Supplementary to “Robust Inference of Bi-Directional Causal Relationships in Presence of Correlated Pleiotropy with GWAS Summary Data”

Haoran Xue¹ and Wei Pan¹, ²

¹Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, Minnesota 55455.
²Corresponding author. Email: panxx014@umn.edu. Phone: 612-624-4655. Fax: 612-626-0660.

Contents

S1 Full Simulation Results

S2 Full Real Data Results
  S2.1 48 Risk Factor-Disease Pairs
  S2.2 Pairs of 4 Diseases
  S2.3 Links to GWAS Summary Datasets

S3 Theoretical Results
  S3.1 Proof of Theorem 1
  S3.2 Proof of Theorem 2
  S3.3 MR-cML with Data Perturbation
S1 Full Simulation Results

S1 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S2 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

\[
\theta_{XY} = 0, \xi \sim \text{Unif}(-0.2,0.2), \text{ X Continuous, Y Continuous}
\]
S3 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0.02$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \to Y$ (left column) and $Y \to X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S4 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0.02$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S5 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0.1$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S6 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0.1$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

\[ \theta_{XY} = 0.1, \xi \sim \text{Unif}(-0.2,0.2), X \text{ Continuous, } Y \text{ Continuous} \]
S7 Fig: When both X and Y are continuous, $\theta_{XY} = 0.2$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S8 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0.2$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

$\theta_{XY} = 0.2$, $\xi \sim \text{Unif}(-0.2,0.2)$, $X$ Continuous, $Y$ Continuous
S9 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0.3$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

$\theta_{XY} = 0.3$, $\xi = 0$, $X$ Continuous, $Y$ Continuous
S10 Fig: When both $X$ and $Y$ are continuous, $\theta_{XY} = 0.3$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S11 Fig: When $X$ is binary, $Y$ is continuous, $\theta_{XY} = 0$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S12 Fig: When $X$ is binary, $Y$ is continuous, $\theta_{XY} = 0$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

\[ \theta_{XY} = 0, \xi \sim \text{Unif}(-0.2,0.2), \text{X Binary, Y Continuous} \]
S13 Fig: When \( X \) is binary, \( Y \) is continuous, \( \theta_{XY} = 0.02 \) and \( \xi = 0 \), the proportions of significant simulation results obtained by the methods for direction \( X \rightarrow Y \) (left column) and \( Y \rightarrow X \) (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S14 Fig: When $X$ is binary, $Y$ is continuous, $\theta_{XY} = 0.02$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

\[ \theta_{XY} = 0.02, \xi \sim \text{Unif}(-0.2,0.2), X \text{ Binary}, Y \text{ Continuous} \]
S15 Fig: When $X$ is binary, $Y$ is continuous, $\theta_{XY} = 0.1$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S16 Fig: When $X$ is binary, $Y$ is continuous, $\theta_{XY} = 0.1$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S17 Fig: When $X$ is binary, $Y$ is continuous, $\theta_{XY} = 0.2$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S18 Fig: When $X$ is binary, $Y$ is continuous, $\theta_{XY} = 0.2$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S19 Fig: When $X$ is binary, $Y$ is continuous, $\theta_{XY} = 0.3$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S20 Fig: When \( X \) is binary, \( Y \) is continuous, \( \theta_{XY} = 0.3 \) and \( \xi \sim \text{Unif}(-0.2,0.2) \), the proportions of significant simulation results obtained by the methods for direction \( X \to Y \) (left column) and \( Y \to X \) (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S21 Fig: When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S22 Fig: When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
When X is continuous, Y is binary, $\theta_{XY} = 0.02$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S24 Fig: When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0.02$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S25 Fig: When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0.1$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

\[ \theta_{XY} = 0.1, \xi = 0, X \text{ Continuous, } Y \text{ Binary} \]
S26 Fig: When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0.1$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \to Y$ (left column) and $Y \to X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S27 Fig: When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0.2$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

\[
\theta_{XY} = 0.2, \xi = 0, X \text{ Continuous, } Y \text{ Binary}
\]
S28 Fig: When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0.2$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0.3$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S30 Fig: When $X$ is continuous, $Y$ is binary, $\theta_{XY} = 0.3$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

$\theta_{XY} = 0.3, \xi \sim \text{Unif}(-0.2,0.2), X \text{ Continuous, Y Binary}$
S31 Fig: When both $X$ and $Y$ are binary, $\theta_{XY} = 0$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S32 Fig: When both $X$ and $Y$ are binary, $\theta_{XY} = 0$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S33 Fig: When both $X$ and $Y$ are binary, $\theta_{XY} = 0.02$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S34 Fig: When both $X$ and $Y$ are binary, $\theta_{XY} = 0.02$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S35 Fig: When both $X$ and $Y$ are binary, $\theta_{XY} = 0.1$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.

$\theta_{XY} = 0.1$, $\xi = 0$, $X$ Binary, $Y$ Binary

![Graph showing proportions of significant results for different methods and directions.](image-url)
S36 Fig: When both $X$ and $Y$ are binary, $\theta_{XY} = 0.1$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \to Y$ (left column) and $Y \to X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
When both $X$ and $Y$ are binary, $\theta_{XY} = 0.2$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S38 Fig: When both $X$ and $Y$ are binary, $\theta_{XY} = 0.2$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
S39 Fig: When both $X$ and $Y$ are binary, $\theta_{XY} = 0.3$ and $\xi = 0$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
When both $X$ and $Y$ are binary, $\theta_{XY} = 0.3$ and $\xi \sim \text{Unif}(-0.2,0.2)$, the proportions of significant simulation results obtained by the methods for direction $X \rightarrow Y$ (left column) and $Y \rightarrow X$ (right column). The first row shows results for four main methods: MR-cML-DP-S, CD-cML-DP-S, CD-Ratio-S, and CD-Egger-S; the second row shows results for four methods without screening: MR-cML-DP, CD-cML-DP, CD-Ratio, and CD-Egger; the third row shows results for other five methods.
## S2  Full Real Data Results

### S2.1  48 Risk Factor-Disease Pairs

S1 Table: Inferring causal effects between first 6 risk factors and CAD. In each cell we show the Bonferroni adjusted 1-0.05/96 \approx 0.9995 confidence intervals (CIs) of $\theta$ for the MR methods, and CIs of $K$ for the CD methods; for Steiger’s method, we show the proportion of SNPs giving significant results. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.
S2 Table: Inferring causal effects between second 6 risk factors and CAD. In each cell we show the Bonferroni adjusted 1-0.05/96 = 0.9995 confidence intervals (CIs) of \( \theta \) for the MR methods, and CIs of \( K \) for the CD methods; for Steiger’s method, we show the proportion of SNPs giving significant result. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.

| Method          | Direction | BW to CAD | CAD to CAD | DBP to CAD | DBP to DBP | SBP to CAD | SBP to DBP | FG to CAD | CAD to FG | Smoke to CAD | CAD to Smoke | Alcohol to CAD | CAD to Alcohol |
|-----------------|-----------|-----------|------------|------------|------------|------------|------------|-----------|-----------|--------------|---------------|----------------|----------------|
| MR-CML-DP-S     | TRUE      | 0.057     | 0.968      | 0.009      | 0.603      | TRUE       | TRUE       | 0.129     | 0.131     | FALSE        | FALSE        | FALSE          | TRUE           |
|                 | FALSE     | 0.027     | 0.997      | 0.068      | 0.074      | FALSE      | TRUE       | 0.369     | 0.265     | FALSE        | TRUE          | FALSE          | TRUE           |
| MR-CML-S        | TRUE      | 0.026     | 0.009      | 0.059      | 0.076      | TRUE       | TRUE       | 0.038     | 0.055     | TRUE         | FALSE        | TRUE           | TRUE           |
|                 | FALSE     | 0.025     | 0.084      | 0.047      | 0.055      | FALSE      | TRUE       | 0.537     | 0.357     | FALSE        | TRUE          | TRUE           | TRUE           |
| CD-3ML-DP-S     | FALSE     | 0.002     | 0.043      | 0.196      | 0.272      | FALSE      | FALSE      | 0.130     | 0.230     | TRUE         | TRUE          | TRUE           | TRUE           |
|                 | TRUE      | 0.012     | 0.008      | 0.024      | 0.061      | TRUE       | TRUE       | 0.009     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| CD-3ML-S        | FALSE     | 0.002     | 0.012      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.012      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| CD-Ratio-S      | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| CD-Egger-S      | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| MR-CML-DP       | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| MR-CML          | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| CD-CML-DP       | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| CD-CML          | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| CD-Ratio       | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| CD-Egger       | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
| Steiger        | FALSE     | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
|                 | TRUE      | 0.002     | 0.009      | 0.014      | 0.029      | FALSE      | FALSE      | 0.012     | 0.014     | FALSE        | FALSE        | FALSE          | FALSE          |
S3 Table: Inferring causal effects between first 6 risk factors and Stroke. In each cell we show the Bonferroni adjusted 1-0.05/96 ≈ 0.9995 confidence intervals (CIs) of θ for the MR methods, and CIs of K for the CD methods; for Steiger’s method, we show the proportion of SNPs giving significant result. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.

| Method          | TG to Stroke | Stroke to TG | LDL to Stroke | Stroke to LDL | HDL to Stroke | Stroke to HDL | Height to Stroke | Stroke to Height | BMI to Stroke | Stroke to BMI | BF to Stroke | Stroke to BF |
|-----------------|--------------|--------------|---------------|---------------|---------------|---------------|-----------------|-----------------|--------------|---------------|--------------|--------------|
| MR-ML-DP-S      | c(0.169, 0.179) | FALSE        | 0.098         | 0.188         | 0.014         | 0.111         | 0.082           | 0.137           | 0.014        | 0.099        | 0.011        | 0.451        |
| MR-ML-S         | c(0.198, 0.217) | FALSE        | 0.188         | 0.083         | 0.089         | 0.095         | 0.064           | 0.054           | 0.015        | 0.098        | 0.013        | 0.451        |
| CD-ML-DP-S      | c(0.022, 0.023) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| CD-ML-S         | c(0.018, 0.021) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| CD-Ratio-S      | c(0.018, 0.021) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| CD-Ratio        | c(0.018, 0.021) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| CD-Egger-S      | c(0.035, 0.062) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| CD-Egger        | c(0.035, 0.062) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| MR-ML-DP        | c(0.169, 0.179) | FALSE        | 0.098         | 0.188         | 0.014         | 0.111         | 0.082           | 0.137           | 0.014        | 0.099        | 0.011        | 0.451        |
| MR-ML           | c(0.088, 0.155) | FALSE        | 0.188         | 0.083         | 0.083         | 0.083         | 0.064           | 0.054           | 0.003        | 0.085        | 0.006        | 0.451        |
| CD-ML-DP        | c(0.022, 0.023) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| CD-ML           | c(0.022, 0.023) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| CD              | c(0.022, 0.023) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| CD-Egger        | c(0.018, 0.018) | FALSE        | 0.098         | 0.029         | 0.029         | 0.029         | 0.005           | 0.018           | 0.018        | 0.018        | 0.018        | 0.018        |
| LHC-ML          | c(0.589, 0.640) | FALSE        | 0.188         | 0.083         | 0.083         | 0.083         | 0.064           | 0.054           | 0.015        | 0.098        | 0.006        | 0.451        |
| Steiger         | TRUE          | FALSE        | TRUE          | TRUE          | TRUE          | TRUE          | TRUE            | TRUE            | TRUE         | TRUE         | TRUE         | TRUE         |
S4 Table: Inferring causal effects between second 6 risk factors and Stroke. In each cell we show the Bonferroni adjusted 1-0.05/96 ≈ 0.9995 confidence intervals (CIs) of $\theta$ for the MR methods, and CIs of $K$ for the CD methods; for Steiger’s method, we show the proportion of SNPs giving significant result. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.

| Method Type | Direction | Stroke to | Stroke to | Stroke to | Stroke to | Stroke to | Stroke to | Stroke to | Stroke to | Alcohol to Stroke | Alcohol to Stroke |
|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------------|------------------|
| MR-cML-DP-S | BW to Stroke | (1.477, 0.088) | FALSE | (1.058, 0.062) | FALSE | (1.057, 0.045) | TRUE | (1.045, 0.049) | TRUE | (1.031, 0.319) | FALSE | (1.044, 0.510) | FALSE |
| MR-cML-S   | BW to Stroke | (1.478, 0.086) | TRUE | (1.058, 0.062) | FALSE | (1.057, 0.045) | TRUE | (1.045, 0.049) | TRUE | (1.031, 0.319) | FALSE | (1.044, 0.510) | FALSE |
| CD-3ML-DP-S| Stroke to Stroke | (1.047, 0.062) | FALSE | (0.108, 0.074) | TRUE | (0.107, 0.212) | TRUE | (0.140, 0.240) | TRUE | (0.103, 0.195) | FALSE | (0.108, 0.474) | FALSE |
| CD-3ML-S   | Stroke to Stroke | (1.047, 0.062) | FALSE | (0.108, 0.074) | TRUE | (0.107, 0.212) | TRUE | (0.140, 0.240) | TRUE | (0.103, 0.195) | FALSE | (0.108, 0.474) | FALSE |
| CD-Ratio-S | Stroke to Stroke | (1.047, 0.062) | FALSE | (0.108, 0.074) | TRUE | (0.107, 0.212) | TRUE | (0.140, 0.240) | TRUE | (0.103, 0.195) | FALSE | (0.108, 0.474) | FALSE |
| CD-Egger-S | Stroke to Stroke | (1.047, 0.062) | FALSE | (0.108, 0.074) | TRUE | (0.107, 0.212) | TRUE | (0.140, 0.240) | TRUE | (0.103, 0.195) | FALSE | (0.108, 0.474) | FALSE |

**Legend:** TRUE/FALSE indicates whether the result is significant or not, and the cells giving significant results are marked in red.
S5 Table: Inferring causal effects between first 6 risk factors and T2D. In each cell we show the Bonferroni adjusted $1 - 0.05/96 \approx 0.9995$ confidence intervals (CIs) of $\theta$ for the MR methods, and CIs of $K$ for the CD methods; for Steiger’s method, we show the proportion of SNPs giving significant result. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.

| Method           | Direction | TG to T2D | T2D to TG | LDL to T2D | T2D to LDL | HDL to T2D | T2D to HDL | Height to T2D | T2D to Height | BMI to T2D | T2D to BMI | TF to T2D | T2D to TF | BF to T2D | T2D to BF |
|------------------|-----------|-----------|-----------|------------|------------|------------|------------|--------------|---------------|------------|------------|-----------|-----------|-----------|-----------|
| MR-cML-IP-S      | TRUE      | 0.164     | 0.217     | 0.203     | 0.024      | 0.035      | 0.021      | 0.154        | 0.057         | FALSE      | 0.409     | 0.042     | FALSE     | 0.289     | 0.089     |
| MR-cML-S         | TRUE      | 0.095     | 0.185     | 0.135     | 0.039      | 0.039      | 0.039      | 0.154        | 0.057         | FALSE      | 0.409     | 0.042     | FALSE     | 0.124     | 0.034     |
| CD-ML-TP-S       | TRUE      | 0.113     | 0.029     | 0.066     | 0.055      | 0.042      | 0.042      | 0.061        | 0.061         | FALSE      | 0.409     | 0.042     | FALSE     | 0.289     | 0.089     |
| MR-cML-TP        | TRUE      | 0.095     | 0.185     | 0.135     | 0.039      | 0.039      | 0.039      | 0.154        | 0.057         | FALSE      | 0.409     | 0.042     | FALSE     | 0.124     | 0.034     |
| CD-ML            | FALSE     | 0.181     | 0.181     | 0.181     | 0.181      | 0.181      | 0.181      | 0.181        | 0.181         | TRUE       | 3.329     | 0.032     | TRUE      | 0.289     | 0.089     |
| CD-ML-IP         | TRUE      | 0.111     | 0.029     | 0.066     | 0.055      | 0.042      | 0.042      | 0.061        | 0.061         | FALSE      | 0.409     | 0.042     | FALSE     | 0.124     | 0.034     |
| CD-ML-TP         | TRUE      | 0.095     | 0.185     | 0.135     | 0.039      | 0.039      | 0.039      | 0.154        | 0.057         | FALSE      | 0.409     | 0.042     | FALSE     | 0.124     | 0.034     |
S6 Table: Inferring causal effects between second 6 risk factors and T2D. In each cell we show the Bonferroni adjusted $1-0.05/96 \approx 0.9995$ confidence intervals (CIs) of $\theta$ for the MR methods, and CIs of $K$ for the CD methods; for Steiger’s method, we show the proportion of SNPs giving significant result. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.

| Method     | Direction | BW to T2D | BW to T2D | BW to DBP | DBP to T2D | DBP to DBP | DBP to SSB | T2D to T2D | T2D to T2D | T2D to SSB | SSB to T2D | T2D to FG | FG to T2D | T2D to Smoke | Smoke to T2D | Alcohol to T2D | Alcohol to T2D | T2D to Alcohol |
|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|----------|----------|----------------|---------------|----------------|----------------|---------------|
| MR-cML-DP-S|           | 0.226     | 0.034     | 0.004     | 0.412     | 0.039     | 0.185     | 0.215     | 1.344     | 0.177     | 0.064     | 0.029    | 0.614    | 0.177     | 0.064          | 0.051          | 0.177         | 0.064         | 0.177         |
| MR-cML-S  |           | 0.015     | 0.044     | 0.001     | 0.041     | 0.030     | 1.677     | 2.842     | 0.106     | 0.241     | 0.023     | 0.608    | 0.177    | 0.064    | 0.177          | 0.064          | 0.177         | 0.064         | 0.177         |
| CD-3ML-DP-S|           | 0.749     | 0.195     | 0.006     | 0.177     | 1.103     | 0.031     | 0.411     | 0.134     | 0.207     | 0.781     | 0.030    | 0.959    | 0.195    | 0.030          | 0.177          | 0.030         | 0.195         | 0.030         |
| CD-3ML-S  |           | 0.012     | 0.019     | 0.005     | 0.160     | 0.021     | 0.666     | 0.614     | 0.218     | 0.041     | 0.036    | 0.027   | 0.014    | 0.021   | 0.041          | 0.041          | 0.014         | 0.027         | 0.041         |
| CD-Ratio-S|           | 0.364     | 0.188     | 0.352     | 0.011     | 0.47     | 0.062     | 0.671     | 0.235     | 0.215     | 0.383    | 0.101   | 0.813    | 0.011   | 0.383          | 0.011         | 0.813         | 0.011         | 0.383         |
| CD-Egger-S|           | 0.201     | 0.123     | 0.413     | 0.057     | 0.549     | 0.012     | 1.055     | 0.276     | 0.018     | 0.015    | 0.016   | 1.056    | 0.014   | 0.016          | 0.015         | 1.060         | 0.006         | 0.015         |
| MR-cML-DP  |           | 0.245     | 0.054     | 0.056     | 0.012     | 0.039     | 1.585     | 3.221     | 0.11      | 0.959     | 0.029    | 0.614   | 0.177    | 0.029   | 0.614          | 0.177         | 0.029         | 0.614         | 0.177         |
| MR-cML-S  |           | 0.015     | 0.044     | 0.041     | 0.034     | 0.069     | 0.033     | 0.016     | 0.039     | 0.016     | 0.021    | 0.218   | 0.006   | 0.021   | 0.218          | 0.006         | 0.021         | 0.218         | 0.006         |
| CD-3ML-DP-S|           | 0.399     | 0.223     | 0.452     | 0.054     | 0.549     | 0.013     | 0.411     | 0.043     | 0.847     | 0.038   | 0.959   | 0.151   | 0.038   | 0.959          | 0.151         | 0.038         | 0.959         | 0.151         |
| CD-3ML-S  |           | 0.012     | 0.019     | 0.356     | 0.005     | 0.466     | 0.020     | 0.019     | 0.167     | 0.020     | 0.019    | 0.019   | 0.019   | 0.019   | 0.019          | 0.019         | 0.019         | 0.019         | 0.019         |
| CD-Ratio-S|           | 0.007     | 0.028     | 0.322     | 0.001     | 0.185     | 0.024     | 0.388     | 0.017     | 0.017     | 0.038    | 0.001   | 0.038   | 0.017   | 0.038          | 0.017         | 0.038         | 0.017         | 0.038         |
| CD-Egger-S|           | 0.172     | 0.123     | 0.569     | 0.087     | 0.106     | 0.017     | 0.017     | 0.421     | 0.017     | 0.421    | 0.017   | 0.017   | 0.421   | 0.017          | 0.017         | 0.017         | 0.421         | 0.017         |
| LHC-MR     |           | 0.118     | 0.068     | 0.789     | 0.248     | 0.001     | 1.007     | 0.216     | 0.5       | 0.216    | 0.5     | 0.216   | 0.5     | 0.216   | 0.5          | 0.216         | 0.5           | 0.216         | 0.5           |
| Steiger    |           | 0.018     | 0.045     | 0.097     | 0.042     | 0.048     | 0.091     | 0.032     | 0.055     | 0.055    | 0.055   | 0.055   | 0.055   | 0.055   | 0.055          | 0.055         | 0.055         | 0.055         | 0.055         |

47
S7 Table: Inferring causal effects between first 6 risk factors and Asthma. In each cell we show the Bonferroni adjusted $1-0.05/96 \approx 0.9995$ confidence intervals (CIs) of $\theta$ for the MR methods, and CIs of $K$ for the CD methods; for Steiger’s method we show the proportion of SNPs giving significant result. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.

| Method     | TG to Asthma | LDL to Asthma | Asthma to LDE | HDL to Asthma | Asthma to BMI | BMI to Asthma | Asthma to BF | BF to Asthma |
|------------|--------------|---------------|---------------|--------------|--------------|--------------|--------------|--------------|
| CD-cML-DP-S | 0.059, 0.043 | 0.013 | 0.033 | 0.032 | 0.026 | 0.041 | 0.023 | 0.026 |
| CD-cML-S   | 0.042, 0.031 | 0.010 | 0.031 | 0.036 | 0.029 | 0.042 | 0.029 | 0.029 |
| MR-cML-DP-S | 0.064, 0.018 | 0.010 | 0.031 | 0.036 | 0.029 | 0.042 | 0.029 | 0.029 |
| MR-cML-S   | 0.056, 0.017 | 0.010 | 0.031 | 0.036 | 0.029 | 0.042 | 0.029 | 0.029 |
| CD-Ratio-S | 0.022, 0.012 | 0.003 | 0.016 | 0.030 | 0.029 | 0.032 | 0.029 | 0.029 |
| CD-Egger-S | 0.027, 0.017 | 0.003 | 0.016 | 0.030 | 0.029 | 0.032 | 0.029 | 0.029 |
| MR-cML-DP  | 0.060, 0.018 | 0.010 | 0.031 | 0.036 | 0.029 | 0.042 | 0.029 | 0.029 |
| MR-cML-S   | 0.056, 0.017 | 0.010 | 0.031 | 0.036 | 0.029 | 0.042 | 0.029 | 0.029 |
| CD-MU-DP   | 0.063, 0.017 | 0.010 | 0.031 | 0.036 | 0.029 | 0.042 | 0.029 | 0.029 |
| CD-MU-S    | 0.056, 0.017 | 0.010 | 0.031 | 0.036 | 0.029 | 0.042 | 0.029 | 0.029 |
| CD-Ratio   | 0.013, 0.010 | 0.003 | 0.016 | 0.030 | 0.029 | 0.032 | 0.029 | 0.029 |
| CD-Egger   | 0.021, 0.010 | 0.003 | 0.016 | 0.030 | 0.029 | 0.032 | 0.029 | 0.029 |
| LH-MR      | 0.065, 0.059 | 0.013 | 0.031 | 0.036 | 0.029 | 0.042 | 0.029 | 0.029 |
| Steiger    | TRUE, FALSE  | TRUE, FALSE  | TRUE, FALSE  | TRUE, FALSE | TRUE, FALSE | TRUE, FALSE | TRUE, FALSE | TRUE, FALSE |
S8 Table: Inferring causal effects between second 6 risk factors and Asthma. In each cell we show the Bonferroni adjusted $1-0.05/96 \approx 0.9995$ confidence intervals (CIs) of $\theta$ for the MR methods, and CIs of $K$ for the CD methods; for Steiger’s method, we show the proportion of SNPs giving significant result. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.

| Method       | Direction | BW to Asthma | Asthma to BW | DBP to Asthma | Asthma to DBP | SBP to Asthma | Asthma to SBP | FG to Asthma | Asthma to FG | Smoke to Asthma | Asthma to Smoke | Alcohol to Asthma | Asthma to Alcohol |
|--------------|-----------|--------------|--------------|---------------|--------------|--------------|--------------|-------------|--------------|----------------|----------------|-----------------|------------------|
| MR-cML-DP-S  |           | -0.123       | 0.043        | 0.012         | 0.443        | 0.001        | 0.686        | 0.022        | 0.711        | 0.018          | 0.052          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | FALSE        | TRUE        | FALSE        | TRUE           | TRUE           | FALSE           | FALSE           |
| MR-cML-S     |           | -0.394       | 0.021        | 0.012         | 0.089        | 0.009        | 0.159        | 0.148        | 0.045        | 0.012          | 0.041          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
| CD-cML-DP-S  |           | -0.018       | 0.076        | 0.057         | 0.069        | 0.066        | 0.059        | 0.065        | 0.289        | 0.123          | 0.073          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | FALSE        | TRUE         | FALSE           | FALSE           | FALSE           | FALSE           |
| CD-cML-S     |           | -0.004       | 0.066        | 0.022         | 0.052        | 0.020        | 0.032        | 0.261        | 0.136        | 0.096          | 0.168          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
| CD-Ratio-S   |           | -0.048       | 0.011        | 0.008         | 0.023        | 0.016        | 0.024        | 0.041        | 0.223        | 0.041          | 0.223          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
| CD-Ratio     |           | -0.007       | 0.060        | 0.021         | 0.065        | 0.006        | 0.098        | 0.261        | 0.136        | 0.096          | 0.168          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
| CD-cML-DP    |           | -0.038       | 0.023        | 0.016         | 0.023        | 0.011        | 0.024        | 0.041        | 0.223        | 0.041          | 0.223          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
| CD-cML       |           | -0.044       | 0.011        | 0.008         | 0.023        | 0.016        | 0.024        | 0.041        | 0.223        | 0.041          | 0.223          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
| CD-Ratio     |           | -0.009       | 0.066        | 0.022         | 0.052        | 0.020        | 0.032        | 0.261        | 0.136        | 0.096          | 0.168          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
| CD-Ratio     |           | -0.007       | 0.060        | 0.021         | 0.065        | 0.006        | 0.098        | 0.261        | 0.136        | 0.096          | 0.168          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
| LHC-MR       |           | -0.394       | 0.076        | 0.057         | 0.069        | 0.066        | 0.059        | 0.065        | 0.289        | 0.123          | 0.073          | FALSE           | FALSE           |
|              |           | TRUE         | FALSE         | TRUE          | FALSE        | TRUE         | TRUE         | TRUE        | TRUE         | TRUE           | TRUE           | FALSE           | FALSE           |
S2.2  Pairs of 4 Diseases

S41 Fig: Causal relationship between pairs of 4 diseases.

| Asthma | CAD | Stroke | T2D |
|--------|-----|--------|-----|
| Asthma |     |        |     |
| CAD    |     |        |     |
| Stroke |     |        |     |
| T2D    |     |        |     |

- ▲: $p > 0.05$
- ▲: $0.004 < p < 0.05$
- ▲: $p < 0.004$
S9 Table: Inferring causal effects between pairs of 4 diseases. In each cell we show the Bonferroni adjusted 1-0.05/12 ≈ 0.996 confidence intervals (CIs) of θ for the MR methods, and CIs of K for the CD methods; for Steiger’s method, we show the proportion of SNPs giving significant result. TRUE/FALSE in each cell indicates whether the result is significant or not, and the cells giving significant results are marked in red.

| Method | CAD to Stroke | Stroke to CAD | CAD to T2D | T2D to CAD | Stroke to Asthma | Asthma to Stroke | Stroke to T2D | T2D to Stroke | Stroke to Asthma | Asthma to Stroke |
|--------|---------------|---------------|------------|------------|----------------|----------------|--------------|--------------|----------------|----------------|
| CD-cML-DP-S | 0.270, 0.386 | TRUE | FALSE | FALSE | TRUE | FALSE | FALSE | FALSE | FALSE | FALSE |
| CD-cML-S | 0.238, 0.396 | TRUE | FALSE | FALSE | TRUE | FALSE | FALSE | FALSE | FALSE | FALSE |
| CD-ML | 0.005, 0.149 | FALSE | 0.046, 0.173 | FALSE | 0.084, 0.187 | TRUE | 0.174, 0.246 | FALSE | 0.04, 0.132 | FALSE |
| CD-cML | 0.136, 0.225 | TRUE | 0.041, 0.165 | FALSE | 0.059, 0.163 | TRUE | 0.157, 0.239 | FALSE | 0.058, 0.155 | FALSE |
| S2.3 Links to GWAS Summary Datasets

We downloaded the GWAS summary datasets from the IEU GWAS database [1], which are the same as the data included in R package TwoSampleMR. The links are shown in S10 Table.
S3  Theoretical Results

S3.1 Proof of Theorem 1

**Theorem 1.** Under Assumptions 1 and 2, if \( m_{YX}^0 \in \mathcal{M} \), we have \( P(\hat{m}_I = m_{YX}^0) \to 1 \) and \( P(\hat{B}_{XY}(\hat{m}_I) = B_{XY}^0) \to 1 \) as \( N_1, N_2 \to \infty \). Furthermore, the cMLE \( \hat{K}_{XY} := \hat{K}_{XY}(\hat{m}_I) \) is consistent and asymptotically normal:

\[
\sqrt{V}(\hat{K}_{XY} - K_{XY}) \xrightarrow{d} N(0,1), \text{ as } N_1, N_2 \to \infty,
\]

where

\[
V = \sum_{g \in (B_{XY}^c)^c} \frac{\rho_{Xg}^2}{\sigma_{Xg}^2} - \frac{K_{XY}^2}{\sigma_{Yg}^2}.
\]

**Proof.** First, we show \( P(\hat{B}_{XY}(m_{YX}^0) = B_{XY}^0) \to 1 \), which is equivalent to show for any \( B_1 \subseteq \{1, \cdots, m\} \) such that \(|B_1| = m_{YX}^0 \) and \( B_1 \neq B_{XY}^0 \), \( P(\hat{B}_{XY}(m_{YX}^0) = B_1) \to 0 \) as \( N_1, N_2 \to \infty \). We have

\[
P(\hat{B}_{XY}(m_{YX}^0) = B_1)
\leq P(\min_{K, \rho_{Xg} \in B_1} \sum_{g \in (B_{XY}^c)^c} \left( \frac{(r_{Xg} - \hat{\rho}_{Xg})^2}{SE(r_{Xg})^2} + \frac{(r_{Yg} - \hat{\rho}_{Yg})^2}{SE(r_{Yg})^2} \right) \leq \min_{K, \rho_{Xg} \in (B_{XY}^c)^c} \sum_{g \in (B_{XY}^c)^c} \left( \frac{(r_{Xg} - \hat{\rho}_{Xg})^2}{SE(r_{Xg})^2} + \frac{(r_{Yg} - \hat{\rho}_{Yg})^2}{SE(r_{Yg})^2} \right)
\]

\[
\leq P(\min_{K, \rho_{Xg} \in B_1} \sum_{g \in (B_{XY}^c)^c} \left( \frac{(r_{Xg} - \hat{\rho}_{Xg})^2}{SE(r_{Xg})^2} + \frac{(r_{Yg} - \hat{\rho}_{Yg})^2}{SE(r_{Yg})^2} \right) \leq \sum_{g \in (B_{XY}^c)^c} \left( \frac{(r_{Xg} - \hat{\rho}_{Xg})^2}{SE(r_{Xg})^2} + \frac{(r_{Yg} - \hat{\rho}_{Yg})^2}{SE(r_{Yg})^2} \right)).
\]
Note that, for \( g \in (B_Y^n)^c \), \( r_X^g - p_X^g \) and \( \frac{r_X^g - K_{XY} \rho_X^g}{SE(\rho_g)} \) are independent, so for any \( \varepsilon > 0 \), there exists \( C > 0 \) such that
\[
P(\sum_{g \in (B_Y^n)^c} \left( \frac{(r_X^g - \rho_X^g)^2}{SE(r_X^g)^2} + \frac{(r_Y^g - K_{XY} \rho_X^g)^2}{SE(\rho_g)^2} \right) > C) < \frac{\varepsilon}{2}.
\]

And we have
\[
P(\min_{\hat{\rho}_X^g \in B_1} \sum_{g \in (B_Y^n)^c} \left( \frac{(r_X^g - \hat{\rho}_X^g)^2}{SE(r_X^g)^2} + \frac{(r_Y^g - \hat{\rho}_X^g)^2}{SE(\rho_g)^2} \right) \leq C) + P(\sum_{g \in (B_Y^n)^c} \left( \frac{(r_X^g - \hat{\rho}_X^g)^2}{SE(r_X^g)^2} + \frac{(r_Y^g - K_{XY} \rho_X^g)^2}{SE(\rho_g)^2} \right) > C).
\]

After profiling out \( \hat{\rho}_X^g \)'s, we get
\[
\min_{\hat{\rho}_X^g \in B_1} \sum_{g \in (B_Y^n)^c} \left( \frac{(r_X^g - \hat{\rho}_X^g)^2}{SE(r_X^g)^2} + \frac{(r_Y^g - \hat{\rho}_X^g)^2}{SE(\rho_g)^2} \right) = \min_{\hat{\rho}_X^g \in B_1} \frac{(r_Y^g - \bar{\hat{K}} \cdot r_X^g)^2}{SE(r_g)^2 + \bar{\hat{K}}^2SE(r_X^g)^2},
\]
so
\[
P(\min_{\hat{\rho}_X^g \in B_1} \sum_{g \in (B_Y^n)^c} \left( \frac{(r_X^g - \hat{\rho}_X^g)^2}{SE(r_X^g)^2} + \frac{(r_Y^g - \hat{\rho}_X^g)^2}{SE(\rho_g)^2} \right) \leq C)
= P(\min_{\hat{\rho}_X^g \in B_1} \frac{(r_Y^g - \bar{\hat{K}} \cdot r_X^g)^2}{SE(r_g)^2 + \bar{\hat{K}}^2SE(r_X^g)^2} \leq C).
\]

We have \( \frac{r_X^g - K_{XY} \rho_X^g}{\sqrt{SE(r_X^g)^2 + K^2SE(r_X^g)^2}} \sim N(K_{XY} \cdot \rho_X^g + b_{XY} \cdot r_X^g, 1) \), so \( \sum_{g \in B_1^n} \frac{(r_X^g - K_{XY} \rho_X^g)^2}{SE(r_X^g)^2 + K^2SE(r_X^g)^2} \) follows non-central \( \chi^2 \) distribution with degrees of freedom \( m - m_Y^n \) and non-centrality parameter \( \lambda_{\bar{\hat{K}}} \) depending on \( \bar{\hat{K}} \)
\[
\lambda_{\bar{\hat{K}}} = \sum_{g \in B_1^n} \frac{(K_{XY} \cdot \rho_X^g + b_{XY} \cdot r_X^g - \bar{\hat{K}} \cdot \rho_X^g)^2}{SE(\rho_g)^2 + \bar{\hat{K}}^2SE(r_X^g)^2}.
\]

With Assumption 2, we get
\[
\lambda_{\bar{\hat{K}}} \geq \sum_{g \in B_1^n} \frac{(K_{XY} \cdot \rho_X^g + b_{XY} \cdot r_X^g - \bar{\hat{K}} \cdot \rho_X^g)^2}{u_Y + \bar{\hat{K}}^2 \cdot \frac{\rho_X}{\sqrt{\nu}}}
= N_2 \cdot \sum_{g \in B_1^n} \frac{(K_{XY} \cdot \rho_X^g + b_{XY} \cdot r_X^g - \bar{\hat{K}} \cdot \rho_X^g)^2}{u_Y + \bar{\hat{K}}^2 \cdot \frac{\rho_X}{\sqrt{\nu}}}.
\]

With Assumption 1, we know
\[
\min_{\hat{\rho}_X^g \in B_1} \sum_{g \in B_1^n} \frac{(K_{XY} \cdot \rho_X^g + b_{XY} \cdot r_X^g - \bar{\hat{K}} \cdot \rho_X^g)^2}{u_Y + \bar{\hat{K}}^2 \cdot \frac{\rho_X}{\sqrt{\nu}}} = v > 0,
\]
here \( v \) is a constant. This is because, with Assumption 1, there is no \( \bar{\hat{K}} \) making \( K_{XY} \cdot \rho_X^g + b_{XY} \cdot r_X^g = \bar{\hat{K}} \cdot \rho_X^g \) for all \( g \in B_1^n \) simultaneously. So we have \( \min_{\bar{\hat{K}}} \lambda_{\bar{\hat{K}}} \geq N_2 \cdot v \). Then as \( N_2 \) large enough, we have
\[
P(\min_{\hat{\rho}_X^g \in B_1} \sum_{g \in (B_Y^n)^c} \frac{(r_Y^g - \bar{\hat{K}} \cdot r_X^g)^2}{SE(r_g)^2 + \bar{\hat{K}}SE(r_X^g)^2} \leq C) \leq \frac{\varepsilon}{2}.
\]
Combining (1) and (2), we get \( P(\hat{B}_{XY}(m_{XY}^0) = B_{XY}^0) \to 1 \) as \( N_1, N_2 \to \infty \).

Next, we show \( P(\hat{m}_1 = m_{XY}^0) \to 1 \). For any \( m_1 < m_{XY}^0 \), we have

\[
P(\hat{m}_1 = m_1) \leq P(BIC(m_1) \leq BIC(m_{XY}^0))
\]

\[
= P\left(-2 \cdot L(\hat{K}_{XY}(m_1), \hat{\rho}_{XY}(m_1), \hat{b}_{XY,g}(m_1)) + \log(n) \cdot m_1 \leq -2 \cdot L(\hat{K}_{XY}(m_{XY}^0), \hat{\rho}_{XY}(m_{XY}^0), \hat{b}_{XY,g}(m_{XY}^0)) + \log(n) \cdot m_{XY}^0\right)
\]

\[
= P\left(2 \cdot L(\hat{K}_{XY}(m_{XY}^0), \hat{\rho}_{XY}(m_{XY}^0), \hat{b}_{XY,g}(m_{XY}^0)) - 2 \cdot L(\hat{K}_{XY}(m_1), \hat{\rho}_{XY}(m_1), \hat{b}_{XY,g}(m_1)) \leq \log(n)(m_{XY}^0 - m_1)\right).
\]

As we have shown \( P(\hat{B}_{XY}(m_{XY}^0) = B_{XY}^0) \to 1 \), with probability goes to 1 we have

\[
2 \cdot L(\hat{K}_{XY}(m_{XY}^0), \hat{\rho}_{XY}(m_{XY}^0), \hat{b}_{XY,g}(m_{XY}^0)) - 2 \cdot L(\hat{K}_{XY}(m_1), \hat{\rho}_{XY}(m_1), \hat{b}_{XY,g}(m_1))
\]

\[
= \min_{\hat{K}, \hat{\rho}_{XY} \in B_{XY}} \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) - \min_{\hat{K}, \hat{\rho}_{XY} \in B_{XY}} \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right).
\]

Then we get

\[
P(\hat{m}_1 = m_1)
\]

\[
\leq P\left(\min_{\hat{K}, \hat{\rho}_{XY} \in B_{XY}} \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) \leq \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - K \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) + \log(n)(m_{XY}^0 - m_1)\right)
\]

\[
\leq \sum_{|B| = m_1} P\left(\min_{\hat{K}, \hat{\rho}_{XY} \in B_{XY}} \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) \leq \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - K \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) + \log(n)(m_{XY}^0 - m_1)\right)
\]

Similar as above, we get

\[
\min_{\hat{K}, \hat{\rho}_{XY} \in B_{XY}} \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) = \min_{\hat{K}} \sum_{g \in B_{XY}} \frac{(r_g - \hat{K} \cdot \hat{r}_g)^2}{SE(r_g)^2 + K^2 \cdot SE(r_g)^2},
\]

and \( \sum_{g \in B_{XY}} \frac{(r_g - \hat{K} \cdot \hat{r}_g)^2}{SE(r_g)^2 + K^2 \cdot SE(r_g)^2} \) follows non-central \( \chi^2 \) distribution with degrees of freedom \( m - m_1 \) and non-centrality parameter \( \lambda_K \) depending on \( \hat{K} \)

\[
\lambda_K = \sum_{g \in B_{XY}} \frac{(K_{XY} \cdot \hat{\rho}_{XY} + b_{XY,g} - \hat{K} \cdot \hat{\rho}_{XY})^2}{SE(r_g)^2 + K^2 \cdot SE(r_g)^2}.
\]

Similarly, since \( m_1 < m_{XY}^0 \), with Assumption 2 we have \( \lambda_K \geq N_2 \cdot \nu \) for some constant \( \nu \), so for any \( |B| = m_1 \), we get

\[
P\left(\min_{\hat{K}, \hat{\rho}_{XY} \in B_{XY}} \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) \leq \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - K \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) + \log(n)(m_{XY}^0 - m_1)\right) \to 0.
\]

This gives us \( P(\hat{m}_1 = m_1) \to 0 \) for any \( m_1 < m_{XY}^0 \). For any \( m_1 > m_{XY}^0 \), we have

\[
P(\hat{m}_1 = m_1)
\]

\[
\leq P\left(\log(n)(m_1 - m_{XY}^0) \leq \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right) - \min_{\hat{K}, \hat{\rho}_{XY} \in B_{XY}} \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right)\right)
\]

\[
\leq P\left(\log(n)(m_1 - m_{XY}^0) \leq \sum_{g \in B_{XY}} \left( \frac{(r_{XY} - \hat{\rho}_{XY})^2}{SE(r_{XY})^2} + \frac{(r_g - \hat{K} \hat{\rho}_{XY})^2}{SE(r_g)^2} \right)\right).
\]
Since $\sum_{g \in \{B^{0}_{XY}\}^c} \left(\frac{(r_{Xg} - \rho_{Xg})^2}{SE(r_{Xg})^2} + \frac{(r_{Yg} - \rho_{Yg})^2}{SE(r_{Yg})^2}\right)$ is a central $\chi^2$ distribution with degrees of freedom $2(m - m_{0_{XY}})$, we get $P(\hat{m}_{I} = m_{1}) \to 0$ for any $m_{1} > m_{0_{XY}}$. So we have $P(\hat{m}_{I} = m_{0_{XY}}) \to 1$ as $N_{1}, N_{2} \to \infty$.

As $P(\hat{B}_{XY}(\hat{m}_{I}) = B^{0}_{XY}) \to 1$, we could consistently select all invalid IVs. Following Theorem 3.2 in [4], we have

$$\frac{V}{\sqrt{V_{1}}} (\hat{K}_{XY} - K_{XY}) \xrightarrow{d} N(0, 1), \text{ as } N_{1}, N_{2} \to \infty,$$

where

$$V = \sum_{g \in \{B^{0}_{XY}\}^c} \frac{\rho_{Xg}^{2} \sigma_{Xg}^{2} + \rho_{Yg}^{2} \sigma_{Yg}^{2}}{(\sigma_{Xg}^{2} \cdot K_{XY}^{2} + \sigma_{Yg}^{2})^{2}} = \sum_{g \in \{B^{0}_{XY}\}^c} \frac{\rho_{Xg}^{2}}{\sigma_{Xg}^{2} \cdot K_{XY}^{2} + \sigma_{Yg}^{2}},$$

and

$$V_{1} = \sum_{g \in \{B^{0}_{XY}\}^c} \frac{\rho_{Xg}^{2} \sigma_{Xg}^{2} + \rho_{Yg}^{2} \sigma_{Yg}^{2} + \sigma_{Xg}^{2} \sigma_{Yg}^{2}}{(\sigma_{Xg}^{2} \cdot K_{XY}^{2} + \sigma_{Yg}^{2})^{2}}.$$

In our model $\rho_{Xg}$’s and $\rho_{Yg}$’s are fixed constants, $\sigma_{Xg}^{2}$’s and $\sigma_{Yg}^{2}$’s are $O(1/n)$, so we have $V/V_{1} \to 1$, and

$$\sqrt{V}(\hat{K}_{XY} - K_{XY}) \xrightarrow{d} N(0, 1), \text{ as } N_{1}, N_{2} \to \infty.$$

---

**S3.2 Proof of Theorem 2**

First we introduce the definition of “converge weakly”, as Definition 2.2 in [2].

**Definition 2.2 by Xiong et al. [2].** $F(\cdot)$ is a distribution function, $F_{n}(\cdot)$ is random distribution function that depends on some random variable. We say $F_{n}(\cdot)$ converges weakly to $F(\cdot)$ in probability if for each continuous point $x$ of $F(\cdot)$, $F_{n}(x) \xrightarrow{p} F(x)$ as $n \to \infty$. This is denoted by $F_{n}(\cdot) \xrightarrow{w.p.} F(\cdot)$.

Now we show the proof of Theorem 2.

**Theorem 2.** Under Assumptions 1 and 2, conditional on the original GWAS summary data, $\sqrt{V}(\hat{K}^{(i)}_{XY} - K_{XY}) \xrightarrow{w.p.} N(0, 1)$ as $N_{1}, N_{2} \to \infty$.

**Proof.** Denote $\bar{B} = \{i : \hat{b}_{XY} \neq 0\}$ as the set of estimated invalid IVs with non-zero direct effects based on perturbed data. First we show that $P(\bar{B} = B_{XY}^{0} | \mathcal{D}) \xrightarrow{P} 1$, which is equivalent to for any $\varepsilon > 0, \delta > 0$, there exists $n$ such that when $n_{1} > n, n_{2} > n$ we have $P(\bar{P}(\bar{B} = B_{XY}^{0} | \mathcal{D}) < 1 - \varepsilon) < \delta$. Following similar argument in Theorem 1, we could get the unconditional probability $P(\bar{B} = B_{XY}^{0}) \to 1$. Suppose we could find a pair of $\varepsilon_{0} > 0, \delta_{0} > 0$ such that $P(\bar{P}(\bar{B} = B_{XY}^{0} | \mathcal{D}) < 1 - \varepsilon_{0}) > \delta_{0}$ for arbitrarily large $n_{1}, n_{2}$, then we can get

$$P(\bar{B} = B_{XY}^{0}) = \int_{\mathcal{D}} \bar{P}(\bar{B} = B_{XY}^{0}) dF(\mathcal{D}) < 1 - \varepsilon_{0} \delta_{0},$$

55
contradicts that \( P(\hat{B} = B_{XY}^{0}) \rightarrow 1 \), thus we have shown that \( P(\hat{B} = B_{XY}^{0} | \emptyset) \overset{w}{\rightarrow} 1 \). Now we could focus on the case that \( \hat{B} = \hat{B} = B_{XY}^{0} \), for simplicity we use \( \hat{K}, \hat{\hat{K}} \) to represent \( \hat{K}_{XY}^{(l)} \), \( \hat{K}_{XY} \). Similar to [4], after profiling out \( \rho_{XY} \)'s in the original log-likelihood function, we have

\[
\hat{K} = \arg \min_{K} \sum_{g \in (B_{XY}^{0})^{c}} \frac{(\hat{r}_{g} - K \cdot \hat{r}_{g})^{2}}{\sigma_{Xg}^{2} \cdot K^{2} + \sigma_{Yg}^{2}}, \quad \hat{\hat{K}} = \arg \min_{K} \sum_{g \in (B_{XY}^{0})^{c}} \frac{(r_{g} - K \cdot r_{g})^{2}}{\sigma_{Xg}^{2} \cdot K^{2} + \sigma_{Yg}^{2}}.
\]

(3)

Denote

\[
f(K) = \sum_{g \in (B_{XY}^{0})^{c}} \frac{(\hat{r}_{g} - K \cdot \hat{r}_{g})^{2}}{\sigma_{Xg}^{2} \cdot K^{2} + \sigma_{Yg}^{2}},
\]

and

\[
\phi(K) = \frac{\partial f(K)}{\partial K} = \sum_{g \in (B_{XY}^{0})^{c}} \frac{(r_{g} - K r_{g})(K \sigma_{Xg}^{2} + r_{g} \sigma_{Yg}^{2})}{(\sigma_{Xg}^{2} K^{2} + \sigma_{Yg}^{2})^{2}},
\]

\[
= \sum_{g \in (B_{XY}^{0})^{c}} \frac{(r_{g} - K r_{g})(K \sigma_{Xg}^{2} + r_{g} \sigma_{Yg}^{2})}{(\sigma_{Xg}^{2} K^{2} + \sigma_{Yg}^{2})^{2}},
\]

\[
= \sum_{g \in (B_{XY}^{0})^{c}} \frac{(r_{g} - K r_{g})(K \sigma_{Xg}^{2} + r_{g} \sigma_{Yg}^{2}) + (\xi_{g} - K \xi_{g})(K r_{g} \sigma_{Xg}^{2} + r_{g} \sigma_{Yg}^{2} + K \sigma_{Xg}^{2} + r_{g} \sigma_{Yg}^{2})}{(\sigma_{Xg}^{2} K^{2} + \sigma_{Yg}^{2})^{2}},
\]

here \( \xi_{g} = \hat{r}_{g} - r_{g} = \xi_{g}, \sigma_{g} = \hat{r}_{g} - r_{g} \sim N(0, \sigma_{g}^{2}) \). We have

\[
0 = \phi(\hat{K}) = \phi(\hat{K}) + \phi'(\hat{K})(\hat{K} - \hat{\hat{K}}) + \frac{1}{2} \phi''(K^{*})(\hat{K} - \hat{\hat{K}})^{2},
\]

with \( K^{*} \) is between \( \hat{K} \) and \( \hat{\hat{K}} \), thus

\[
\sqrt{V}(\hat{K} - \hat{\hat{K}}) = -\frac{\phi(\hat{K})/\sqrt{V}}{\phi'(\hat{K})/V + (1/2)(\hat{K} - \hat{\hat{K}})\phi''(K^{*})/V}.
\]

Next we show \( \phi(\hat{K})/\sqrt{V} \overset{w}{\rightarrow} N(0, 1) \). From equation (6), we can get

\[
\phi(\hat{K}) = \sum_{g \in (B_{XY}^{0})^{c}} \frac{(r_{g} - K r_{g})(\hat{K} \xi_{g} \sigma_{Xg}^{2} + \epsilon_{g} \sigma_{Yg}^{2}) + (\xi_{g} - K \xi_{g})(K r_{g} \sigma_{Xg}^{2} + r_{g} \sigma_{Yg}^{2} + \hat{K} \sigma_{Xg}^{2} + \epsilon_{g} \sigma_{Yg}^{2})}{(\sigma_{Xg}^{2} K^{2} + \sigma_{Yg}^{2})^{2}}.
\]

Note that \( \xi_{g} \)'s and \( \epsilon_{g} \)'s are \( O_{p}(1/\sqrt{n}) \), \( n = \min(N_{1}, N_{2}) \), thus

\[
\phi(\hat{K}) = \sum_{g \in (B_{XY}^{0})^{c}} \frac{(r_{g} - K r_{g})(\hat{K} \xi_{g} \sigma_{Xg}^{2} + \epsilon_{g} \sigma_{Yg}^{2}) + (\xi_{g} - K \xi_{g})(K r_{g} \sigma_{Xg}^{2} + r_{g} \sigma_{Yg}^{2} + \hat{K} \sigma_{Xg}^{2} + \epsilon_{g} \sigma_{Yg}^{2})}{(\sigma_{Xg}^{2} K^{2} + \sigma_{Yg}^{2})^{2}} + O_{p}(1),
\]

thus \( \phi(\hat{K})/\sqrt{V} \overset{w}{\rightarrow} N(0, V^{*}/V) \overset{w}{\rightarrow} N(0, 1) \), with

\[
V^{*} = \sum_{g \in (B_{XY}^{0})^{c}} \frac{\sigma_{Xg}^{2}(r_{g} - K r_{g})^{2} + \frac{1}{2} \sigma_{Xg}^{2}(r_{g} - K r_{g})^{2}}{(\sigma_{Xg}^{2} K^{2} + \sigma_{Yg}^{2})^{4}}.
\]

as \( r_{g} \overset{p}{\rightarrow} \rho_{X}, r_{g} \overset{p}{\rightarrow} \rho_{Y}, \hat{K} \overset{p}{\rightarrow} K_{0} \), we can get \( V^{*}/V \overset{p}{\rightarrow} 1 \), thus we get \( \phi(\hat{K})/\sqrt{V} \overset{w}{\rightarrow} N(0, 1) \).

Next we show \( -\phi'(\hat{K})/V \overset{w}{\rightarrow} 1 \). After some calculation we get

\[
\phi'(K) = \sum_{g \in (B_{XY}^{0})^{c}} \frac{2 \sigma_{Xg}^{2} r_{g} \hat{r}_{g} - 6 \sigma_{Xg}^{2} r_{g} \hat{r}_{g} (r_{g} - K r_{g})^{2} + 3 \sigma_{Xg}^{2} r_{g} \hat{r}_{g} (r_{g} - K r_{g})^{2}}{(\sigma_{Xg}^{2} K^{2} + \sigma_{Yg}^{2})^{3}},
\]

(5)
as \( r_{Xg} \xrightarrow{p} p_{Xg}, \hat{r}_{Yg} \xrightarrow{p} p_{Yg}, \hat{K} \xrightarrow{p} K_0 \), we get \(-\varphi'(\hat{K})/V \xrightarrow{w} 1\), with Theorem 3.3 in [2], \(-\varphi'(\hat{K})/V \xrightarrow{D} w \rightarrow 1\).

Based on equation (8), we can see \( \delta''(K) \) has its numerator of order \( n^5 \) and its denominator of order \( n^6 \), thus \( \delta''(K^*)/V = O_p(1) \). As \( \hat{K} \xrightarrow{D} K_0, \hat{K} \xrightarrow{p} K_0 \), we have \( \hat{K} - \hat{K} \xrightarrow{D} \rightarrow 0 \), again with Theorem 3.3 in [2] we get \( \hat{K} - \hat{K} \xrightarrow{D} w \rightarrow 0 \). Thus we can get \( \frac{1}{2} \delta''(K^*)(\hat{K} - \hat{K}) \xrightarrow{D} w \rightarrow 0 \). Now with Theorem 3.2 in [2], we can get \( \sqrt{V}(\hat{K} - \hat{K}) \xrightarrow{D} w \rightarrow N(0,1) \), completing the proof. \( \square \)

S3.3 MR-cML with Data Perturbation

Now we show that the data perturbation scheme is also consistent for MR-cML in [3]. We use the following notations: the true effects on \( X \) are \( \beta_{Xi} \)'s, and those on \( Y \) are \( \beta_{Yi} \)'s; the estimated/observed effects on \( X \) are \( \hat{\beta}_{Xi} \sim N(\beta_{Xi}, \sigma_{Xi}^2) \), and those on \( Y \) are \( \hat{\beta}_{Yi} \sim N(\beta_{Yi}, \sigma_{Yi}^2) \). Here \( \sigma_{Xi} \)'s and \( \sigma_{Yi} \)'s are the true standard deviations; in practice we have the standard errors \( \hat{\sigma}_{Xi} \)'s and \( \hat{\sigma}_{Yi} \)'s as their estimates from GWAS datasets, thus approximately we have \( \hat{\beta}_{Xi} \sim N(\beta_{Xi}, \hat{\sigma}_{Xi}^2) \) and \( \hat{\beta}_{Yi} \sim N(\beta_{Yi}, \hat{\sigma}_{Yi}^2) \). For simplicity and without ambiguity, we treat the standard errors \( \hat{\sigma}_{Xi} \)'s and \( \hat{\sigma}_{Yi} \)'s as the true values of \( \sigma_{Xi} \)'s and \( \sigma_{Yi} \)'s in the following. The perturbed effects on \( X \) are \( \hat{\beta}_{Xi} \sim N(\beta_{Xi}, \sigma_{Xi}^2) \), and the perturbed effects on \( Y \) are \( \hat{\beta}_{Yi} \sim N(\beta_{Yi}, \sigma_{Yi}^2) \). The true causal effect is \( \theta \), the estimated causal effect based on the observed data with cML-BIC is \( \hat{\theta} \), and the estimated causal effect based on a perturbed dataset with cML-BIC is \( \hat{\theta} \). Let \( \mathcal{D} = \{ (\hat{\beta}_{Xi}, \hat{\beta}_{Yi}) | i = 1, \cdots, m \} \) denote the observed data.

**Assumption 1 for MR-cML.** (Plurality condition.) Suppose that \( B_0 \) is the index set of the invalid IVs with non-zero direct effects, i.e. \( r_i \neq 0 \) if and only if \( i \in B_0 \), and \( K_0 = |B_0| \). For any \( B \subseteq \{1, \cdots, m\} \) and \( |B| = K_0 \), if \( B \neq B_0 \), then there does not exist any constant \( S \) such that \( r_i = S \cdot \beta_{Xi} \) for all \( i \in B^c \).

**Assumption 2 for MR-cML.** (Orders of the variances and sample sizes.) There exist positive constants \( l_x, l_y, l_n \) and \( u_x, u_y, u_n \) such that we have \( l_x/n_1 \leq \sigma_{Xi}^2 \leq u_x/n_1, l_y/n_2 \leq \sigma_{Yi}^2 \leq u_y/n_2, \) and \( l_n \cdot n_2 \leq n_1 \leq u_n \cdot n_2 \) for \( i = 1, \cdots, m \).

Denote

\[
V = \sum_{i \in B_0} \frac{\beta_{Xi}^2}{\sigma_{Xi}^2, \beta_{Xi}, \sigma_{Yi}^2}.
\]

**Theorem 2 for MR-cML.** Under Assumptions [1] for MR-cML and [2] for MR-cML according to Definition 2.2 in [2], \( \sqrt{V}(\hat{\theta} - \theta) \xrightarrow{D} w \rightarrow N(0,1) \) as \( n_1, n_2 \rightarrow \infty \).

**Proof.** Denote \( \bar{B} = \{ i : \tilde{r}_i \neq 0 \} \) as the set of estimated invalid IVs with non-zero direct effects based on perturbed data. First we show that \( P(\bar{B} = B_0 | \mathcal{D}) \xrightarrow{p} 1 \), which is equivalent to for any \( \varepsilon > 0, \delta > 0 \), there exists \( n \) such that when \( n_1 > n, n_2 > n \) we have \( P(\bar{B} = B_0 | \mathcal{D}) < 1 - \varepsilon < \delta \). Following similar argument in Theorem 1, we could get the unconditional probability \( P(\bar{B} = B_0) \rightarrow 1 \). Suppose we could find a pair of \( \varepsilon_0 > 0, \delta_0 > 0 \) such that \( P(\bar{B} = B_0 | \mathcal{D}) < 1 - \varepsilon_0) > \delta_0 \) for arbitrarily large \( n_1, n_2 \), then we can get

\[
P(\bar{B} = B_0) = \int_{\mathcal{D}} P(\bar{B} = B_0 | \mathcal{D}) dF(\mathcal{D}) < 1 - \varepsilon_0 \delta_0,
\]

57
contradicts that \( P(\hat{B} = B_0) \to 1 \), thus we have shown that \( P(\hat{B} = B_0|\mathcal{G}) \xrightarrow{P} 1 \). Now we could focus on the case that \( \hat{B} = \hat{B} = B_0 \). Similar to [4], after profiling out \( b Xi; i \)'s in the original log-likelihood function, we have

\[
\hat{\theta} = \arg \min_{\theta} \sum_{i \in B_0} \frac{(\hat{\beta}_{yi} - \theta \cdot \hat{\beta}_{Xi})^2}{\sigma^2_{Xi} + \sigma^2_{yi}}, \quad \hat{\theta} = \arg \min_{\theta} \sum_{i \in B_0} \frac{(\hat{\beta}_{yi} - \theta \cdot \hat{\beta}_{Xi})^2}{\sigma^2_{Xi} + \sigma^2_{yi}}.
\]

(6)

Denote

\[
f(\theta) = \sum_{i \in B_0} \frac{(\hat{\beta}_{yi} - \theta \cdot \hat{\beta}_{Xi})^2}{\sigma^2_{Xi} + \sigma^2_{yi}},
\]

and

\[
\phi(\theta) = \frac{\partial f(\theta)}{\partial \theta} = \sum_{i \in B_0} \frac{(\hat{\beta}_{yi} - \theta \cdot \hat{\beta}_{Xi})(\theta \hat{\beta}_{yi} \sigma^2_{Xi} + \hat{\beta}_{Xi} \sigma^2_{yi})}{(\sigma^2_{Xi} + \sigma^2_{yi})^2}
\]

\[
= \sum_{i \in B_0} \frac{(\hat{\beta}_{yi} - \theta \cdot \hat{\beta}_{Xi})(\hat{\beta}_{yi} \cdot \xi_i \sigma^2_{Xi} + \hat{\beta}_{Xi} \cdot \xi_i \sigma^2_{yi} + (\hat{\beta}_{yi} - \theta \cdot \hat{\beta}_{Xi})(\theta \hat{\beta}_{yi} \sigma^2_{Xi} + \hat{\beta}_{Xi} \sigma^2_{yi} + \theta \xi_i \sigma^2_{Xi} + \xi_i \sigma^2_{yi})}{(\sigma^2_{Xi} + \sigma^2_{yi})^2},
\]

here \( \xi_i = \hat{\beta}_{yi} - \hat{\beta}_{Xi} \sim N(0, \sigma^2_{yi}), \xi_i = \hat{\beta}_{Xi} - \hat{\beta}_{Xi} \sim N(0, \sigma^2_{Xi}) \). We have

\[
0 = \phi(\hat{\theta}) = \phi(\hat{\theta}) + \phi'(\hat{\theta})(\hat{\theta} - \hat{\theta}) + \frac{1}{2} \phi''(\theta^*)(\hat{\theta} - \hat{\theta})^2,
\]

with \( \theta^* \) is between \( \hat{\theta} \) and \( \hat{\theta} \), thus

\[
\sqrt{V(\hat{\theta} - \hat{\theta})} = \frac{-\phi(\hat{\theta})/\sqrt{V}}{\phi'(\hat{\theta})/V + (1/2)(\hat{\theta} - \hat{\theta})\phi''(\theta^*)/V}.
\]

Next we show \( \phi(\hat{\theta})/\sqrt{V}|\mathcal{G} \xrightarrow{w.P} N(0,1) \). From equation (6), we can get

\[
\phi(\hat{\theta}) = \sum_{i \in B_0} \left( \frac{(\hat{\beta}_{yi} - \hat{\beta}_{Xi})(\hat{\beta}_{yi} \cdot \xi_i \sigma^2_{Xi} + \hat{\beta}_{Xi} \cdot \xi_i \sigma^2_{yi} + (\hat{\beta}_{yi} \cdot \theta \cdot \xi_i \sigma^2_{Xi} + \hat{\beta}_{Xi} \cdot \theta \cdot \xi_i \sigma^2_{yi})}{(\sigma^2_{Xi} + \sigma^2_{yi})^2} \right) + O_p(1),
\]

(7)

Note that \( \xi_i \)'s and \( \xi_i \)'s are \( O_p(1/\sqrt{n}), n = \min(n_1, n_2) \), thus

\[
\phi(\hat{\theta}) = \sum_{i \in B_0} \left( \frac{\xi_i (\hat{\beta}_{yi} \cdot \hat{\beta}_{Xi} \sigma^2_{Xi} - \hat{\beta}_{yi} \cdot \hat{\beta}_{Xi} \sigma^2_{Xi} + \hat{\beta}_{Xi} \cdot \hat{\beta}_{Xi} \sigma^2_{yi} - \hat{\beta}_{Xi} \cdot \hat{\beta}_{Xi} \sigma^2_{yi})}{(\sigma^2_{Xi} + \sigma^2_{yi})^2} \right) + O_p(1),
\]

(8)

so that \( V^* \xrightarrow{P} \beta_{Xi}, \hat{\beta}_{yi} \xrightarrow{P} \beta_{Xi}, \hat{\beta}_{yi} \xrightarrow{P} \theta, \theta \xrightarrow{P} \theta_0 \), we can get \( V^*/V \xrightarrow{P} 1 \), thus we get \( \phi(\hat{\theta})/\sqrt{V}|\mathcal{G} \xrightarrow{w.P} N(0,1) \).

Next we show \( -\phi'(\hat{\theta})/V|\mathcal{G} \xrightarrow{w.P} 1 \). After some calculation we get

\[
\phi'(\theta) = \sum_{i \in B_0} \left( \frac{2 \sigma^2_{Xi} \beta_{Xi} \beta_{yi} \cdot \theta^3 + 3(\sigma^2_{Xi} \beta_{Xi} \beta_{yi} \cdot \theta^3 - \sigma^2_{Xi} \beta_{Xi} \beta_{yi} \cdot \theta^3)(\sigma^2_{Xi} \beta_{Xi} \beta_{yi} \cdot \theta^3 - \sigma^2_{Xi} \beta_{Xi} \beta_{yi} \cdot \theta^3)}{(\sigma^2_{Xi} + \sigma^2_{yi})^3} \right) + O_p(1),
\]

(8)

58
as $\tilde{\beta}_{Xi} \xrightarrow{P} \beta_{Xi}$, $\tilde{\beta}_{Yi} \xrightarrow{P} \beta_{Yi}$, $\tilde{\theta} \xrightarrow{P} \theta_0$, we get $-\phi'(\hat{\theta})/V \xrightarrow{P} 1$, with Theorem 3.3 in [2], $-\phi'(\hat{\theta})/V \xrightarrow{w.p.} 1$.

Based on equation (8), we can see $\phi''(\theta)$ has its numerator of order $n^5$ and its denominator of order $n^6$, thus $\phi''(\theta^\ast)/V = O_p(1)$. As $\tilde{\theta} \xrightarrow{P} \theta_0$, $\hat{\theta} \xrightarrow{P} \theta_0$, we have $\tilde{\theta} - \hat{\theta} \xrightarrow{P} 0$, again with Theorem 3.3 in [2] we get $\tilde{\theta} - \hat{\theta} \xrightarrow{w.p.} 0$. Thus we can get $\frac{1}{2} \phi''(\theta^\ast)(\tilde{\theta} - \hat{\theta}) \xrightarrow{w.p.} 0$. Now with Theorem 3.2 in [2], we can get $\sqrt{V(\tilde{\theta} - \hat{\theta})} \xrightarrow{w.p.} N(0,1)$, completing the proof. □

References

[1] Lyon, M. S., Andrews, S. J., Elsworth, B., Gaunt, T. R., Hemani, G., & Marcora, E. (2021). The variant call format provides efficient and robust storage of GWAS summary statistics. Genome Biology, 22(1), 1-10.

[2] Xiong, S., & Li, G. (2008). Some results on the convergence of conditional distributions. Statistics & probability letters, 78(18), 3249-3253.

[3] Xue, H., Shen, X., & Pan, W. (2021). Constrained maximum likelihood-based Mendelian randomization robust to both correlated and uncorrelated pleiotropic effects. The American Journal of Human Genetics, 108(7), 1251-1269.

[4] Zhao, Q., Wang, J., Hemani, G., Bowden, J., & Small, D. S. (2020). Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. The Annals of Statistics, 48(3), 1742-1769.