



**HAL**  
open science

## Mendelian randomization analysis with survival outcomes

Youngjoo Cho, Andrea Rau, Alex Reiner, Paul Auer

► **To cite this version:**

Youngjoo Cho, Andrea Rau, Alex Reiner, Paul Auer. Mendelian randomization analysis with survival outcomes. *Genetic Epidemiology*, 2021, 45 (1), pp.16-23. 10.1002/gepi.22354 . hal-03155097

**HAL Id: hal-03155097**

**<https://hal.science/hal-03155097>**

Submitted on 28 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Mendelian Randomization Analysis with Survival Outcomes

Youngjoo Cho, Andrea Rau, Alex Reiner, Paul L. Auer

April 20, 2020

## 1 Introduction

Mendelian Randomization (MR) has become a standard approach for evaluating the causal effects of heritable exposures on outcomes of interest (Davey Smith et al. 2003; Burgess et al. 2012; Davey Smith et al. 2014). Often, investigators do not have access to raw genetic data and methods have been developed that allow MRs to be carried out with only summary statistics of exposure-to-instrument and outcome-to-instrument regressions (Burgess et al. 2013). In addition, many MRs are implemented using multiple instruments in order to maximize the power to detect a true causal effect and to protect against directional pleiotropy (e.g., Voight et al. 2012; Day et al. 2015). A common approach to conducting an MR under this paradigm (using multiple instruments with access only to summary statistics), is an inverse-variance weighted (IVW) meta-analysis of estimated causal effects across the instruments. A critical assumption of this approach is the so called "exclusion criteria" which specifies that the instrument cannot be associated with the outcome except through its effect on the exposure (i.e., only an indirect effect through the exposure). When this assumption is met, the IVW method provides an unbiased estimate of the true causal effect of the exposure on the outcome. When this assumption is violated, other more robust methods can be used to estimate the causal effect. One such approach is the often used MR-Egger regression

(Bowden et al. 2015). MR-Egger regression has become a popular approach to conducting MR analyses due to its flexibility in allowing direct effects of instruments on the outcome and providing a robust estimate of the causal effect in the presence of modest violations to the exclusion restriction. An added benefit of the MR-Egger approach is that it too only requires access to summary statistics, i.e., access to raw genotypes is not necessary.

Given the popularity of MR-Egger regression, it has been used in a variety of MRs including when the outcome of interest is failure time. Indeed, previous publications have used MR-Egger regression with summary statistics on proportional hazard models of instruments on failure time (e.g., Guo et al. 2015). Here we show through mathematical derivations and simulations that when the outcome of interest is failure time, a naive analysis can lead to significant bias. We recommend that raw genetic data be accessed in order to assess the proportionality assumption of the standard Cox model (Cox 1972) across instruments. After assessing the validity of this assumption, the appropriate model (an additive hazards model or a proportional hazards model) can be fit to the data and causal inference can proceed. Finally, we provide an application of this principle on data from the Women’s Health Initiative on the causal effect of body mass index (BMI) on breast cancer survival.

## 2 Methods

### 2.1 Causal Inference with survival outcomes and an additive hazard function

In the context of survival data, we let time to event and censoring be  $T$  and  $C$ , respectively, where we observe the minimum,  $\tilde{T} = \min(T, C)$ . We denote a  $p \times 1$  vector of covariates by  $\mathbf{X}$ . As proposed in the Li et al. (2014), we assume an additive hazard model for the exposure and  $T$ . That is, the hazard function is an additive function of covariates.

$$h(t|\mathbf{X}) = h_0(t) + \mathbf{X}^T \boldsymbol{\beta}$$

where  $\boldsymbol{\beta}$  is a  $p \times 1$  vector of regression coefficients. In an MR with multiple genetic variants as instruments, we let  $G_1, \dots, G_J$  represent the  $J$  genetic variants (coded as 0, 1, or 2),  $X_E$  represent the exposure of interest and  $\mathbf{X}_U$  represent a vector of confounders. We assume that  $\mathbf{X}_U$  is unknown and that the following relationships hold:

$$\begin{aligned} X_E &= \gamma_0 + \mathbf{G}^T \boldsymbol{\gamma}_G + \mathbf{X}_U^T \boldsymbol{\gamma}_U + \epsilon^X \\ h(t|\mathbf{X}) &= h_0(t) + \beta_E X_E + \mathbf{X}_U^T \boldsymbol{\beta}_U \end{aligned}$$

When we have a single genetic variant  $j$ , as Li et al. (2014) discussed, we can estimate  $X_E$  with

$$\hat{X}_E = \hat{\gamma}_0 + \mathbf{G}^T \hat{\boldsymbol{\gamma}}_G$$

To carry out a two-stage regression approach to causal inference we then fit an additive hazard model by using  $\hat{X}_E$  with the observed time and censoring indicator. This approach can be extended to multiple genetic variants by plugging in

$$h(t|\mathbf{X}) = h_0(t) + \beta_E (\mathbf{G}^T \boldsymbol{\gamma}_G + \mathbf{X}_U^T \boldsymbol{\gamma}_U + \epsilon^X) + \mathbf{X}_U^T \boldsymbol{\beta}_U$$

Now, consider the survival function:

$$\begin{aligned} S(t|X_E, \mathbf{X}_U, \epsilon^X) &= S_0(t) \exp\{-t(\beta_E (\mathbf{G}^T \boldsymbol{\gamma}_G + \mathbf{X}_U^T \boldsymbol{\gamma}_U + \epsilon^X) + \mathbf{X}_U^T \boldsymbol{\beta}_U)\} \\ &= S_0(t) \exp\{-t(\beta_E \mathbf{G}^T \boldsymbol{\gamma}_G)\} \exp\{-t(\mathbf{X}_U^T (\beta_E \boldsymbol{\gamma}_U + \boldsymbol{\beta}_U) + \beta_E \epsilon^X)\} \end{aligned}$$

By integration with respect to  $(\mathbf{X}_U, \epsilon^X)$ , we divide the survival function given  $X_E$  into two parts : one which depends only on  $X_E$  and another which depends on  $(\mathbf{X}_U, \epsilon^X)$ . Hence from this approach, the additivity of the hazard function is preserved and the effect of interest is  $\beta_E$ .

To summarize, we extend the approach of Li et al. (2014), as follows:

1. Estimate  $X_E$  by fitting a linear regression model on  $G_1, \dots, G_J$ .
2. Fit additive hazard model by using estimate  $X_E$ .

The estimator from this procedure exactly falls into framework from Li et al. (2014), and it is therefore consistent and asymptotically normal.

## 2.2 Causal Inference with survival outcomes and a proportional hazard function

If we assume a proportional hazard model, we have

$$\begin{aligned} X_E &= \gamma_0 + \mathbf{G}^T \boldsymbol{\gamma}_G + \mathbf{X}_U^T \boldsymbol{\gamma}_U + \epsilon^X \\ h(t|\mathbf{X}) &= h_0(t) \exp(\beta_E X_E + \mathbf{X}_U^T \boldsymbol{\beta}_U) \end{aligned}$$

In a two-stage approach, we then plug  $X_E$  into the function  $h(t|\mathbf{X})$ , and get

$$h(t|\mathbf{X}) = h_0(t) \exp\{\beta_E(\gamma_0 + \mathbf{G}^T \boldsymbol{\gamma}_G + \mathbf{X}_U^T \boldsymbol{\gamma}_U + \epsilon^X) + \mathbf{X}_U^T \boldsymbol{\beta}_U\}$$

By integration to obtain the survival function, we get

$$S(t|\mathbf{X}) = \{S_0(t)\}^{\exp(-\beta_E(\gamma_0 + \mathbf{G}^T \boldsymbol{\gamma}_G + \mathbf{X}_U^T \boldsymbol{\gamma}_U + \epsilon^X) + \mathbf{X}_U^T \boldsymbol{\beta}_U)}$$

In this case, unlike the additive hazard model, we cannot separate the part only depending on  $X_E$  although we integrate out  $\mathbf{X}_U$ . Therefore, the proportionality of the hazard function is not preserved with respect to  $X_E$  and  $\beta_E$  is not the causal effect.

## 2.3 Extensions to MR-egger regression

In the context of MR-egger regression, we assume that the hazard function is related to  $G_j$  as well as  $\beta_E$  and  $\mathbf{X}_U$  as follows:

$$\begin{aligned} X_E &= \gamma_{0j}^* + \gamma_{1j}^* G_j + \boldsymbol{\gamma}_U^{*T} \mathbf{X}_U + \epsilon^{X*} \\ h(t) &= h_0(t) + \alpha_j^* G_j + \beta_E X_E + \boldsymbol{\beta}_U^T \mathbf{X}_U \end{aligned}$$

where  $\epsilon^{X*}$  is a new error term in the new linear regression with respect to  $G_j$ . In this case, other than unmeasured confounders, we focus on the relation between individual genetic variant and the exposure, and between the hazard and the exposure with individual genetic variant. Then by plugging in  $X_E$  into  $h(t)$ , we obtain

$$\begin{aligned} h(t) &= h_0(t) + \alpha_j^* G_j + \beta_E(\gamma_{0j}^* + \gamma_{1j}^* G_j + \boldsymbol{\gamma}_U^{*T} \mathbf{X}_U + \epsilon^{X*}) + \boldsymbol{\beta}_U^T \mathbf{X}_U \\ &= h_0(t) + G_j(\alpha_j^* + \gamma_{1j}^* \beta_E) + \epsilon^* \end{aligned}$$

where  $\epsilon^* = \beta_E \gamma_{0j}^* + \beta_E \beta_U^T \mathbf{X}_U + \gamma_U^{*T} \mathbf{X}_U + \epsilon^{*X}$ . By two-stage modeling, we obtain  $\frac{\alpha_j^* + \beta_E \gamma_{1j}^*}{\gamma_{1j}^*} = \beta_E + \frac{\alpha_j^*}{\gamma_{1j}^*}$ . We can see from this that when genetic variants are valid instrument variables, they satisfy the following conditions (Bowen et al. 2015)

1. Each  $G_j$  and  $\mathbf{X}_U$  are independent.
2. Each  $G_j$  is associated with  $X_E$ .
3. Each  $G_j$  is independent of  $(T, C)$  conditional on the  $X_E$  and  $\mathbf{X}_U$ .

In this case,  $\alpha_j^* = 0$ , so the causal effect is  $\beta_E$ .

This method is particularly useful. For censored data, suppose we have:

- Summary statistics from running an additive hazard model for failure time on each genetic variant, i.e., We run the additive hazard model to estimate  $\Gamma_j = \alpha_j^* + \beta_E \gamma_{1j}^*$ . Denote the statistics by  $\hat{\Gamma}_{G_j}$  and its estimated standard error as  $\hat{s}_{\hat{\Gamma}_{G_j}}$ .
- Summary statistics from running a simple linear regression model of exposure on each genetic variant. Denote the statistic for the estimated slope as  $\hat{\gamma}_{G_j}$

The estimation procedure for regression coefficients for time-independent covariates is the least-square method (Lin and Ying, 1994; Li et al. 2014). Hence by employing an additive hazard model, we can follow estimation procedures for MRs in censored data. As we discussed, due to additivity and linearity of our model, we can apply the inverse variance weighting method (Johnson, 2013; Bowen et al. 2015). Let  $\hat{\gamma}_j$  be

$$\frac{\sum_{j=1}^J \hat{\gamma}_{G_j} \hat{s}_{\hat{\Gamma}_{G_j}}^{-2} \hat{\Gamma}_{G_j}}{\sum_{j=1}^J \hat{\gamma}_{G_j}^2 \hat{s}_{\hat{\Gamma}_{G_j}}^{-2}}$$

However, this inverse weighting method has bias when  $\alpha_j^* \neq 0$  (i.e., directional pleiotropy). As discussed in Bowen et al. (2015), this meta-analysis framework is applicable to an additive hazard model. As in Bowen et al. (2015), for each  $j = 1, \dots, J$ ,

$$\hat{\Gamma}_{G_j}^{MR} = \beta_{0E} + \beta_E \hat{\gamma}_{G_j}^{MR}$$

Then by estimation of  $\beta_E$ , we obtain the estimated causal effect. The intercept  $\beta_{0E}$  estimates the overall directional pleiotropy (Bowen et al. 2015)

Due to the linearity and additivity of the additive hazard model, our additive hazard model has the same characteristics as a linear regression model for uncensored data. In other words, our estimate of  $\beta_E$  is the causal effect when  $\beta_{0E} = 0$ .

Under a Cox proportional hazard model,

$$\begin{aligned} & h_0(t) \exp(\alpha_j G_j + \beta_E X_E + \boldsymbol{\beta}_U^T \mathbf{U}) \\ &= h_0(t) \exp(\alpha_j G_j + \beta_E (\gamma_{0j}^* + \gamma_{1j}^* G_j + \boldsymbol{\gamma}_U^{*T} \mathbf{X}_U + \epsilon^{*X}) + \boldsymbol{\beta}_U^T \mathbf{U}) \end{aligned}$$

Due to multiplicity of the proportional hazard model, as argued in Li et al. (2014),  $\beta_E$  does not represent the causal effect. Therefore the IVW and MR-egger methods based on a Cox model do not provide valid procedures for estimating the causal effect of  $\beta_E$ .

### 3 Simulation study

#### 3.1 Models and parameter settings

To evaluate the performance of the estimation procedures under a variety of settings, let  $G_j$  represent the  $j$ th genetic variant drawn from a multinomial distribution  $((0,1,2); (0.25, 0.5, 0.25))$  with  $\gamma_j$  drawn from  $\text{Uniform}(0.5, 4)$ . Let  $\alpha_j$  also be a fixed value that is generated from

1. Case 1 :  $\alpha_j = 0$  for all  $j$ , i.e., no pleiotropy.
2. Case 2:  $\alpha_j \sim N(0.25, 0.1^2)$  for all  $j$ , pleiotropy.

$U$  and  $\epsilon^X$  are independently generated from  $N(0, 4)$ . Let  $\gamma_0 = 0$ . We generate data from an additive hazard as:

$$X = \gamma_0 + \sum_{j=1}^J \gamma_j G_j + U + \epsilon^X$$

$$h(t) = h_0(t) + \sum_{j=1}^J \alpha_j G_j + \beta_E X + U$$

We set  $h_0(t) = 5$  which implies a constant hazard. In this scenario, we fit an additive hazard model. We investigated four different causal effect sizes  $\beta_E = 0, 0.1, 0.5, 1$ . In this section, we display results from  $\beta_E = 0.1$  and  $\beta_E = 0.5$ , with results from  $\beta_E = 0$  and  $\beta_E = 1$  in the Appendix.

We also generated failure times from the Cox model in order to investigate the properties of estimates obtained by fitting a Cox model. The posited model is

$$X = \gamma_0 + \sum_{j=1}^J \gamma_j^* G_j + U^* + \epsilon^{X^*}$$

$$h(t) = \exp\left(\sum_{j=1}^J \alpha_j^* G_j + \beta_E X + U^*\right)$$

where  $\gamma_j \sim \text{Uniform}(0, 0.5), j = 1, \dots, 25$ ,  $U^*, \epsilon^* \sim N(0, 1)$  and  $U^*$  and  $\epsilon^*$  are independently generated. We considered causal effects when  $\beta_E = 0$  and  $\beta_E = 0.5$ . In this section, we display result from  $\beta_E = 0.5$  with result from  $\beta_E = 0$  in the Appendix. Let  $\alpha_j^*$  also be a fixed value that is generated from

1. Case 1\* :  $\alpha_j^* = 0$  for all  $j$ , i.e., no pleiotropy.
2. Case 2\* :  $\alpha_j^* \sim N(0.1, 0.05^2)$  for all  $j$ , pleiotropy.

We generate time by using a probability integral transformation. Let  $U^* \sim \text{Unif}(0, 1)$  and  $S(t)$  be the survival function. Note that

$$S(t) = e^{-H(t)}$$



where  $H(t)$  is the cumulative hazard function. Since the survival function also follows  $Unif(0, 1)$ , the failure time is generated from

$$T = -\frac{\log U^*}{H(t)}$$

For the additive hazard model, for both Case 1 and Case 2, the censoring variable  $C$  is generated from an exponential distribution with rate 2. For the Cox model, for Case 1\* and Case 2\*,  $C$  is generated from an exponential distribution with rate 10 and 50, respectively.

We compared the performance of 3 different models: 2-stage regression as in Li et al. (2014) (Multi), the IVW method, and the MR-Egger method for both additive hazard and Cox models. We considered two different sample sizes,  $n = 250$  and  $n = 500$ .

## 3.2 Results

Under our simulation settings, we calculated bias (Bias), the empirical standard deviation (Emp), standard error (SE), and 95% coverage (Coverage, based on the Wald method), and power (evaluated at a significance level of 0.05) When  $\beta_E = 0.1$ , the censoring rate for Case 1 is approximately 29% and for Case 2 is approximately 28%. For  $\beta_E = 0.5$ , the censoring rate for Case 1 is approximately 12% and Case 2 is approximately 14%.

Tables 1 and 2 shows the results for  $\beta_E = 0.1$  and  $\beta_E = 0.5$ , respectively, for the additive hazard model. When conditions 1, 2 and 3 are satisfied (denoted Case 1), then Multi, IVW, and Egger all display good performance with coverage close to 0.95 and small amounts of bias. Note that Multi and IVW are both more efficient than Egger, under Case 1. When the conditions for valid instruments are not satisfied (denoted Case 2), Multi and IVW display significant bias and poor coverage at the nominal level. Note that Egger provides smaller bias and correct coverage in the nominal level though it is less efficient than Multi and IVW. Under the null (Table S1), Egger is the only method that consistently provides correct coverage under all settings with the appropriate Type I error control.

Table 3 shows the result for  $\beta_E = 0.5$  for the Cox model. The censoring rate is approximately 42% for Case 1\* and 33% for Case 2\*. Although we fit the correct model, there is

substantial bias for all three methods. Interestingly, in Case 1\*, the Egger method has lowest bias while bias of the Multi method is the lowest one in Case 2\*. Under the null (Table S3) and no pleiotropy, IVW and Multi perform the best in terms of coverage and Type I error control. However, under the null with pleiotropy, Egger performs the best.

## 4 Data analysis

As a case study in the context of real data, we considered BMI as a potential causal factor for breast cancer (BC) survival (Guo et al. 2015). Using data from the Women’s Health Initiative (WHI), we considered a total of 1367 BC patients. The WHI is a prospective cohort study and one of the largest studies of women’s health ever conducted in the U.S. Recruitment for the WHI began in 1993 and ended in 1998 and was conducted by 40 Clinical Centers in 24 states and the District of Columbia. There are two major components of WHI: (1) a clinical trial (CT) that enrolled and randomized 68,132 women ages 50-79 into at least one of three placebo control clinical trials (hormone therapy, dietary modification, and supplementation with calcium and vitamin D); and (2) an observational study (OS) that enrolled 93,676 women of the same age range into a parallel prospective cohort study (Prentice et al. 1998). All of the 1367 BC patients from the WHI were selected from the OS. Of the 1367 patients, 484 died due to breast cancer and the remaining 883 were alive at the time of the last follow-up.

We used the same 94 SNPs as in Guo et al. (2015) as instruments for BMI. First, we ran a cox proportional hazards model independently for each SNP against survival status and considered the global goodness-of-fit test as described in Grambsch and Therneau (1994). The cox model showed poor goodness of fit across the 94 SNPs ( $p=0.0122$ ), so we also ran the Aalen additive hazards model between each SNP and survival status. To implement the MR-Egger causal estimation procedure, we also ran linear models between each SNP and BMI. Finally, we fit the MR-Egger regression line to obtain causal effects for both the Cox model and the additive hazard. Both the estimate based on the additive hazard ( $\beta =$

0.00096,  $p=0.7018$ ) and the proportional hazard ( $\beta = 0.0086$ ,  $p = 0.557$ ) provided non-statistically significant results for the causal effect of BMI on BC survival. However, Figure 1 shows the differences in the fitted MR-Egger lines. Note that the appropriate model for these data appears to be the additive hazard (based on the GoF test) and the estimated effects differ by an order of magnitude.

## 5 Discussion

In Burgess and Thompson (2017), the authors comprehensively discuss the use of MR-Egger regression and the importance of assessing the assumptions inherent to that model. Specifically, they draw attention to the linearity assumption in MR-Egger regression and classify this as a '2nd order' issue as it does not effect valid *inferences* on causal effects but may effect valid *estimates* of causal effects. The issue we raise in this article is exactly this type of 2nd order concern, i.e., obtaining valid estimates of causal effects when the outcome of interest is failure time. In this context the additive hazard preserves linearity and the resulting causal effect estimates from MR-Egger regression are valid. Though the Cox model does not preserve linearity, it is still the appropriate model when the data generating mechanism is a proportional hazard. In such cases, causal inferences can proceed with the understanding that the causal effect estimates can be biased (as our simulations indicate).

It is important to note that contemporary MRs are often performed on publicly available summary statistics from large-scale genetics consortia. However, in the context of a MR analysis with censored survival times as the outcome of interest, summary statistics do not suffice to assess the appropriateness of the proportional hazards assumption. In these contexts, we recommend only using data for which genotypes and phenotypes are available so that the analyst can run the Grambsch and Therneau (1994)'s GoF test, to assess whether the proportionality assumptions is met. Then the appropriate model (additive vs proportional hazard) can be fit to the data and MR-Egger inference can proceed.

## 6 Acknowledgements

We thank Dr. Cheng Zheng for helpful discussions. We also acknowledge the participants in the Women's Health Initiative study. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C.

## References

- [1] Davey Smith, G., Ebrahim, S. (2003) Mendelian randomization: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 32:, 122.
- [2] Burgess S., Butterworth A., Malarstig A., Thompson S. (2012) Use of Mendelian randomisation to assess potential benefit of clinical intervention. *BMJ* 345, e7325.
- [3] Davey Smith G. Hemani G. (2014) Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 23, 8998.
- [4] Burgess S., Butterworth A., Thompson S. (2013) Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 37, 65865.
- [5] Voight, B.F., Peloso, G.M., Orho-Melander, et al. (2012) Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* 380, 572-580.
- [6] Day, F.R., Ruth, K.S., Thompson, D.J., et al. (2015) Large-scale genomic analyses link reproductive ageing to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nature Genetics* 47, 12941303.
- [7] Cox, D.R. (1972) Regression Models and Life-Tables. *Journal of the Royal Statistical Society, Series B.* 34, 187-220.

- [8] Li, J., Fine, J. and Brookhart, A. (2015). Instrumental variable additive hazards models. *Biometrics*, 71, 122-130.
- [9] Guo, Q., Burgess, S., Turman, C., Bolla, M. K., Wang, Q., Lush, M., Abraham, J. et al. (2017). Body mass index and breast cancer survival: a Mendelian randomization analysis. *International Journal of Epidemiology*, 46, 1814-1822.
- [10] Bowden, J., Davey Smith, G. and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International Journal of Epidemiology*, 44, 512-525.
- [11] Johnson, T. (2015). Efficient calculation for multi-SNP genetic risk scores: presented at the American Society of Human Genetics Annual Meeting, San Francisco, November 6-10, 2012. <https://rdr.io/cran/gtx/f/inst/doc/ashg2012.pdf> [16 April 2020, date last accessed].
- [12] Lin, D. Y. and Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika*, 81, 61-71.
- [13] Prentice, R., Rossouw, J., Furberg, C., et al. (1998) Design of the WHI Clinical Trial and Observational Study. *Control Clin Trials* 19, 61-109.
- [14] P. Grambsch and T. Therneau (1994), Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81, 515-526.
- [15] Burgess, S. and Thompson, S.G. (2017) Interpreting findings from Mendelian randomization using the MR-Egger method. *European Journal of Epidemiology* 32,377-389.

## 7 Tables

			Bias	Emp	SE	Cover	Power
250	Case 1	Multi	0.01	0.093	0.095	0.952	0.214
		IVW	0.008	0.094	0.107	0.968	0.164
		Egger	0.027	0.194	0.211	0.958	0.084
	Case 2	Multi	0.101	0.147	0.149	0.888	0.256
		IVW	0.098	0.149	0.169	0.92	0.204
		Egger	0.042	0.31	0.333	0.958	0.068
500	Case 1	Multi	0.003	0.064	0.066	0.948	0.344
		IVW	0.003	0.067	0.071	0.954	0.284
		Egger	0.012	0.151	0.156	0.952	0.106
	Case 2	Multi	0.091	0.1	0.104	0.86	0.464
		IVW	0.091	0.105	0.112	0.874	0.39
		Egger	0.014	0.236	0.247	0.962	0.06

Table 1: Simulation results from data generated from additive hazard model when  $\beta_E = 0.1$ .  
Multi : adaption of Li et al. (2014) IVW : inverse variance weighting Egger : adaption of  
Egger regression

			Bias	Emp	SE	Cover	Power
250	Case 1	Multi	0.018	0.256	0.262	0.948	0.512
		IVW	0.008	0.263	0.293	0.966	0.422
		Egger	0.005	0.537	0.578	0.96	0.122
	Case 2	Multi	0.106	0.309	0.315	0.94	0.486
		IVW	0.096	0.316	0.353	0.954	0.4
		Egger	0.021	0.65	0.697	0.962	0.104
500	Case 1	Multi	0.003	0.181	0.183	0.94	0.792
		IVW	0.003	0.186	0.195	0.948	0.744
		Egger	0.01	0.401	0.427	0.956	0.222
	Case 2	Multi	0.092	0.219	0.219	0.928	0.774
		IVW	0.092	0.226	0.234	0.94	0.714
		Egger	0.016	0.49	0.514	0.96	0.168

Table 2: Simulation results from data generated from additive hazard model when  $\beta_E = 0.5$ . Multi : adaption of Li et al. (2014) IVW : inverse variance weighting Egger : adaption of Egger regression

			Bias	Emp	SE	Cover	Power
250	Case 1*	Multi	-0.157	0.073	0.077	0.47	0.98
		IVW	-0.175	0.067	0.075	0.322	0.974
		Egger	-0.116	0.104	0.108	0.832	0.932
	Case 2*	Multi	-0.046	0.072	0.074	0.888	1
		IVW	-0.084	0.06	0.066	0.784	1
		Egger	-0.141	0.096	0.099	0.714	0.938
500	Case 1*	Multi	-0.183	0.056	0.055	0.09	1
		IVW	-0.195	0.053	0.053	0.038	1
		Egger	-0.15	0.09	0.086	0.592	0.964
	Case 2*	Multi	-0.06	0.052	0.053	0.786	1
		IVW	-0.091	0.046	0.047	0.476	1
		Egger	-0.184	0.081	0.08	0.338	0.956

Table 3: Simulation results from data generated from Cox model with  $\beta_E = 0.5$ . Multi : adaption of Li et al. (2014) IVW : inverse variance weighting Egger : adaption of Egger regression



## 8 Figure Legends

Figure 1: Causal diagram showing instrumental variable mechanism (Bowden et al. 2015)

Figure 2: Results from MR-Egger regression investigating the causal effect of BMI on Breast Cancer survival. SNP-BMI effects are represented on the x-axis, SNP-BC survival effects are represented on the y-axis. Results from both an additive hazard (blue) and a proportional hazard (black) are shown.

## Appendix

### A Additional simulation results

This section describes results from  $\beta_E = 0$  and  $\beta_E = 1$  from additive hazard model, and  $\beta_E = 0$  from Cox model. In additive hazard model, when  $\beta_E = 0$ , (i.e., no causal effect) the censoring rate is approximately 46% for Case 1 and 37% for Case 2. For  $\beta_E = 1$ , censoring rate for Case 1 is approximately 7% and Case 2 is approximately 9%. When  $\beta_E = 0$  and Case 1 (i.e., no pleiotropy and no causal effect), fitting additive hazard model show good performance with proper coverage at the nominal level. For Case 2 and no causal effect, the Multi and IVW estimators show poor coverage. However, MR-egger provides reasonable bias and good coverage. When  $\beta_E = 1$ , results are very similar to  $\beta_E = 0.1$  and  $\beta_E = 0.5$ . In Case 1, all three methods perform well. In Case 2, MR-Egger has the smallest bias and good coverage, though it is still less efficient than Multi and IVW.

For data generated from Cox model, in  $\beta_E = 0$ , censoring rates for Case 1\* and Case 2\* are approximately 88% and 80%, respectively. For bias point of view, in the Case 1\*, Egger has the worst performance while it is the best performance in Case 2\*. For Case 1\*, when  $n = 500$ , Multi and IVW are slightly biased but their coverage and power are closed to nominal level.

			Bias	Emp	SE	Cover	Power
250	Case 1	Multi	0.006	0.05	0.052	0.944	0.056
		IVW	0.005	0.053	0.059	0.95	0.05
		Egger	0.018	0.106	0.116	0.948	0.052
	Case 2	Multi	0.1	0.104	0.108	0.844	0.156
		IVW	0.097	0.106	0.122	0.878	0.122
		Egger	0.035	0.216	0.24	0.964	0.036
500	Case 1	Multi	0.005	0.036	0.036	0.946	0.054
		IVW	0.005	0.037	0.039	0.95	0.05
		Egger	0.013	0.084	0.086	0.952	0.048
	Case 2	Multi	0.093	0.072	0.076	0.776	0.224
		IVW	0.093	0.076	0.082	0.796	0.204
		Egger	0.013	0.174	0.178	0.944	0.056

Table S1: Simulation results from data generated from additive hazard model when  $\beta_E = 0$ . Multi : adaption of Li et al. (2014) IVW : inverse variance weighting Egger : adaption of Egger regression

			Bias	Emp	SE	Cover	Power
250	Case 1	Multi	0.025	0.462	0.466	0.946	0.608
		IVW	0.004	0.47	0.519	0.972	0.516
		Egger	-0.002	0.948	1.028	0.972	0.166
	Case 2	Multi	0.123	0.507	0.521	0.952	0.596
		IVW	0.102	0.519	0.581	0.958	0.516
		Egger	0.013	1.067	1.148	0.966	0.14
500	Case 1	Multi	0.0004	0.324	0.324	0.938	0.876
		IVW	-0.001	0.331	0.345	0.944	0.83
		Egger	0.008	0.715	0.76	0.96	0.252
	Case 2	Multi	0.085	0.365	0.363	0.924	0.854
		IVW	0.085	0.374	0.385	0.932	0.8
		Egger	0.011	0.804	0.849	0.96	0.226

Table S2: Simulation results when  $\beta_E = 1$ . Multi : adaption of Li et al. (2014) IVW : inverse variance weighting Egger : adaption of Egger regression

			Bias	Emp	SE	Cover	Power
250	Case 1*	Multi	0.084	0.188	0.187	0.916	0.084
		IVW	0.078	0.192	0.203	0.932	0.068
		Egger	0.165	0.316	0.308	0.894	0.106
	Case 2*	Multi	0.268	0.141	0.145	0.532	0.468
		IVW	0.259	0.142	0.157	0.614	0.386
		Egger	0.155	0.232	0.237	0.88	0.12
500	Case 1*	Multi	0.046	0.129	0.135	0.948	0.052
		IVW	0.045	0.131	0.143	0.958	0.042
		Egger	0.111	0.237	0.237	0.924	0.076
	Case 2*	Multi	0.25	0.101	0.106	0.312	0.688
		IVW	0.247	0.104	0.111	0.368	0.632
		Egger	0.09	0.185	0.186	0.906	0.094

Table S3: Simulation results from data generated from Cox model when  $\beta_E = 0$ . Multi : adaption of Li et al. (2014) IVW : inverse variance weighting Egger : adaption of Egger regression

