in the analyses and, if applicable, how their weights were selected	4	Independent single-nucleotide polymorphisms (SNPs) associated with each considered exposure, mediator, and outcome at genome-wide significance (P<5e-8) were selected as genetic instruments from the original literature.
		c) Describe the MR estimator (e.g. two-stage least squares, Wald ratio) and related statistics. Detail the included covariates and, in case of two-sample MR, whether the same covariate set was used for adjustment in the two samples	3	See Table 1
		d) Explain how missing data were addressed		Not applicable.
		e) If applicable, indicate how multiple testing was addressed	5	As we tested three types of CVDs as well as three mediators, a two-sided P<0.05/3=0.017 corrected by Bonferroni method was applied to indicate significant associations in the primary analysis.
7	**Assessment of assumptions**	Describe any methods or prior knowledge used to assess the assumptions or justify their validity	4	F statistic and horizontal pleiotropy tests.
8	**Sensitivity analyses and additional analyses**	Describe any sensitivity analyses or additional analyses performed (e.g. comparison of effect estimates from different approaches, independent replication, bias analytic techniques, validation of instruments, simulations)	4-5	We therefore did additional pleiotropy-robust MR as sensitivity analyses, including weighted median, MR-Egger regression, and MR-PRESSO. In addition to multisite chronic pain as the main considered exposure, we evaluated the chronic widespread pain, another important pain phenotype that may represent the extreme of a pain state.
9	**Software and pre-registration**			
		a) Name statistical software and package(s), including version and settings used	6	All statistical analyses were performed using R (v3.6.3) with "TwoSampleMR" (v0.5.6), "MendelianRandomization" (v0.5.1), and "MRPRESSO" (v1.0) packages.
		b) State whether the study protocol and details were pre-registered (as well as when and where)		Not applicable.
	RESULTS			
10	**Descriptive data**			
		a) Report the numbers of individuals at each stage of included studies and reasons for exclusion. Consider use of a flow diagram		Not applicable.
		b) Report summary statistics for phenotypic exposure(s), outcome(s), and other relevant variables (e.g. means, SDs, proportions)	6	See Table 1 and the MR estimates represent the change in the outcome per one number increase of pain areas in multisite chronic pain, per one standard deviation (4.8 kg/m2) unit increase in BMI, and per one unit increase in log odds scale of the binary exposures (e.g., major depressive disorder).
		c) If the data sources include meta-analyses of previous studies, provide the assessments of heterogeneity across these studies		The heterogeneity was mainly assessed using the Cochran's Q statistic.
		d) For two-sample MR: i. Provide justification of the similarity of the genetic variant-exposure associations between the exposure and outcome samples	3	The majority of participants included in GWAS meta-analyses were of European ancestry.

			ii. Provide information on the number of individuals who overlap between the exposure and outcome studies	3	To avoid potential bias related to sample overlap, we selected GWAS summary data for three types of CVDs with UK Biobank cohort removed, as the chronic pain data is solely composed of the UK Biobank.
11	**Main results**				
		a)	Report the associations between genetic variant and exposure, and between genetic variant and outcome, preferably on an interpretable scale	6	See Supplementary Table 1-3.
		b)	Report MR estimates of the relationship between exposure and outcome, and the measures of uncertainty from the MR analysis, on an interpretable scale, such as odds ratio or relative risk per SD difference	7-8	See Table 2-4.
		c)	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period		Not applicable.
		d)	Consider plots to visualize results (e.g. forest plot, scatterplot of associations between genetic variants and outcome versus between genetic variants and exposure)	7	See Supplementary Figure 1.
12	**Assessment of assumptions**				
		a)	Report the assessment of the validity of the assumptions	7	Although the MR-Egger intercept did not report clear evidence for the presence of directional pleiotropy (except for the analysis of BMI and coronary artery disease), the MR-PRESSO global test indicated that horizontal pleiotropy is pervasive in the current MR study.
		b)	Report any additional statistics (e.g., assessments of heterogeneity across genetic variants, such as I^2, Q statistic or E-value)	7	We observed significant heterogeneity in most MR analyses as indicated by Cochran's Q statistic.
13	**Sensitivity analyses and additional analyses**				
		a)	Report any sensitivity analyses to assess the robustness of the main results to violations of the assumptions	7	Nevertheless, pleiotropy-robust MR approaches basically did not change the direction and magnitude of point estimates, except for MR-Egger regression that produced causal estimates with reduced precision (wide CIs).
		b)	Report results from other sensitivity analyses or additional analyses	7-8	In the "leave-one-out" analysis, no individual SNP was observed to drive the overall results. Using chronic widespread pain as the secondary exposure, we found a similar pattern of the primary results.
		c)	Report any assessment of direction of causal relationship (e.g., bidirectional MR)	7-8	In reverse MR analyses, there was little evidence for effects of genetic liability to CVDs on multisite chronic pain, BMI, smoking, and major depressive disorder (Supplementary Table 5).
		d)	When relevant, report and compare with estimates from non-MR analyses		Not applicable.
		e)	Consider additional plots to visualize results (e.g., leave-one-out analyses)	7	See Supplementary Figure 2.
	DISCUSSION				
14	**Key results**		Summarize key results with reference to study objectives	8	We found robust evidence that multisite chronic pain is a causative factor for coronary artery disease.
15	**Limitations**		Discuss limitations of the study, taking into account the validity of the IV assumptions, other sources of potential bias, and imprecision. Discuss both direction and magnitude of any potential bias and any efforts to address them	11	There are also several limitations to this work.

16	**Interpretation**				
		a)	Meaning: Give a cautious overall interpretation of results in the context of their limitations and in comparison with other studies	8-9	Our results are consistent with earlier MR studies.
		b)	Mechanism: Discuss underlying biological mechanisms that could drive a potential causal relationship between the investigated exposure and the outcome, and whether the gene-environment equivalence assumption is reasonable. Use causal language carefully, clarifying that IV estimates may provide causal effects only under certain assumptions	8-9	These findings together provided supportive evidence for BMI, smoking, and depression as causal mediators (biological mechanisms) from greater number of chronic pain sites to higher risk of developing coronary artery disease. Therefore, our MR results may be more suitable for indicating whether the null hypothesis is supported, rather than estimating the average causal effects.
		c)	Clinical relevance: Discuss whether the results have clinical or public policy relevance, and to what extent they inform effect sizes of possible interventions	10	Our findings could have implications for future public health and clinical practice.
17	**Generalizability**		Discuss the generalizability of the study results (a) to other populations, (b) across other exposure periods/timings, and (c) across other levels of exposure	12	Lastly, this MR investigation was based on data drawn from subjects of predominantly European ancestry, thus limiting the extrapolation of our findings to other ethnic populations.
	OTHER INFORMATION				
18	**Funding**		Describe sources of funding and the role of funders in the present study and, if applicable, sources of funding for the databases and original study or studies on which the present study is based	12	No specific funding was received from any bodies in the public, commercial or not-for-profit sectors to carry out the work described in this article.
19	**Data and data sharing**		Provide the data used to perform all analyses or report where and how the data can be accessed, and reference these sources in the article. Provide the statistical code needed to reproduce the results in the article, or report whether the code is publicly accessible and if so, where	12	GWAS summary data used in the main analyses was provided in Supplementary Table 1-3.
20	**Conflicts of Interest**		All authors should declare all potential conflicts of interest	12	None.

1. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomization (STROBE-MR) Statement. JAMA. 2021;under review.

2. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomisation (STROBE-MR): Explanation and Elaboration. BMJ. 2021;375:n2233.