pIVW: A novel Mendelian Randomization Method Accounting for Weak Instruments and Horizontal Pleiotropy with Applications to the COVID-19 Outcomes

Siqi Xu[1], Wing Kam Fung*[1], and Zhonghua Liu†[1]

[1]Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong, China

Abstract

Mendelian randomization (MR) utilizes genetic variants as instrumental variables (IVs) to estimate the causal effect of an exposure variable on an outcome of interest even in the presence of unmeasured confounders. However, many MR methods including the most popular inverse-variance weighted (IVW) estimator could be biased by the weak IVs that are weakly associated with the exposure. In this article, we develop a novel method called penalized inverse-variance weighted (pIVW) estimator, where we adjust the IVW estimator to account for the weak IVs by a proposed penalization method to prevent the denominator of the pIVW estimator from being close to zero. Moreover, we account for the horizontal pleiotropy—a widespread phenomenon in human genome that could bias the inference for the causal effect—by adjusting the variance estimation of the pIVW estimator. The proposed pIVW estimator can reduce to the debiased IVW (dIVW) estimator—another extension of the the IVW estimator—when the number of

*Correspondence: wingfung@hku.hk
†Correspondence: zhhliu@hku.hk

IVs and the IV strength increase. More generally, we prove that the pIVW estimator can achieve smaller bias and variance than the dIVW estimator under some regularity conditions. We also illustrate the improved performance of the proposed pIVW estimator over competing MR methods through a comprehensive simulation study. Further, we analyze the causal effects of the obesity-related traits and diseases on the Coronavirus disease 2019 (COVID-19). Notably, we find that hypertensive disease is associated with increased risk of hospitalized COVID-19, while peripheral vascular disease and higher body mass index are associated with increased risks of COVID-19 infection, hospitalized COVID-19 and critically ill COVID-19. The R package for the pIVW method is publicly available at https://github.com/siqixu/mr.pivw.

1 Introduction

It is of major interest in health studies to identify causal risk factors associated with various clinical outcomes. For example, identifying the causal risk factors of the Coronavirus disease 2019 (COVID-19) is currently one of the most pressing global public health problems (Jordan et al., 2020; Zheng et al., 2020). The COVID-19 pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has posed a serious threat to human health all over the world (Pascarella et al., 2020). As of July 1, 2021, the COVID-19 pandemic has led to over 180 million cases and more than 3.9 million deaths, which has brought unprecedented medical and economic burdens worldwide (Dong et al., 2020). To reduce the risks of COVID-19 incidence and mortality, it is crucial to identify the causal risk factors for the development of public health policies and clinical strategies for prevention and intervention. So far, associations between some exposure variables such as obesity and the COVID-19 outcomes have been reported by several epidemiological studies (Popkin et al., 2020; Stefan et al., 2020). However, the associations identified from the observational data might be subject to unmeasured confounding of the exposure-outcome relationship.

To address the unmeasured confounding issue in observational studies, Mendelian randomization (MR) utilizes genetic variants as instrumental variables (IVs) to estimate the causal effect of an exposure variable on an outcome of interest even in the presence of unmeasured confounders (Smith and Ebrahim, 2003, 2004; Sheehan et al., 2008). Due to the availability of summary-level data from the genome-wide association studies (GWASs), a plenty of methods were developed for GWAS summary-level data using multiple genetic variants as IVs (Zheng et al., 2017; Lawlor, 2016), which facilitate the wide use of MR analysis in health studies. However, the validity of MR strictly depends on the following three core assumptions defining a valid IV (Lawlor et al., 2008; Didelez and Sheehan, 2007):

- IV Relevance: the IV is associated with the exposure;
- IV Independence: the IV is independent of any confounders of the exposure-outcome relationship;
- Exclusion Restriction: the IV only affects the outcome via the exposure.

When either of these IV assumptions is violated, the MR analysis may also yield biased estimation of the causal effect. In particular, the first IV assumption (IV Relevance) can be nearly violated when the IVs are only weakly associated with the exposure variable (Burgess and Thompson, 2011; Burgess et al., 2011; Davies et al., 2015). In MR studies, the weak IV bias may occur when the genetic variants only explain a small proportion of variance for the exposure variable. On the other hand, the widespread horizontal pleiotropy in human genome can also lead to the violation of the third IV assumption (Exclusion Restriction) (Verbanck et al., 2018; Hemani et al., 2018), which is a phenomenon that the genetic variants directly affect the outcome not mediated by the exposure variable (see Figure 1 for a graphical illustration).

The inverse-variance weighted (IVW) estimator (Burgess et al., 2013) is the most popular MR method being widely used in health studies. It has a simple and explicit expression, which combines the estimated causal effects from multiple IVs into a weighted average with the idea borrowing from the fixed-effect meta-analysis literature. Despite its widespread popularity, studies (Zhao et al., 2020; Ye et al., 2019) pointed out that the IVW estimator can be seriously biased by the weak IVs. Besides the IVW estimator, many common MR methods were found to be subject to the weak IV bias, such as MR-Egger (Bowden et al., 2015) which estimates the causal effect by a weighted linear regression, and MR-Median (Bowden et al., 2016) which is a weighted median estimator of the estimated causal effects from each of multiple IVs. MR-RAPS (Zhao et al., 2020) is a maximum profile likelihood estimator, which was shown to be robust to weak IVs. However, in contrast to the IVW estimator, MR-RAPS has no closed form of solution and might have multiple roots. Recently, the debiased IVW (dIVW) estimator (Ye et al., 2019) was proposed to account for the weak IV by a simple modification to the IVW estimator. Comparing to the IVW estimator, the dIVW estimator was proved to be consistent under weaker conditions that allow the presence of many weak IVs. Nevertheless, as a ratio estimator, the dIVW estimator is likely to yield an inflated estimate when its denominator is close to zero, which might occur in the presence of many weak IVs. Moreover, when the denominator is close to zero, a ratios estimator may have heavy tailed distributions and may not possess finite moments, such as the ratio of two normal random variables (Piegorsch and Casella, 1985; Marsaglia et al., 2006; Press, 1969). In the economic literature, many IV methods as ratio estimators also encounter similar issues especially under the the weak IV setting (Nelson and Startz, 1990; Andrews et al., 2019).

In this article, we develop a novel method called penalized inverse-variance weighted (pIVW) estimator, where the IVW estimator is further adjusted by a proposed penalized log-likelihood function. Through the penalization, we prevent the denominator in the ratio estimator from being close to zero and thus provide improved causal estimation. Moreover, we account for the horizontal pleiotropy by adjusting the variance estimation of the pIVW estimator. The proposed pIVW estimator have some nice features. First, our theoretical and numerical results show that the proposed pIVW estimator can achieve smaller bias and variance than the dIVW estimator under some regularity conditions. Second, it is consistent and asymptotically normal even in the presence of many weak IVs, and requires no further assumptions than the dIVW estimator. Third, it has a unique and explicit form of solution, whereas some other robust MR methods like MR-RAPS might have multiple roots and diverging solutions. We also illustrate the improved performance of the proposed pIVW estimator over the competing MR methods by a comprehensive simulation study. Furthermore, we focus our interest on the causal effects of the obesity-related exposures on the COVID-19 outcomes. Some recent epidemiological studies have found that obesity and some obesity-related diseases (e.g., hypertension) were associated with increased risks of COVID-19 infection and severity (Popkin et al., 2020; Nakeshbandi et al., 2020; Klang et al., 2020; Zhang et al., 2020; de Almeida-Pititto et al., 2020), however, their associations might arise from the unobserved confounders rather than the causality. Although some

MR studies have also found associations between body mass index (BMI) on COVID-19 outcomes (Leong et al., 2021; Ponsford et al., 2020; Aung et al., 2020; Freuer et al., 2021), but the associations between the obesity-related diseases and COVID-19 outcomes have seldom been studied. Moreover, current MR analyses mostly relied on the traditional MR methods such as the IVW estimator that may suffer from the weak IV bias. In particular, the weak IVs are likely to exist when the exposure variables are some complex traits and diseases. Therefore, we aim to provide more robust estimation by utilizing the proposed pIVW estimator to account for the weak IVs. Specifically, we apply the pIVW estimator to analyzing the causal effects of five obesity-related traits and diseases (i.e., peripheral vascular disease, dyslipidemia, hypertensive disease, type 2 diabetes and BMI) on three COVID-19 outcomes (i.e., COVID-19 infection, hospitalized COVID-19 and critically ill COVID-19). Among them, we find that hypertensive disease is significantly associated with increased risk of hospitalized COVID-19, while peripheral vascular disease and higher BMI are significantly associated with increased risks of COVID-19 infection, hospitalized COVID-19 and critically ill COVID-19.

2 Preliminary

2.1 Linear Structural Models

Suppose that there are p independent genetic variants $\{G_j\}_{j=1}^{p}$. When there exists no horizontal pleiotropy, the relationships among the genetic variants G_js, the exposure X, the outcome Y and the unmeasured confounder U depicted in Figure 1 can be formulated by the linear structural models as follows (Bowden et al., 2015):

$$X = \sum_{j=1}^{p} \gamma_j G_j + U + \epsilon_X, \tag{1}$$

$$Y = \beta X + U + \epsilon_Y, \tag{2}$$

where γ_j is the genetic effect of G_j on X, β is the causal effect of our interest, and ϵ_X and ϵ_Y are mutually independent random errors. Let Γ_j denotes the effect of G_j on Y, then we have $\Gamma_j = \beta\gamma_j$ by substituting Equation (1) for X in Equation (2). When Y is a binary outcome and Equation (2) is replaced by a logistic model, $\Gamma_j = \beta\gamma_j$ can still be justified by a probit approximation of the logistic model (Zhao et al., 2020; Vansteelandt et al., 2011).

Let $\hat{\gamma}_j$ and $\hat{\Gamma}_j$ be the estimates of γ_j and Γ_j with the variances $\sigma^2_{\hat{\gamma}_j}$ and $\sigma^2_{\hat{\Gamma}_j}$, respectively. In the two-sample MR design (Lawlor, 2016), $\{\hat{\gamma}_j, \sigma_{\hat{\gamma}_j}\}_{j=1}^p$ and $\left\{\hat{\Gamma}_j, \sigma_{\hat{\Gamma}_j}\right\}_{j=1}^p$ can be obtained from two independent GWASs. Since the GWASs generally involve large population, it is common to assume that $\hat{\gamma}_j$ and $\hat{\Gamma}_j$ are independently distributed as $\hat{\gamma}_j \sim N\left(\gamma_j, \sigma^2_{\hat{\gamma}_j}\right)$ and $\hat{\Gamma}_j \sim N\left(\Gamma_j, \sigma^2_{\hat{\Gamma}_j}\right)$ with known $\sigma^2_{\hat{\gamma}_j}$ and $\sigma^2_{\hat{\Gamma}_j}$, respectively.

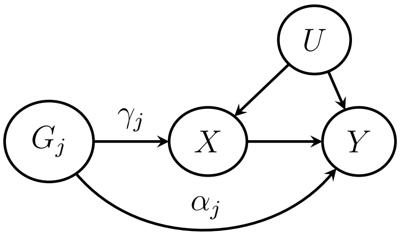

Figure 1: The relationships among the jth genetic variant G_j, the exposure X, the outcome Y and the unmeasured confounder U. The effect of G_j on X is γ_j, the direct effect (pleiotropic effect) of G_j on Y is α_j, and the causal effect of X on Y is β.

2.2 IVW Estimator and Debiased IVW (dIVW) Estimator

The popular inverse-variance weighted (IVW) estimator combines the estimated causal effects $\hat{\beta}_j = \hat{\Gamma}_j/\hat{\gamma}_j$ from each genetic variant with the weights $w_j = \sigma^{-2}_{\hat{\Gamma}_j}\hat{\gamma}^2_j$ as follows:

$$\hat{\beta}_{IVW} = \frac{\sum_{j=1}^p w_j\hat{\beta}_j}{\sum_{j=1}^p w_j} = \frac{\sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\hat{\gamma}_j\hat{\Gamma}_j}{\sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\hat{\gamma}^2_j}.$$

Let $\mu_1 = \sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\gamma_j\Gamma_j$ and $\mu_2 = \sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\gamma^2_j$. We have $\beta = \mu_1/\mu_2$ since $\Gamma_j = \beta\gamma_j$ under models (1)-(2). Then, as shown by Zhao et al. (2020), the IVW estimator can be approximated by

$$\hat{\beta}_{IVW} \approx \frac{E\left[\sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\hat{\gamma}_j\hat{\Gamma}_j\right]}{E\left[\sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\hat{\gamma}^2_j\right]} = \frac{\mu_1}{\mu_2 + \sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\sigma^2_{\hat{\gamma}_j}} = \frac{\beta}{1 + \frac{\sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\sigma^2_{\hat{\gamma}_j}}{\sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\gamma^2_j}}.$$

When there is no measurement error for γ_j (i.e., $\sigma^2_{\hat{\gamma}_j} = 0$), we have $\hat{\beta}_{IVW} \approx \mu_1/\mu_2 = \beta$.

However, studies (Ye et al., 2019; Zhao et al., 2020) indicated that the IVW estimator can be seriously biased toward zero for ignoring the measurement errors of γ_j especially in the presence of many weak IVs that have small $\sigma_{\hat{\gamma}_j}^{-2}\gamma_j^2$.

To handle the bias due to weak IVs, the debiased IVW (dIVW) estimator replaces the denominator in the IVW estimator with an unbiased estimator $\hat{\mu}_2$ of μ_2 as

$$\hat{\beta}_{dIVW} = \frac{\hat{\mu}_1}{\hat{\mu}_2} = \frac{\sum_{j=1}^{p} \sigma_{\hat{\Gamma}_j}^{-2}\hat{\gamma}_j\hat{\Gamma}_j}{\sum_{j=1}^{p} \sigma_{\hat{\Gamma}_j}^{-2}\left(\hat{\gamma}_j^2 - \sigma_{\hat{\gamma}_j}^2\right)}.$$

The dIVW estimator was shown to be consistent and asymptotically normal under weaker conditions than the IVW estimator that allow the presence of weak IVs. However, the dIVW estimator is more likely to yield extreme estimates in the presence of weak IVs, because its denominator $\hat{\mu}_2$ has larger coefficient of variation and thus higher probability of being close to zero than that of the IVW estimator. Therefore, in the following Section 3, we extend the IVW estimator to account for the weak IVs by proposing a penalized log-likelihood function for μ_1 and μ_2, which can prevent the estimator of μ_2 from being close to zero and then provide improved causal estimation.

3 Method

3.1 The Penalized IVW (pIVW) Estimator

Assume that the estimators $\hat{\mu}_1 = \sum_{j=1}^{p} \sigma_{\hat{\Gamma}_j}^{-2}\hat{\gamma}_j\hat{\Gamma}_j$ of μ_1 and $\hat{\mu}_2 = \sum_{j=1}^{p} \sigma_{\hat{\Gamma}_j}^{-2}\left(\hat{\gamma}_j^2 - \sigma_{\hat{\gamma}_j}^2\right)$ of μ_2 follow a bivariate normal distribution as $\begin{pmatrix}\hat{\mu}_1\\ \hat{\mu}_2\end{pmatrix} \sim N_2\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \Sigma\right)$ with the covariance matrix $\Sigma = \begin{pmatrix} v_1 & v_{12} \\ v_{12} & v_2 \end{pmatrix}$. We propose a penalized log-likelihood function to adjust the estimates of μ_1 and μ_2 as

$$l_p(\mu_1, \mu_2) = \log f(\hat{\mu}_1, \hat{\mu}_2) + \lambda \log |\mu_2|,$$

where $f(\hat{\mu}_1, \hat{\mu}_2)$ denotes the bivariate normal density function of $(\hat{\mu}_1, \hat{\mu}_2)'$ and $\lambda > 0$ is the penalty parameter. We adopt the penalty term $\log|\mu_2|$ inspiring by Wang et al. (2020), which was also used to handle the issue of near-zero denominator when deriving the confidence interval for a ratio estimator. Through the penalty, the penalized log-likelihood $l_p(\mu_1, \mu_2)$ will become smaller when μ_2 approaches to zero. Therefore, an estimate of μ_2 being close to zero is less preferable by $l_p(\mu_1, \mu_2)$. Specifically, by maximizing $l_p(\mu_1, \mu_2)$, we can obtain the adjusted estimators of μ_1 and μ_2 as

$$\tilde{\mu}_1 = \hat{\mu}_1 + \frac{\hat{v}_{12}}{\hat{v}_2}(\tilde{\mu}_2 - \hat{\mu}_2), \tag{3}$$

$$\tilde{\mu}_2 = \left(\frac{1}{2} + \sqrt{\frac{1}{4} + \lambda\frac{\hat{v}_2}{\hat{\mu}_2^2}}\right)\hat{\mu}_2, \tag{4}$$

respectively, where $\hat{v}_{12} = 2\sum_{j=1}^{p} \sigma_{\hat{\gamma}_j}^2 \sigma_{\hat{\Gamma}_j}^{-4} \hat{\Gamma}_j \hat{\gamma}_j$ and $\hat{v}_2 = \sum_{j=1}^{p} \sigma_{\hat{\Gamma}_j}^{-4}\left[4\left(\hat{\gamma}_j^2 - \sigma_{\hat{\gamma}_j}^2\right)\sigma_{\hat{\gamma}_j}^2 + 2\sigma_{\hat{\gamma}_j}^4\right]$ are the estimates of v_{12} and v_2, respectively. Then, we propose the penalized IVW (pIVW) estimator as a ratio of $\tilde{\mu}_1$ and $\tilde{\mu}_2$, that is

$$\hat{\beta}_{pIVW} = \frac{\tilde{\mu}_1}{\tilde{\mu}_2} = \frac{\hat{\mu}_1}{\tilde{\mu}_2} + \frac{\hat{v}_{12}}{\hat{v}_2}\left(1 - \frac{\hat{\mu}_2}{\tilde{\mu}_2}\right). \tag{5}$$

Note that when the penalty parameter $\lambda = 0$, we have $\tilde{\mu}_2 = \hat{\mu}_2$ and therefore $\hat{\beta}_{pIVW}$ reduces to $\hat{\beta}_{dIVW}$. When $\lambda > 0$, $\hat{\beta}_{pIVW}$ will also be close to $\hat{\beta}_{dIVW}$ when $\hat{v}_2/\hat{\mu}_2^2$ approaches to zero. To see this, we define the IV strength as

$$\kappa = \frac{\sum_{j=1}^{p} \gamma_j^2 \sigma_{\hat{\gamma}_j}^{-2}}{p},$$

which can be estimated by $\hat{\kappa} = \frac{1}{p}\sum_{j=1}^{p}\left(\hat{\gamma}_j^2 - \sigma_{\hat{\gamma}_j}^2\right)\sigma_{\hat{\gamma}_j}^{-2}$. Further, we make the following assumptions:

Assumption 1. The number of IVs, p, diverges to infinity.

Assumption 2. $\left\{\hat{\gamma}_j, \hat{\Gamma}_j\right\}_{j=1}^{p}$ are independently distributed as $\hat{\gamma}_j \sim N\left(\gamma_j, \sigma_{\hat{\gamma}_j}^2\right)$ and $\hat{\Gamma}_j \sim N\left(\beta\gamma_j, \sigma_{\hat{\Gamma}_j}^2\right)$ with known variances $\sigma_{\hat{\gamma}_j}^2$ and $\sigma_{\hat{\Gamma}_j}^2$. The ratio of variances $\sigma_{\hat{\gamma}_j}^2/\sigma_{\hat{\Gamma}_j}^2$ is bounded away from zero and infinity for all j.

We notice that Assumptions 1 and 2 are also required for the consistency of the dIVW estimator (Ye et al., 2019). They are reasonable in the two-sample MR design since the GWASs often involve large population and a large amount of genetic variants. And the independence among $\left\{\hat{\gamma}_j, \hat{\Gamma}_j\right\}_{j=1}^{p}$ can be guaranteed by the linkage-disequilibrium clumping. Then, under Assumptions 1 and 2, $\hat{v}_2/\hat{\mu}_2^2 = O_p\left(\frac{1}{\kappa p} + \frac{1}{\kappa^2 p}\right)$ and hence $\hat{\beta}_{pIVW}$ reduces to $\hat{\beta}_{dIVW}$ as $\kappa\sqrt{p} \to \infty$. More importantly, in the following Theorem 3.1 (a)-(b), we show that $\hat{\beta}_{pIVW}$ can achieve smaller bias and variance than $\hat{\beta}_{dIVW}$ under proper choice of the penalty parameter λ. Further, in the Theorem 3.1 (c), we show that $\hat{\beta}_{pIVW}$ is consistent and asymptotically normal under some regularity conditions.

Theorem 3.1. *Suppose that models (1)-(2) and Assumptions 1-2 hold. Then, as $\kappa\sqrt{p} \to \infty$, we have the following conclusions.*

(a) The approximate bias of $\hat{\beta}_{dIVW}$ is of order $O\left(\frac{1}{\kappa p} + \frac{1}{\kappa^2 p}\right)$, while that of $\hat{\beta}_{pIVW}$ is only of order $o\left(\frac{1}{\kappa p} + \frac{1}{\kappa^2 p}\right)$ when $\lambda = 1$.

(b) The approximate variance of $\hat{\beta}_{pIVW}$ is smaller than that of $\hat{\beta}_{dIVW}$ at the order $O\left(\frac{1}{\kappa p} + \frac{1}{\kappa^2 p}\right)$ when $\lambda > 0$.

(c) Further assume that $\max_j \gamma_j^2 \sigma_{\hat{\gamma}_j}^{-2}/(\kappa p + p) \to 0$. Then, $\hat{\beta}_{pIVW}$ is consistent and asymptotically normally distributed as

$$V^{-\frac{1}{2}}\left(\hat{\beta}_{pIVW} - \beta\right) \xrightarrow{d} N(0,1),$$

where

$$V = \mu_2^{-2}\sum_{j=1}^{p}\left[\sigma_{\hat{\Gamma}_j}^{-2}\left(\gamma_j^2 + \sigma_{\hat{\gamma}_j}^2\right) + \beta^2\sigma_{\hat{\gamma}_j}^2\sigma_{\hat{\Gamma}_j}^{-4}\left(\gamma_j^2 + 2\sigma_{\hat{\gamma}_j}^2\right)\right].$$

The consistency and asymptotic normality of $\hat{\beta}_{pIVW}$ still hold, when we replace V with the estimator

$$\hat{V} = \tilde{\mu}_2^{-2}\sum_{j=1}^{p}\left[\sigma_{\hat{\Gamma}_j}^{-2}\hat{\gamma}_j^2 + \hat{\beta}_{pIVW}^2\sigma_{\hat{\gamma}_j}^2\sigma_{\hat{\Gamma}_j}^{-4}\left(\hat{\gamma}_j^2 + \sigma_{\hat{\gamma}_j}^2\right)\right].$$

Under the setting of many weak IVs, we may have $\kappa \to 0$ as $p \to \infty$ because more weak IVs are likely to be included into the analysis as the number IVs p increases, which may reduce the IV strength κ. Nevertheless, the above theorem can still hold in this case as long as $\kappa\sqrt{p} \to \infty$, which means that it allows the presence of many weak IVs. On the other hand, although many ratio estimators may not have finite moments when their denominators are close to zero (Piegorsch and Casella, 1985; Marsaglia et al., 2006; Press, 1969), for the denominator $\hat{\mu}_2$ of $\hat{\beta}_{dIVW}$, we have $P(|\hat{\mu}_2| > \epsilon) \to 1$ for any $\epsilon > 0$ as $\kappa\sqrt{p} \to \infty$. And we have same result for the denominator $\tilde{\mu}_2$ of $\hat{\beta}_{pIVW}$ since $\tilde{\mu}_2$ is in larger magnitude than $\hat{\mu}_2$ due to Equation (6). Therefore, we can still approximate the biases and variances of $\hat{\beta}_{dIVW}$ and $\hat{\beta}_{pIVW}$ in this case. Moreover, when $\kappa\sqrt{p} \to \infty$, Therorem 3.1 (a)-(b) indicate that the approximate bias of $\hat{\beta}_{pIVW}$ converges to zero in a faster rate than that of $\hat{\beta}_{dIVW}$ when $\lambda = 1$, while the approximate variance of $\hat{\beta}_{pIVW}$ is also smaller than that of $\hat{\beta}_{dIVW}$, which implies that we may choose $\lambda = 1$ for the pIVW estimator in practice to achieve improved performances over the dIVW estimator. Further, in the Theorem 3.1 (c), we show that $\hat{\beta}_{pIVW}$ is consistent and asymptotically normal, which requires no further assumptions comparing to $\hat{\beta}_{dIVW}$.

In the following sections, we show that these nice properties of the proposed pIVW estimator still hold when we conduct selection to exclude some weak IVs from the analysis and account for the horizontal pleiotropy.

3.2 Selection for Candidate Instruments

To handle the weak IV bias, it is common in MR studies to exclude the weak IVs with small $\sigma_{\hat{\gamma}_j}^{-2}\hat{\gamma}_j^2$ from the analysis. However, the IV selection based on the exposure dataset $\{\hat{\gamma}_j, \sigma_{\hat{\gamma}_j}\}_{j=1}^p$ might lead to selection bias or winner's curse that would also bias the causal estimation (Zhao et al., 2020). To address this problem, in the three-sample MR design (Zhao et al., 2019), IV selection is performed on a third dataset $\left\{\hat{\gamma}_j^*, \sigma_{\hat{\gamma}_j}^*\right\}_{j=1}^p$ which is independent of the exposure dataset and the outcome dataset. Specifically, an IV is included into the analysis when $|\hat{\gamma}_j^*| > \delta\sigma_{\hat{\gamma}_j}^*$ with a pre-set threshold $\delta > 0$. Ye et al. (2019) showed that IV selection with an appropriate threshold δ could reduce the bias of the IVW estimator and improve the efficiency of the dIVW estimator. They also recommended a threshold $\delta = \sqrt{2\log p}$ to guarantee small probability of selecting any null IVs (i.e., $\gamma_j = 0$).

When IV selection is performed at a threshold δ, we define the IV strength as

$$\kappa_\delta = \frac{\sum_{j=1}^p \gamma_j^2 \sigma_{\hat{\gamma}_j}^{-2} q_{\delta,j}}{p_\delta},$$

where $q_{\delta,j} = P\left(|\hat{\gamma}_j^*| > \delta\sigma^*_{\hat{\gamma}_j}\right)$ and $p_\delta = \sum_{j=1}^p q_{\delta,j}$. Let $S_\delta = \left\{j : |\hat{\gamma}_j^*| > \delta\sigma^*_{\hat{\gamma}_j}\right\}$ be the set of selected IVs. Then, we can estimate κ_δ by $\hat{\kappa}_\delta = \hat{p}_\delta^{-1}\sum_{j\in S_\delta}\left(\hat{\gamma}_j^2 - \sigma^2_{\hat{\gamma}_j}\right)\sigma^{-2}_{\hat{\gamma}_j}$, where $\hat{p}_\delta$ is the number of selected IVs within S_δ. To study the theoretical properties of the proposed pIVW estimator under IV selection, we have the following Assumption 3 for the summary-level data in the selection dataset.

Assumption 3. $\left\{\hat{\gamma}_j^*, \hat{\gamma}_j, \hat{\Gamma}_j\right\}_{j=1}^p$ are mutually independent and $\hat{\gamma}_j^* \sim N\left(\gamma_j, \sigma^{*2}_{\hat{\gamma}_j}\right)$ with known variance $\sigma^{*2}_{\hat{\gamma}_j}$ for every j. The ratio of variances $\sigma^2_{\hat{\gamma}_j}/\sigma^{*2}_{\hat{\gamma}_j}$ is bounded away from zero and infinity for all j.

Given a selection threshold δ, we evaluate the dIVW estimator $\hat{\beta}_{\delta,dIVW} = \hat{\mu}_{1,\delta}/\hat{\mu}_{2,\delta}$ and the proposed pIVW estimator $\hat{\beta}_{\delta,pIVW} = \tilde{\mu}_{1,\delta}/\tilde{\mu}_{2,\delta}$ using the selected IVs within the set S_δ. In the following Theorems 3.2, we show that, $\hat{\beta}_{\delta,pIVW}$ can still have smaller bias when $\lambda = 1$ and has smaller variance when $\lambda > 0$ than $\hat{\beta}_{\delta,dIVW}$, and $\hat{\beta}_{\delta,pIVW}$ is consistent and asymptotically normal under some regularity conditions.

Theorem 3.2. *Suppose that models (1)-(2) and Assumptions 1-3 hold, and* $\kappa_\delta\sqrt{p_\delta}/\max(1,\delta^2) \to \infty$. *Then, we have the following conclusions.*
(a) The approximate bias of $\hat{\beta}_{\delta,dIVW}$ *is of order* $O\left(\frac{1}{\kappa_\delta p_\delta} + \frac{1+\delta^4}{\kappa_\delta^2 p_\delta}\right)$, *while that of* $\hat{\beta}_{\delta,pIVW}$ *is only of order* $o\left(\frac{1}{\kappa_\delta p_\delta} + \frac{1+\delta^4}{\kappa_\delta^2 p_\delta}\right)$ *when* $\lambda = 1$.
(b) The approximate variance of $\hat{\beta}_{\delta,pIVW}$ *is smaller than that of* $\hat{\beta}_{\delta,dIVW}$ *at the order* $O\left(\frac{1}{\kappa_\delta p_\delta} + \frac{1+\delta^4}{\kappa_\delta^2 p_\delta}\right)$ *when* $\lambda > 0$.
(c) Further assume that $\max_j \gamma_j^2\sigma^{-2}_{\hat{\gamma}_j}q_{\delta,j}/(\kappa_\delta p_\delta + p_\delta) \to 0$. *Then,* $\hat{\beta}_{\delta,pIVW}$ *is consistent and asymptotically normally distributed as*

$$V_\delta^{-\frac{1}{2}}\left(\hat{\beta}_{\delta,pIVW} - \beta\right) \xrightarrow{d} N(0,1),$$

where

$$V_\delta = \left[\sum_{j=1}^p \sigma^{-2}_{\hat{\Gamma}_j}\gamma_j^2 q_{\delta,j}\right]^{-2}\sum_{j=1}^p\left[\sigma^{-2}_{\hat{\Gamma}_j}\left(\gamma_j^2 + \sigma^2_{\hat{\gamma}_j}\right) + \beta^2\sigma^2_{\hat{\gamma}_j}\sigma^{-4}_{\hat{\Gamma}_j}\left(\gamma_j^2 + 2\sigma^2_{\hat{\gamma}_j}\right)\right]q_{\delta,j}.$$

The consistency and asymptotic normality of $\hat{\beta}_{\delta,pIVW}$ *still hold, when we replace* V_δ *with the* *estimator*

$$\hat{V}_\delta = \tilde{\mu}_{2,\delta}^{-2} \sum_{j \in S_\delta} \left[\sigma_{\hat{\Gamma}_j}^{-2} \hat{\gamma}_j^2 + \hat{\beta}_{\delta,pIVW}^2 \sigma_{\hat{\gamma}_j}^2 \sigma_{\hat{\Gamma}_j}^{-4} \left(\hat{\gamma}_j^2 + \sigma_{\hat{\gamma}_j}^2 \right) \right].$$

Under the IV selection at a threshold δ, Theorem 3.2 provides similar conclusions to Theorems 3.1. It indicates that the pIVW estimator can have smaller bias and variance than the dIVW estimator, and it is consistent and asymptotically normal regardless of whether or not the IV selection is conducted. In comparison, Theorem 3.2 also involves the IV selection threshold δ which may vary with the number of IV p, for instance, $\delta = \sqrt{2 \log p}$. In the special case when $\delta = 0$ (i.e., no IV selection), Theorems 3.2 is equivalent to Theorems 3.1.

3.3 Accounting for Horizontal Pleiotropy

When there exists horizontal pleiotropy (i.e., direct effect of G_j on Y not mediated by X), the linear structural model (2) can be modified as

$$Y = \beta X + \sum_{j=1}^{p} \alpha_j G_j + U + \epsilon_Y, \tag{6}$$

where α_j denotes the direct genetic effect of G_j on the outcome Y (i.e., pleiotropic effect). In this case, we have $\Gamma_j = \beta\gamma_j + \alpha_j$. We follow a common practice in many MR methods to assume that the horizontal pleiotropy is balanced (i.e., the pleiotropic effect has mean zero) (Ye et al., 2019; Zhao et al., 2020; Bowden et al., 2017; Cheng et al., 2020) and treat α_j as random effect following $\alpha_j \sim N(0, \tau^2)$. Then, $\hat{\Gamma}_j \sim N\left(\beta\gamma_j, \sigma_{\hat{\Gamma}_j}^2 + \tau^2\right)$ has larger variance in the presence of horizontal pleiotropy due to the non-zero τ^2. To account for the horizontal pleiotropy, we adjust the variance estimation of $\hat{\beta}_{\delta,pIVW}$ by

$$V_\delta^* = \left[\sum_{j=1}^{p} \sigma_{\hat{\Gamma}_j}^{-2} \gamma_j^2 q_{\delta,j} \right]^{-2} \sum_{j=1}^{p} \left[\sigma_{\hat{\Gamma}_j}^{-2} \left(\gamma_j^2 + \sigma_{\hat{\gamma}_j}^2 \right) \left(1 + \tau^2 \sigma_{\hat{\Gamma}_j}^{-2} \right) + \beta^2 \sigma_{\hat{\gamma}_j}^2 \sigma_{\hat{\Gamma}_j}^{-4} \left(\gamma_j^2 + 2\sigma_{\hat{\gamma}_j}^2 \right) \right] q_{\delta,j}.$$

Following Ye et al. (2019), we derive an estimator of τ^2 as

$$\hat{\tau}_2 = \frac{\sum_{j=1}^{p} \left[\left(\hat{\Gamma}_j - \hat{\beta}_{pIVW} \hat{\gamma}_j \right)^2 - \sigma_{\hat{\Gamma}_j}^2 - \hat{\beta}_{pIVW}^2 \sigma_{\hat{\gamma}_j}^2 \right] \sigma_{\hat{\Gamma}_j}^{-2}}{\sum_{j=1}^{p} \sigma_{\hat{\Gamma}_j}^{-2}}. \tag{7}$$

Then, we estimate V_δ^* by

$$\hat{V}_\delta^* = \tilde{\mu}_{2,\delta}^{-2} \sum_j \left[\sigma_{\hat{\Gamma}_j}^{-2} \hat{\gamma}_j^2 \left(1 + \hat{\tau}^2 \sigma_{\hat{\Gamma}_j}^{-2}\right) + \hat{\beta}_{\delta,pIVW}^2 \sigma_{\hat{\gamma}_j}^2 \sigma_{\hat{\Gamma}_j}^{-4} \left(\hat{\gamma}_j^2 + \sigma_{\hat{\gamma}_j}^2\right)\right].$$

To establish the theoretical results for the bias, the variance and the asymptotic property for $\hat{\beta}_{\delta,pIVW}$, we replace $\hat{\Gamma}_j \sim N\left(\Gamma_j, \sigma_{\hat{\Gamma}_j}^2\right)$ in the Assumption 2 by $\hat{\Gamma}_j \sim N\left(\Gamma_j, \sigma_{\hat{\Gamma}_j}^2 + \tau^2\right)$. We further assume that $\tau^2 < c_1 \sigma_{\hat{\Gamma}_j}^2$ with a constant c_1 for all j, and that $\max_j \sigma_{\hat{\Gamma}_j}^{-2} < c_2 p^{-1} \sum_{j=1}^p \sigma_{\hat{\Gamma}_j}^{-2}$ for a constant c_2 in Theorem 3.2 (c). Then, Theorem 3.2 still hold with V_δ and $\hat{V}_\delta$ replaced by V_δ^* and $\hat{V}_\delta^*$, respectively.

4 Simulation Study

4.1 Simulation Settings

We randomly generate the summary-level data for 1000 IVs from $\hat{\gamma}_j \sim N\left(\gamma_j, \sigma_{\hat{\gamma}_j}^2\right)$ and $\hat{\Gamma}_j \sim N\left(\Gamma_j, \sigma_{\hat{\Gamma}_j}^2\right)$ independently. For the true γ_j, we consider three scenarios as follows.

- Scenario A (some strong IVs and many null IVs): 10 IVs have $\gamma_j \sim N(0, 0.03^2)$ and the rest of IVs have $\gamma_j = 0$. As such, the IV strength of 10 non-null IVs is around 15.54 and that of all 1000 IVs is around 0.16.

- Scenario B (many weak IVs and many null IVs): 100 IVs have $\gamma_j \sim N(0, 0.02^2)$ and the rest of IVs have $\gamma_j = 0$. As such, the IV strength of 100 non-null IVs is around 3.11 and that of all 1000 IVs is around 0.31.

- Scenario C (all weak IVs):all IVs have $\gamma_j \sim N(0, 0.02^2)$. The IV strength is around 4.07.

Then, we let $\Gamma_j = \beta\gamma_j + \alpha_j$, where $\alpha_j \sim N(0, \tau^2)$. We set $\beta = 0.5$, and $\tau = 0$ and 0.01 representing the absence and the presence of horizontal pleiotropy, respectively. For the variances $\sigma_{\hat{\gamma}_j}^2$ and $\sigma_{\hat{\Gamma}_j}^2$, we calculate them by $\sigma_{\hat{\gamma}_j}^2 = \frac{\text{var}(X) - \gamma_j^2 \text{var}(G_j)}{n_X \text{var}(G_j)}$ and $\sigma_{\hat{\Gamma}_j}^2 = \frac{\text{var}(Y) - \Gamma_j^2 \text{var}(G_j)}{n_Y \text{var}(G_j)}$, where n_X and n_Y denote the sample sizes of the GWASs for the exposure and the outcome, respectively. We set $n_X = n_Y = 100,000$. For $\text{var}(G_j)$, we let $G_j \sim Bin(2, MAF_j)$ and randomly generate the minor allele frequencies by $MAF_j \sim U(0.1, 0.5)$. For $\text{var}(X)$ and $\text{var}(Y)$, we calculate them from Equations (1) and (6) with the variances of U, ϵ_X and ϵ_Y being 2, respectively. Furthermore, we generate an independent dataset with $\hat{\gamma}_j^* \sim N\left(\gamma_j, \sigma^2_{\hat{\gamma}_j}/2\right)$ for the IV selection at threshold $\delta = \sqrt{2\log p} = 3.72$. The simulation is based on 10,000 replicates.

In the simulation study, we first investigate the impact of the penalty parameter λ on the performance of the proposed pIVW estimator, where λ increases from 0 to 2.5 by 0.5. Then, we compare the proposed pIVW estimator with $\lambda = 1$ to competing MR methods, including the IVW, MR-Egger (Bowden et al., 2015), MR-Median (Bowden et al., 2016), MR-RAPS (Zhao et al., 2020) and dIVW estimators. For fair comparison, we use l_2 loss function for MR-RAPS. The overview of the competing methods can be found in the Supplementary. The performances among various methods are compared in terms of the relative bias (bias divided by the true β) and the empirical standard error of the estimated causal effect, as well as the coverage probability of 95% confidence interval.

4.2 Simulation Results

In Table 1, we investigate the performances of the pIVW estimator with various penalty parameter λ when no horizontal pleiotropy exists and no IV selection is conducted. Note that, when $\lambda = 0$, the pIVW estimator reduces to the dIVW estimator. The pIVW estimator has smallest bias at $\lambda = 1$ under Scenarios A and B. Specifically, the bias decreases as λ increases from 0 to 1 and increases as λ increases from 1 to 2.5. On the other hand, the empirical standard error decreases as the increase of λ. The coverage probability of confidence interval is generally around 95%. In Scenario C where there is larger IV strength than Scenarios A and B, the pIVW estimator has nearly the same performance when λ varies from 0 to 2.5. We observe similar impact of λ on the performances of the pIVW estimator when there exists horizontal pleiotropy and when the IVs are selected from an independent dataset (Supplementary Tables S1.1-S1.3).

In Table 2, we compare the proposed pIVW estimator with the competing MR methods when no horizontal pleiotropy exists and no IV selection is conducted. Here, we use the penalty parameter $\lambda = 1$ for the pIVW estimator. In Scenarios A and B, the IVW, MR-Egger and MR-Median estimators have serious biases and poor coverage probabilities.

Table 1: The performances of the pIVW estimator with various penalty parameter λ. The relative bias (bias divided by the true causal effect), the empirical standard error and the coverage probability of the 95% confidence interval. The true causal effect $\beta = 0.5$. No horizontal pleiotropy exists ($\tau = 0$). No IV selection is conducted. The simulation is based on 10,000 replicates

Scenario	Penalty parameter λ	Relative bias (%)	Empirical standard error	Coverage probability (%)
A (some strong and many null IVs)	0.0	15.3	1.272	96.1
	0.5	4.7	0.382	95.6
	1.0	-1.1	0.341	94.9
	1.5	-5.5	0.317	94.2
	2.0	-9.1	0.299	93.6
	2.5	-12.1	0.285	92.8
B (many weak and many null IVs)	0.0	3.4	0.190	96.1
	0.5	1.7	0.185	95.9
	1.0	0.2	0.181	95.7
	1.5	-1.2	0.177	95.3
	2.0	-2.6	0.173	95.0
	2.5	-3.8	0.170	94.6
C (all weak IVs)	0.0	0.0	0.025	95.2
	0.5	0.0	0.025	95.2
	1.0	0.0	0.025	95.2
	1.5	-0.1	0.025	95.1
	2.0	-0.1	0.025	95.1
	2.5	-0.2	0.025	95.1

In comparison, the MR-RAPS and dIVW estimators have smaller biases, but we find that they are likely to produce extreme estimates which result in their relatively large empirical standard errors in Scenarios A (see Supplementary Figure S1.1 for the box plot of the estimates from 10,000 simulation replicates). In contrast, the proposed pIVW estimator has smallest bias among six methods and also has smaller empirical standard errors than the dIVW estimators. In Scenario C where the IV strength is larger, all the methods have improved performances. However, the IVW, MR-Egger and MR-Median estimators still have substantial biases, whereas the proposed pIVW estimator is nearly unbiased. The MR-RAPS and dIVW estimators perform similarly to the pIVW estimator in this case. In addition, we study the impact of the number of IVs (i.e., p) by increasing p from 1000 to 5000 in all three scenarios and using 50 and 500 non-null IVs in Scenarios A and B, respectively (Supplementary Tables S1.4). We find that the increase in p has little impact on the biases

Table 2: Comparison among six methods. The relative bias (bias divided by the true causal effect), the empirical standard error and the coverage probability of the 95% confidence interval. The true causal effect $\beta = 0.5$. No horizontal pleiotropy exists ($\tau = 0$). No IV selection is conducted. The simulation is based on 10,000 replicates

Scenario	Method	Relative bias (%)	Empirical standard error	Coverage probability (%)
A (some strong and many null IVs)	IVW	-86.5	0.039	0.0
	MR-Egger	-73.1	0.060	0.0
	MR-Median	-70.9	0.061	0.0
	MR-RAPS	9.2	1.595	94.2
	dIVW	15.3	1.272	96.1
	pIVW*	-1.1	0.341	94.9
B (many weak and many null IVs)	IVW	-76.8	0.037	0.0
	MR-Egger	-63.2	0.057	0.0
	MR-Median	-65.7	0.054	0.0
	MR-RAPS	1.6	0.167	95.0
	dIVW	3.4	0.190	96.1
	pIVW*	0.2	0.181	95.7
C (all weak IVs)	IVW	-19.7	0.019	0.1
	MR-Egger	-18.4	0.031	16.5
	MR-Median	-18.8	0.026	9.4
	MR-RAPS	0.0	0.024	95.1
	dIVW	0.0	0.025	95.2
	pIVW*	0.0	0.025	95.2

*The pIVW estimator with penalty parameter $\lambda = 1$

of the IVW, MR-Egger and MR-Median estimators, while it can further decrease the biases and empirical standard errors of the MR-RAPS, dIVW and pIVW estimators.

Table 3 lists the results with IV selection at threshold $\delta = \sqrt{2 \log p}$. In Scenarios A and B with many null IVs, the IV selection reduces the biases of the IVW, MR-Egger and MR-Median estimators to a great extent, and it also reduces the empirical standard errors of the MR-RAPS and dIVW and pIVW estimators. However, the IV selection is not necessarily helpful for all methods in Scenario C where all the IVs are weak. In this case, the IV selection increases the bias of MR-Egger and it also raises the empirical standard errors of the MR-RAPS and dIVW and pIVW estimators. It is probably because the number of IVs largely decreases under the IV selection but the gain in the IV strength is small, which therefore reduces the efficiencies of methods. Nevertheless, the proposed pIVW estimator still has smallest bias among six methods and smaller empirical standard error than the dIVW

Table 3: Comparison among six methods. The relative bias (bias divided by the true causal effect), the empirical standard error and the coverage probability of the 95% confidence interval. The true causal effect $\beta = 0.5$. No horizontal pleiotropy exists ($\tau = 0$). The IV selection threshold $\delta = \sqrt{2 \log p}$. The simulation is based on 10,000 replicates

Scenario	Method	Relative bias (%)	Empirical standard error	Coverage probability (%)
A (some strong and many null IVs)	IVW	-1.8	0.116	95.3
	MR-Egger	-4.7	0.285	95.4
	MR-Median	-1.8	0.127	96.4
	MR-RAPS	0.8	0.120	95.7
	dIVW	1.5	0.121	95.6
	pIVW*	0.1	0.119	95.1
B (many weak and many null IVs)	IVW	-7.6	0.106	93.8
	MR-Egger	-43.6	0.277	87.0
	MR-Median	-11.6	0.131	95.6
	MR-RAPS	1.1	0.117	95.4
	dIVW	1.8	0.120	95.5
	pIVW*	0.4	0.118	95.2
C (all weak IVs)	IVW	-7.4	0.025	70.3
	MR-Egger	-40.6	0.064	11.3
	MR-Median	-10.8	0.033	74.4
	MR-RAPS	0.0	0.028	95.2
	dIVW	0.0	0.028	95.2
	pIVW*	0.0	0.028	95.3

*The pIVW estimator with penalty parameter $\lambda = 1$

estimator regardless of the use of IV selection. When there exists horizontal pleiotropy (Supplementary Tables S1.5-S1.6), we find similar results to those in Tables 2-3 except that all the methods have increased standard errors in these cases.

On the other hand, although the IVW estimator suffers from substantial bias in the presence of weak IVs, Ye et al. (2019) indicated that the IVW estimator was still consistent when the the true causal effect β was zero. Therefore, we conduct further simulation under $\beta = 0$ (Supplementary Table S1.7). In this case, we find that all six methods are nearly unbiased even in the the presence of weak IVs under Scenarios A-C. However, the MR-RAPS and dIVW estimators still have some extremely large estimates and relatively large empirical standard errors in Scenario A (Supplementary Figure S1.2). Overall, the proposed pIVW estimator is applicable to more general situations than the competing methods, which has smallest bias among all six methods and has smaller standard error than the dIVW estimator under the situations considered.

5 Obesity-related exposures and COVID-19 outcomes

In this section, we focus our interest on the causal effects of five obesity-related diseases and traits (i.e., peripheral vascular disease, dyslipidemia, hypertensive disease, type 2 diabetes and BMI) on COVID-19. Specifically, we consider three categories of COVID-19 outcomes including (1) COVID-19 infection, (2) hospitalized COVID-19 and (3) critically ill COVID-19, which involve population controls versus (1) cases with reported SARS-CoV-2 infection, (2) cases with moderate or severe COVID-19 defined as those hospitalized due to symptoms associated with the infection, and (3) cases required respiratory support in hospital or were deceased due to COVID-19, respectively (COVID-19 Host Genetics Initiative, 2021). The GWAS summary-level data of three categories of COVID-19 outcomes is obtained from the GWAS meta-analysis round 5 of the COVID19 Host Genetics Initiative (COVID-19 Host Genetics Initiative, 2020), which involves up to 49,562 cases and two million controls from 47 distinct studies with populations of different genetic ancestries, including European, Admixed American, African, Middle Eastern, South Asian and East Asian. For BMI, the selection dataset is from Akiyama et al. (Akiyama et al., 2017) with 173,430 individuals of Asian ancestry and exposure dataset is from UK BioBank (Abbott et al., 2018) with 359,983 individuals of European ancestry. For the other four obesity-related diseases, the selection datasets and the exposure datasets are from the GWAS meta-analyses (Zhu et al., 2018) of Genetic Epidemiology Research on Adult Health and Aging (GERA) (Banda et al., 2015) with 53,991 individuals of European ancestry and UK BioBank (Abbott et al., 2018) with 108,039 individuals of European ancestry, respectively. More detailed data description is provided in the Supplementary Table S2.1. To exclude correlated IVs, we perform the linkage-disequilibrium clumping to remove the correlated genetic variants within 10Mb pairs and with the linkage disequilibrium $r^2 < 0.001$.

Figures 2 displays the estimated causal effects, the estimated standard errors and the corresponding P values when no IV selection is conducted. Among the five obesity-related diseases and traits, we find significant positive causal effects of hypertensive disease on hospitalized COVID-19 and BMI on three COVID-19 outcomes by the pIVW estimator at 0.05

(a) Outcome: COVID-19 infection

	IVW	MR-Egger	MR-Median	MR-RAPS	dIVW	pIVW
Peripheral vascular disease (no. of IVs:1783, IV strength:0.04)	3e-03 (0.007) 0.672	3e-03 (0.01) 0.754	5e-03 (0.011) 0.610	0.097 (0.223) 0.664	0.102 (0.27) 0.706	0.057 (0.14) 0.685
Dyslipidemia (no. of IVs:2163, IV strength:0.82)	7e-03 (0.013) 0.567	1e-02 (0.017) 0.579	-4e-03 (0.025) 0.861	0.017 (0.033) 0.600	0.018 (0.032) 0.569	0.018 (0.032) 0.569
Hypertensive disease (no. of IVs:2149, IV strength:0.62)	0.025 (0.018) 0.148	0.032 (0.025) 0.195	0.019 (0.031) 0.532	0.083 (0.055) 0.131	0.074 (0.052) 0.151	0.074 (0.052) 0.151
Type 2 diabetes (no. of IVs:2155, IV strength:0.31)	-5e-03 (0.01) 0.586	-6e-03 (0.013) 0.634	-0.012 (0.018) 0.521	-0.023 (0.042) 0.593	-0.023 (0.042) 0.588	-0.022 (0.041) 0.587
Body mass index (no. of IVs:1893, IV strength:5.01)	0.095 (0.041) 0.020 *	0.119 (0.055) 0.031 *	0.038 (0.082) 0.641	0.116 (0.05) 0.021 *	0.116 (0.05) 0.021 *	0.115 (0.05) 0.021 *

* P<0.05
Estimated causal effect: 0.08, 0.04, 0.00

(b) Outcome: hospitalized COVID-19

	IVW	MR-Egger	MR-Median	MR-RAPS	dIVW	pIVW
Peripheral vascular disease (no. of IVs:1783, IV strength:0.04)	7e-03 (0.013) 0.564	0.022 (0.019) 0.250	0.032 (0.02) 0.108	0.379 (0.55) 0.491	4.413 (95.942) 0.963	0.219 (0.429) 0.610
Dyslipidemia (no. of IVs:2163, IV strength:0.82)	-5e-03 (0.025) 0.850	-8e-03 (0.033) 0.812	-0.013 (0.046) 0.781	-0.013 (0.063) 0.830	-0.011 (0.059) 0.852	-0.011 (0.059) 0.851
Hypertensive disease (no. of IVs:2149, IV strength:0.62)	0.089 (0.033) 7e-03 *	0.068 (0.047) 0.147	0.109 (0.059) 0.064	0.241 (0.095) 0.011 *	0.246 (0.093) 9e-03 *	0.244 (0.093) 8e-03 *
Type 2 diabetes (no. of IVs:2155, IV strength:0.31)	-5e-04 (0.019) 0.981	-0.031 (0.026) 0.244	-0.038 (0.041) 0.352	-3e-03 (0.085) 0.973	-2e-03 (0.085) 0.981	-2e-03 (0.083) 0.977
Body mass index (no. of IVs:1893, IV strength:5.01)	0.382 (0.077) 8e-07 *	0.455 (0.105) 2e-05 *	0.549 (0.141) 1e-04 *	0.466 (0.097) 1e-06 *	0.468 (0.096) 1e-06 *	0.468 (0.096) 1e-06 *

* P<0.05
Estimated causal effect: 0.6, 0.4, 0.2, 0.0

(c) Outcome: critically ill COVID-19

	IVW	MR-Egger	MR-Median	MR-RAPS	dIVW	pIVW
Peripheral vascular disease (no. of IVs:1783, IV strength:0.04)	9e-03 (0.019) 0.645	0.014 (0.029) 0.618	0.030 (0.029) 0.305	0.320 (0.578) 0.580	0.464 (1.326) 0.726	0.202 (0.449) 0.653
Dyslipidemia (no. of IVs:2163, IV strength:0.82)	-8e-03 (0.036) 0.829	0.014 (0.047) 0.764	-0.022 (0.066) 0.737	-0.020 (0.078) 0.799	-0.017 (0.077) 0.831	-0.017 (0.077) 0.830
Hypertensive disease (no. of IVs:2149, IV strength:0.62)	0.092 (0.049) 0.062	0.051 (0.07) 0.473	0.081 (0.089) 0.363	0.247 (0.133) 0.064	0.244 (0.132) 0.063	0.243 (0.131) 0.063
Type 2 diabetes (no. of IVs:2155, IV strength:0.31)	-0.033 (0.028) 0.247	-0.065 (0.041) 0.109	-0.042 (0.058) 0.475	-0.146 (0.124) 0.239	-0.139 (0.122) 0.254	-0.138 (0.12) 0.251
Body mass index (no. of IVs:1893, IV strength:5.01)	0.366 (0.115) 1e-03 *	0.431 (0.155) 5e-03 *	0.391 (0.206) 0.057	0.439 (0.141) 2e-03 *	0.441 (0.14) 2e-03 *	0.441 (0.14) 2e-03 *

* P<0.05
Estimated causal effect: 0.4, 0.2, 0.0

Figure 2: Estimated causal effects of five obesity-related diseases and traits on (a) COVID-19 infection (b) hospitalized COVID-19 (c) critically ill COVID-19. No IV selection is conducted. In each cell, the first row displays the estimated causal effect and the estimated standard error (in bracket), and the second row displays the P value which is labelled with "*" when less than 0.05 significance level. The pIVW estimator with penalty parameter $\lambda = 1$

significance level, which are concordant with some recent epidemiological studies indicating that hypertensive disease and higher BMI were significantly associated with increased risks of the COVID-19 outcomes (Popkin et al., 2020; Nakeshbandi et al., 2020; Klang et al., 2020; Zhang et al., 2020; de Almeida-Pititto et al., 2020). Some previous MR studies have also found significant causal effects of BMI on the COVID-19 outcomes (Leong et al., 2021; Aung et al., 2020; Freuer et al., 2021; Ponsford et al., 2020), but there is still no MR analysis on the hypertensive disease to the best of our knowledge. Additionally, the pIVW estimator suggests that peripheral vascular disease is significantly associated with higher risks of three COVID-19 outcomes under the IV selection at threshold $\delta = \sqrt{2 \log p}$ (see Supplementary Figure S2.1), which may improve the efficiency of the pIVW estimator. To our knowledge, there also lack epidemiological studies and MR studies on the associations between peripheral vascular disease and the COVID-19 outcomes, despite a high incidence of peripheral vascular disease as complication in COVID-19 patients (García-Ortega et al., 2021; Hanff et al., 2020; Piazza and Morrow, 2020). For type 2 diabetes and dyslipidemia, the pIVW estimator gives no evidence of their associations with the COVID-19 outcomes regardless of the IV selection, while their associations with COVID-19 remain unclear due to discordant findings among previous epidemiological studies and MR studies (Ponsford et al., 2020; Leong et al., 2021; Zheng et al., 2020; Hariyanto and Kurniawan, 2020; Yang et al., 2021; Aung et al., 2020).

Like the pIVW estimator, the competing MR methods also have significant findings for the causal effects of peripheral vascular disease, hypertension disease and BMI on the COVID-19 outcomes. But there are still large differences in the magnitude of estimates among the methods in some cases. In particular, all methods provide very different estimates for peripheral vascular disease, where the IV strength is relatively small ($\hat{\kappa} = 0.04$) indicating the presence of a large amount of weak or null IVs. In this case, the IVW, MR-Egger and MR-Median estimators have much smaller estimates than the pIVW estimator since they might be biased toward zero in the presence of many weak IVs. On the contrary, the RAPS and dIVW estimators generally have larger estimates than the pIVW estimator. In particular, in the peripheral vascular disease and hospitalized COVID-19 data, the dIVW estimator has a relatively large causal effect estimate (4.413) with an extreme estimated standard error (95.942) when no IV selection is conducted. It is possibly because the denominator of the dIVW estimator is very close to zero in the presence many weak IVs. For dyslipidemia ($\hat{\kappa} = 0.82$) and type 2 diabetes ($\hat{\kappa} = 0.31$), all methods have insignificant findings and have similar estimates even in the presence many weak IVs. It is likely because the competing MR methods could be nearly unbiased when the true causal effect is zero as shown in the simulation study. For hypertension disease ($\hat{\kappa} = 0.62$), the RAPS and dIVW estimators have similar estimates to the pIVW estimator, whereas the IVW, MR-Egger and MR-Median estimators have much smaller estimates which are likely to be biased toward zero under many weak IVs. When the IV strength becomes larger under IV selection ($\hat{\kappa}_\delta = 4.41$; see Supplementary Figure S2.1), the estimates of the IVW, MR-Egger and MR-Median estimators are closer to those of the pIVW estimator because their biases are likely to be reduced by the IV selection. For the BMI which has larger IV strength ($\hat{\kappa} = 5.02$), there is smaller discrepancy among the methods. The pIVW estimator also reduces to the dIVW estimator in this case and therefore they provide nearly the same results.

6 Discussion

In MR analysis, the estimation of causal effect can be biased by the presence of weak IVs. The IVW estimator is one of the most popular MR methods but also suffers from substantial bias due to the weak IVs. In this paper, we develop a penalized IVW (pIVW) estimator, where the IVW estimator is adjusted by a proposed penalization method which prevents the denominator of the ratio from being zero in order to decrease the bias and the variance of the estimator in the presence of many weak IVs. Moreover, we allow for the horizontal pleiotropy by adjusting the variance estimation of the pIVW estimator.

The proposed pIVW estimator have some nice features. First, our theoretical and numerical results show that the proposed pIVW estimator can achieve smaller bias and variance compared to the dIVW estimator—another recently proposed extension of the IVW estimator—in the presence of many weak IVs. Second, it is consistent and asymptotically normal even in the presence of many weak IVs, and requires no further assumptions than the dIVW estimator. Third, it has a unique and explicit form of solution, whereas many competing MR methods that are robust to the weak IVs may not have a closed form of solution and may have multiple diverging roots. We also illustrate the superior performance of the pIVW estimator over other popular MR methods through comprehensive simulation study. Our theoretical and numerical results also provide guidance for the choice of the penalty parameter λ in the pIVW estimator, which plays a role in the trade-off between the bias and the variance. We show that the pIVW estimator has smallest bias when $\lambda = 1$ and generally has decreased variance as the increase of λ. Thus, to guarantee the pIVW estimator to achieve smallest bias and valid inference, we recommend to choose the penalty parameter to be 1 in practice. Note that, the dIVW estimator is also a special case of the pIVW estimator when $\lambda = 0$. As the number of IVs and the IV strength increase, the choice of λ tends to have less influence on the performance of the pIVW estimator, and the pIVW estimator can also reduce to the dIVW estimator in this case. Therefore, our pIVW estimator is applicable to more general situations and is recommended for practical use.

Further, we analyze the causal effects of five obesity-related diseases and traits (i.e., peripheral vascular disease, dyslipidemia, hypertensive disease, type 2 diabetes and BMI) on three COVID-19 outcomes (i.e., COVID-19 infection, hospitalized COVID-19 and critically ill COVID-19). We find significant positive causal effects of hypertensive disease on hospitalized COVID-19 and BMI on three COVID-19 outcomes by the pIVW estimator at 0.05 significance level, which are accordant with the findings from some recent epidemiological studies. Additionally, we find significant positive causal effects of peripheral vascular disease on three COVID-19 outcomes under IV selection, while there still lack epidemiological and MR studies on the association between peripheral vascular disease and COVID-19 despite its high incidence in COVID-19 patients. And we find no evidence supporting the associations of type 2 diabetes and dyslipidemia with the COVID-19 outcomes, while their roles on the COVID-19 outcomes also remain unclear due to discordant findings among previous epidemiological and MR studies. Overall, our findings suggest that the preventive strategies aiming at obesity and some related diseases, such as control of body weight and prevention of hypertensive disease and peripheral vascular disease, may help to reduce the risks of COVID-19 infection and severity. On the other hand, comparing to other MR estimators, the proposed pIVW estimator may have smaller bias in the presence of many weak IVs. For instance, the peripheral vascular disease has small IV strengths implying the existence of many weak or null IVs. In this case, the IVW estimator has an estimate very close to zero, but which might be underestimated as shown in the simulation studies. In contrast, the dIVW estimator yields a very large estimate with an extreme standard error, but which might also be inflated because its denominator could be close to zero when many weak IVs exist. The causal effect estimate from the proposed pIVW estimator may be more reliable and have smaller bias, which lies between those of the IVW and dIVW estimators and has much smaller standard error than that of the dIVW estimator.

In future study, it would also be possible to extend the proposed penalization method in the pIVW estimator to more flexible model frameworks for handling the weak IVs and other problems in MR simultaneously, such as the extension to MR-Egger for the unbalanced pleiotropy.

References

Abbott, L., Bryant, S., Churchhouse, C., Ganna, A., Howrigan, D., Palmer, D., et al. (2018). Round 2 gwas results of thousands of phenotypes in the uk biobank. http://www.nealelab.is/uk-biobank/ Accessed November 14, 2018.

Akiyama, M., Okada, Y., Kanai, M., Takahashi, A., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., et al. (2017). Genome-wide association study identifies 112 new loci for body mass index in the japanese population. *Nature genetics*, 49(10):1458.

Andrews, I., Stock, J. H., and Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11:727–753.

Aung, N., Khanji, M. Y., Munroe, P. B., and Petersen, S. E. (2020). Causal inference for genetic obesity, cardiometabolic profile and covid-19 susceptibility: A mendelian randomization study. *Frontiers in genetics*, 11:1417.

Banda, Y., Kvale, M. N., Hoffmann, T. J., Hesselson, S. E., Ranatunga, D., Tang, H., Sabatti, C., Croen, L. A., Dispensa, B. P., Henderson, M., et al. (2015). Characterizing race/ethnicity and genetic ancestry for 100,000 subjects in the genetic epidemiology research on adult health and aging (gera) cohort. *Genetics*, 200(4):1285–1295.

Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N., and Thompson, J. (2017). A framework for the investigation of pleiotropy in two-sample summary data mendelian randomization. *Statistics in medicine*, 36(11):1783–1802.

Bowden, J., Smith, G. D., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *International journal of epidemiology*, 44(2):512–525.

Bowden, J., Smith, G. D., Haycock, P. C., and Burgess, S. (2016). Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genetic epidemiology*, 40(4):304–314.

Burgess, S., Butterworth, A., and Thompson, S. G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology*, 37(7):658–665.

Burgess, S. and Thompson, S. G. (2011). Bias in causal estimates from mendelian randomization studies with weak instruments. *Statistics in medicine*, 30(11):1312–1323.

Burgess, S., Thompson, S. G., and Collaboration, C. C. G. (2011). Avoiding bias from weak instruments in mendelian randomization studies. *International journal of epidemiology*, 40(3):755–764.

Cheng, Q., Yang, Y., Shi, X., Yeung, K. F., Yang, C., Peng, H., and Liu, J. (2020). Mr-ldp: a two-sample mendelian randomization for gwas summary statistics accounting for linkage disequilibrium and horizontal pleiotropy. *NAR Genomics and Bioinformatics*, 2(2):lqaa028.

COVID-19 Host Genetics Initiative (2020). The covid-19 host genetics initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the sars-cov-2 virus pandemic. *European Journal of Human Genetics*, 28(6):715.

COVID-19 Host Genetics Initiative (2021). Mapping the human genetic architecture of covid-19 by worldwide meta-analysis. *MedRxiv*.

Davies, N. M., von Hinke Kessler Scholder, S., Farbmacher, H., Burgess, S., Windmeijer, F., and Smith, G. D. (2015). The many weak instruments problem and mendelian randomization. *Statistics in medicine*, 34(3):454–468.

de Almeida-Pititto, B., Dualib, P. M., Zajdenverg, L., Dantas, J. R., De Souza, F. D., Rodacki, M., and Bertoluci, M. C. (2020). Severity and mortality of covid 19 in patients with diabetes, hypertension and cardiovascular disease: A meta-analysis. *Diabetology & metabolic syndrome*, 12(1):1–12.

Didelez, V. and Sheehan, N. (2007). Mendelian randomization as an instrumental variable approach to causal inference. *Statistical methods in medical research*, 16(4):309–330.

Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, 20(5):533–534.

Freuer, D., Linseisen, J., and Meisinger, C. (2021). Impact of body composition on covid-19 susceptibility and severity: A two-sample multivariable mendelian randomization study. *Metabolism*, 118:154732.

García-Ortega, A., Oscullo, G., Calvillo, P., López-Reyes, R., Méndez, R., Gómez-Olivas, J. D., Bekki, A., Fonfría, C., Trilles-Olaso, L., Zaldívar, E., et al. (2021). Incidence, risk factors, and thrombotic load of pulmonary embolism in patients hospitalized for covid-19 infection. *Journal of Infection*, 82(2):261–269.

Hanff, T. C., Mohareb, A. M., Giri, J., Cohen, J. B., and Chirinos, J. A. (2020). Thrombosis in covid-19. *American Journal of Hematology*, 95(12):1578–1589.

Hariyanto, T. I. and Kurniawan, A. (2020). Dyslipidemia is associated with severe coronavirus disease 2019 (covid-19) infection. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(5):1463–1465.

Hemani, G., Bowden, J., and Davey Smith, G. (2018). Evaluating the potential role of pleiotropy in mendelian randomization studies. *Human molecular genetics*, 27(R2):R195–R208.

Jordan, R. E., Adab, P., and Cheng, K. (2020). Covid-19: risk factors for severe disease and death.

Klang, E., Kassim, G., Soffer, S., Freeman, R., Levin, M. A., and Reich, D. L. (2020). Severe obesity as an independent risk factor for covid-19 mortality in hospitalized patients younger than 50. *Obesity*, 28(9):1595–1599.

Lawlor, D. A. (2016). Commentary: Two-sample mendelian randomization: opportunities and challenges. *International journal of epidemiology*, 45(3):908.

Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N., and Davey Smith, G. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Statistics in medicine*, 27(8):1133–1163.

Leong, A., Cole, J. B., Brenner, L. N., Meigs, J. B., Florez, J. C., and Mercader, J. M. (2021). Cardiometabolic risk factors for covid-19 susceptibility and severity: A mendelian randomization analysis. *PLoS medicine*, 18(3):e1003553.

Marsaglia, G. et al. (2006). Ratios of normal variables. *Journal of Statistical Software*, 16(4):1–10.

Nakeshbandi, M., Maini, R., Daniel, P., Rosengarten, S., Parmar, P., Wilson, C., Kim, J. M., Oommen, A., Mecklenburg, M., Salvani, J., et al. (2020). The impact of obesity on covid-19 complications: a retrospective cohort study. *International journal of obesity*, 44(9):1832–1837.

Nelson, C. R. and Startz, R. (1990). The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one. *Journal of business*, pages S125–S140.

Pascarella, G., Strumia, A., Piliego, C., Bruno, F., Del Buono, R., Costa, F., Scarlata, S., and Agrò, F. E. (2020). Covid-19 diagnosis and management: a comprehensive review. *Journal of internal medicine*, 288(2):192–206.

Piazza, G. and Morrow, D. A. (2020). Diagnosis, management, and pathophysiology of arterial and venous thrombosis in covid-19. *Jama*, 324(24):2548–2549.

Piegorsch, W. W. and Casella, G. (1985). The existence of the first negative moment. *The American Statistician*, 39(1):60–62.

Ponsford, M. J., Gkatzionis, A., Walker, V. M., Grant, A. J., Wootton, R. E., Moore, L. S., Fatumo, S., Mason, A. M., Zuber, V., Willer, C., et al. (2020). Cardiometabolic traits, sepsis, and severe covid-19: A mendelian randomization investigation. *Circulation*, 142(18):1791–1793.

Popkin, B. M., Du, S., Green, W. D., Beck, M. A., Algaith, T., Herbst, C. H., Alsukait, R. F., Alluhidan, M., Alazemi, N., and Shekar, M. (2020). Individuals with obesity and covid-19: A global perspective on the epidemiology and biological relationships. *Obesity Reviews*, 21(11):e13128.

Press, S. J. (1969). The t-ratio distribution. *Journal of the American Statistical Association*, 64(325):242–252.

Sheehan, N. A., Didelez, V., Burton, P. R., and Tobin, M. D. (2008). Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med*, 5(8):e177.

Smith, G. D. and Ebrahim, S. (2003). 'mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *International journal of epidemiology*, 32(1):1–22.

Smith, G. D. and Ebrahim, S. (2004). Mendelian randomization: prospects, potentials, and limitations. *International journal of epidemiology*, 33(1):30–42.

Stefan, N., Birkenfeld, A. L., Schulze, M. B., and Ludwig, D. S. (2020). Obesity and impaired metabolic health in patients with covid-19. *Nature Reviews Endocrinology*, 16(7):341–342.

Vansteelandt, S., Bowden, J., Babanezhad, M., Goetghebeur, E., et al. (2011). On instrumental variables estimation of causal odds ratios. *Statistical Science*, 26(3):403–422.

Verbanck, M., Chen, C. Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. *Nature genetics*, 50(5):693–698.

Wang, P., Xu, S., Wang, Y.-X., Wu, B., Fung, W. K., Gao, G., Liang, Z., and Liu, N. (2020). Penalized fieller's confidence interval for the ratio of bivariate normal means. *Biometrics*.

Yang, H., Hou, H., Liang, X., Xu, J., and Wang, Y. (2021). Lack of significant association between dyslipidemia and covid-19 mortality. *The Journal of Infection*.

Ye, T., Shao, J., and Kang, H. (2019). Debiased inverse-variance weighted estimator in two-sample summary-data mendelian randomization. *arXiv preprint arXiv:1911.09802*.

Zhang, J., Wu, J., Sun, X., Xue, H., Shao, J., Cai, W., Jing, Y., Yue, M., and Dong, C. (2020). Association of hypertension with the severity and fatality of sars-cov-2 infection: A meta-analysis. *Epidemiology & Infection*, 148.

Zhao, Q., Chen, Y., Wang, J., and Small, D. S. (2019). Powerful three-sample genome-wide design and robust statistical inference in summary-data mendelian randomization. *International journal of epidemiology*, 48(5):1478–1492.

Zhao, Q., Wang, J., Hemani, G., Bowden, J., Small, D. S., et al. (2020). Statistical inference in two-sample summary-data mendelian randomization using robust adjusted profile score. *Annals of Statistics*, 48(3):1742–1769.

Zheng, J., Baird, D., Borges, M. C., Bowden, J., Hemani, G., Haycock, P., Evans, D. M., and Smith, G. D. (2017). Recent developments in mendelian randomization studies. *Current epidemiology reports*, 4(4):330–345.

Zheng, Z., Peng, F., Xu, B., Zhao, J., Liu, H., Peng, J., Li, Q., Jiang, C., Zhou, Y., Liu, S., et al. (2020). Risk factors of critical & mortal covid-19 cases: A systematic literature review and meta-analysis. *Journal of Infection*.

Zhu, Z., Zheng, Z., Zhang, F., Wu, Y., Trzaskowski, M., Maier, R., Robinson, M. R., McGrath, J. J., Visscher, P. M., Wray, N. R., et al. (2018). Causal associations between risk factors and common diseases inferred from gwas summary data. *Nature communications*, 9(1):1–12.