Addressing **Instrument-Outcome Confounding** in **Mendelian Randomization** through **Representation Learning**

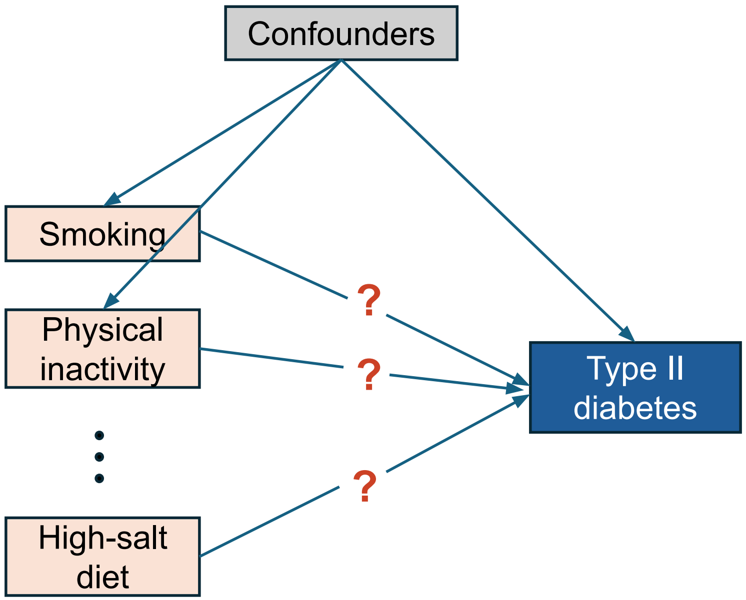
Shimeng Huang

in collaboration with Matthew Robinson and Francesco Locatello

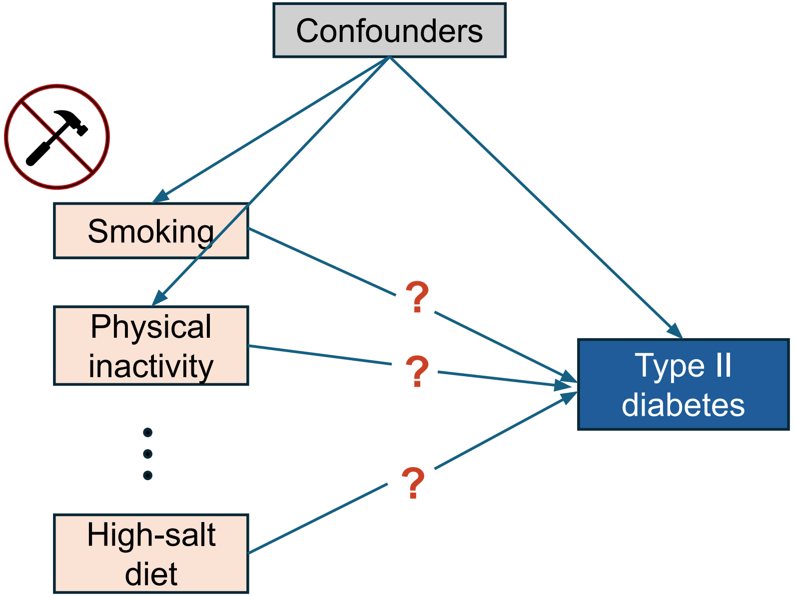
@ICSCS 2025 in Seville, Spain

Institute of Science and Technology Austria

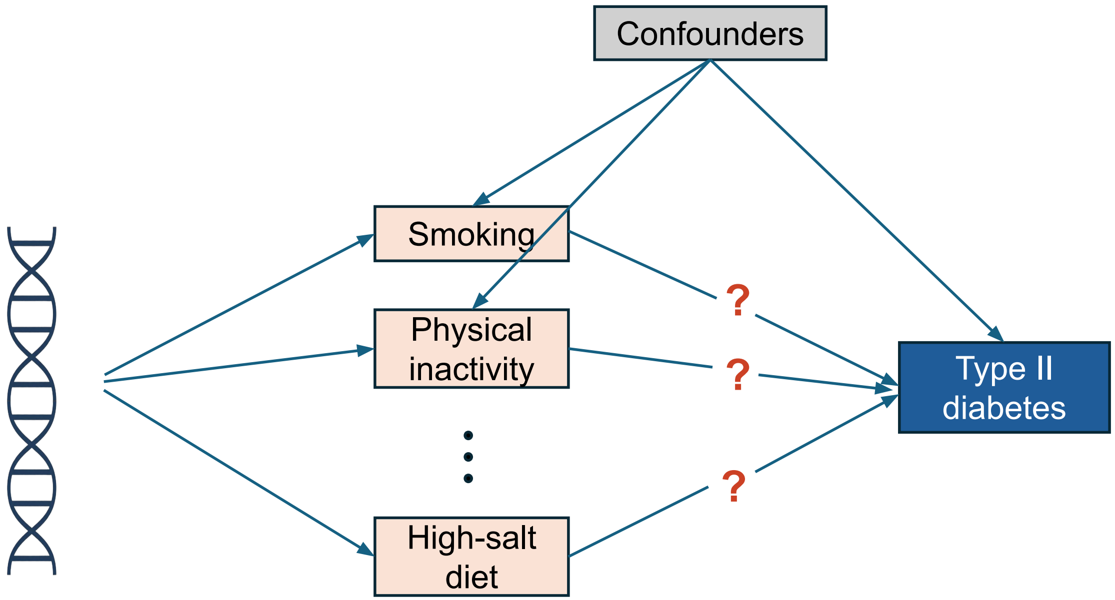
Background: Causality in Epidemiology



Background: Causality in Epidemiology



Background: Causality in Epidemiology

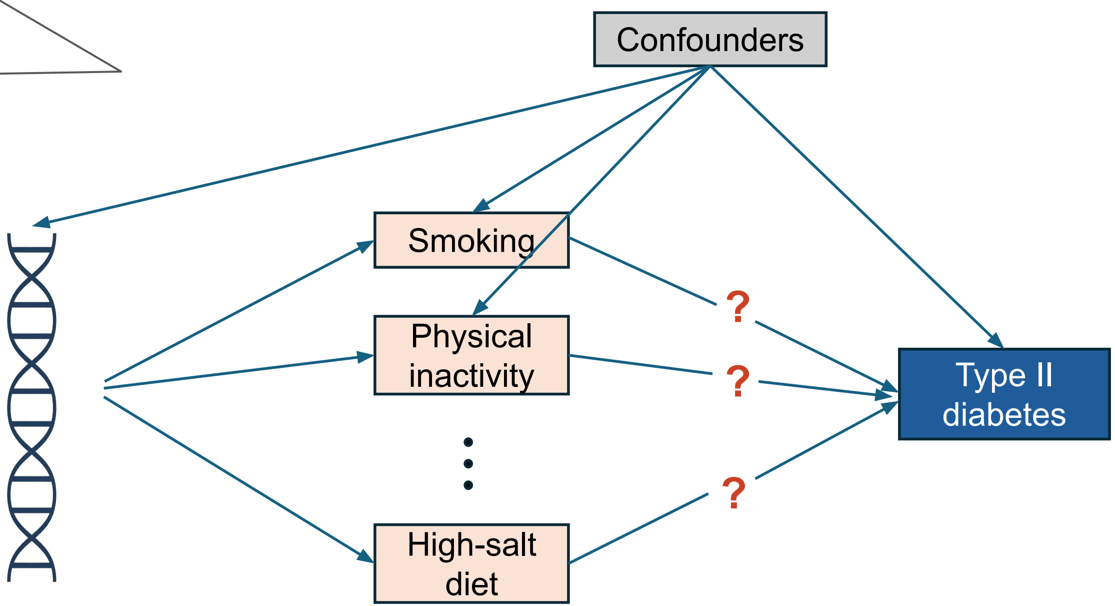


Background: Causality in Epidemiology

Instrument-outcome confounding is common in Mendelian Randomization studies

⟶Violates IV assumptions

⟶Biased causal effect



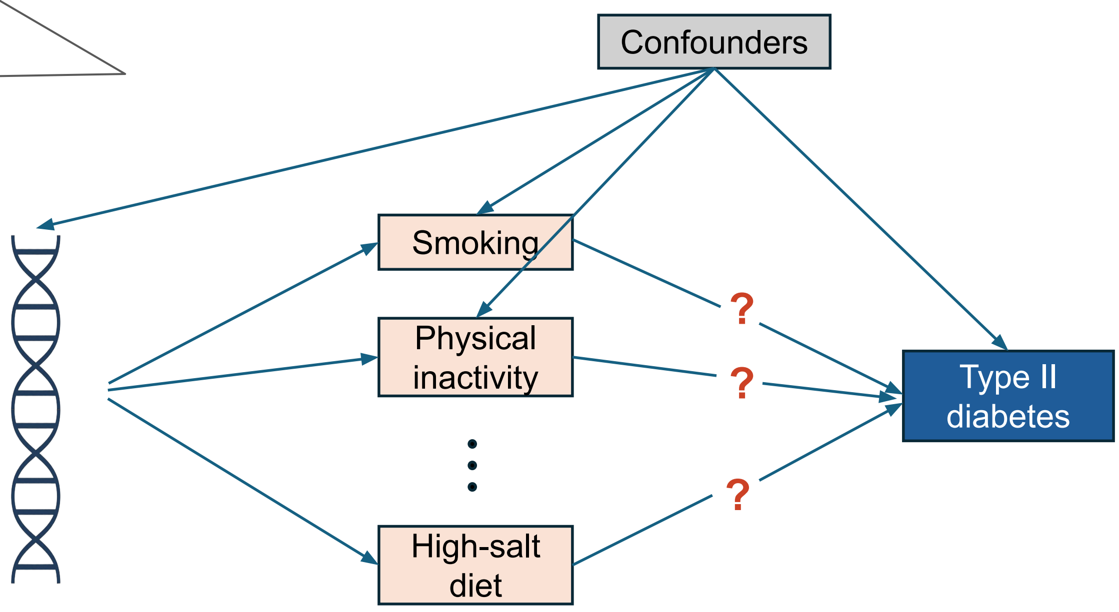
Background: Causality in Epidemiology

Instrument-outcome confounding is common in Mendelian Randomization studies

⟶Violates IV assumptions

⟶Biased causal effect

Is it possible to remove this confounding?



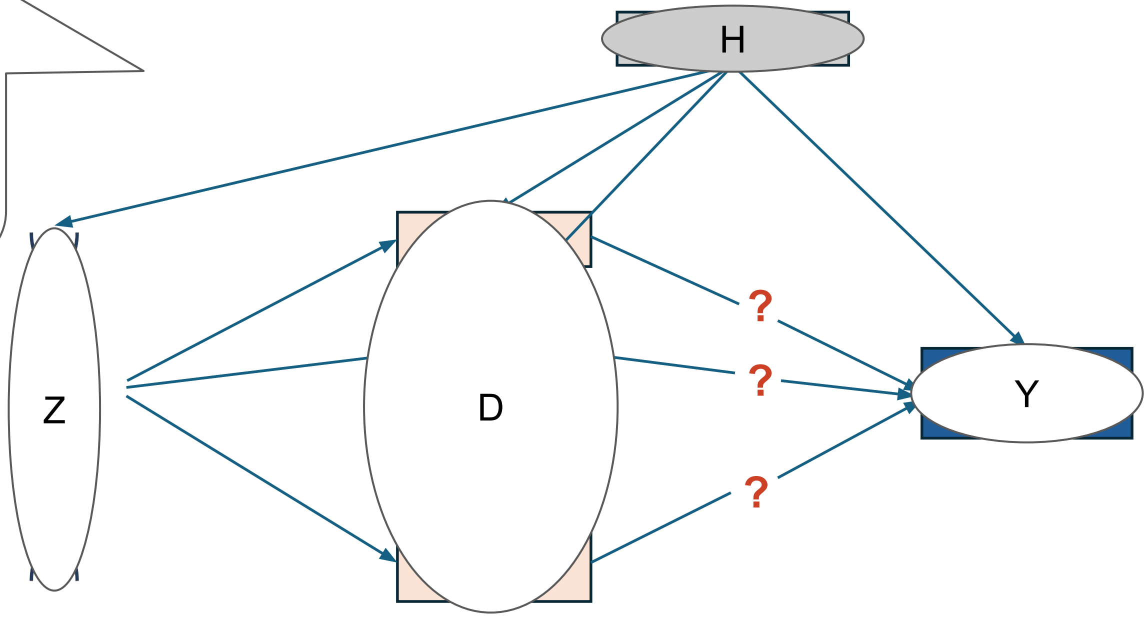
Background: Causality in Epidemiology

Instrument-outcome confounding is common in Mendelian Randomization studies

⟶Violates IV assumptions

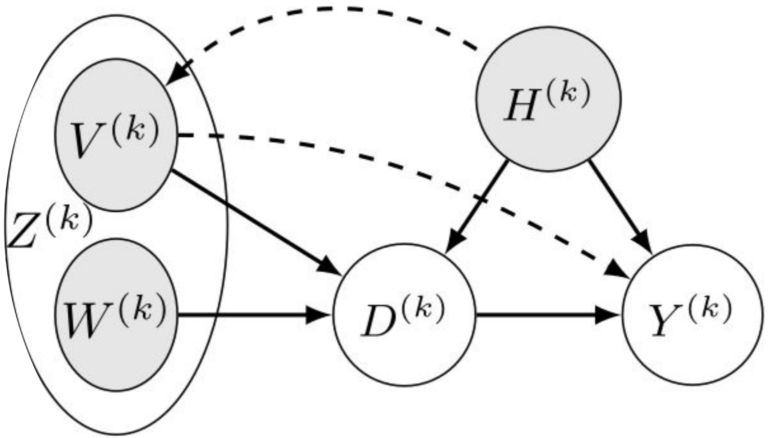
⟶Biased causal effect

Is it possible to remove this confounding?



Method: Contrastive Learning using Multi-Population Data

Suppose we can decompose Z into two parts, V and W…

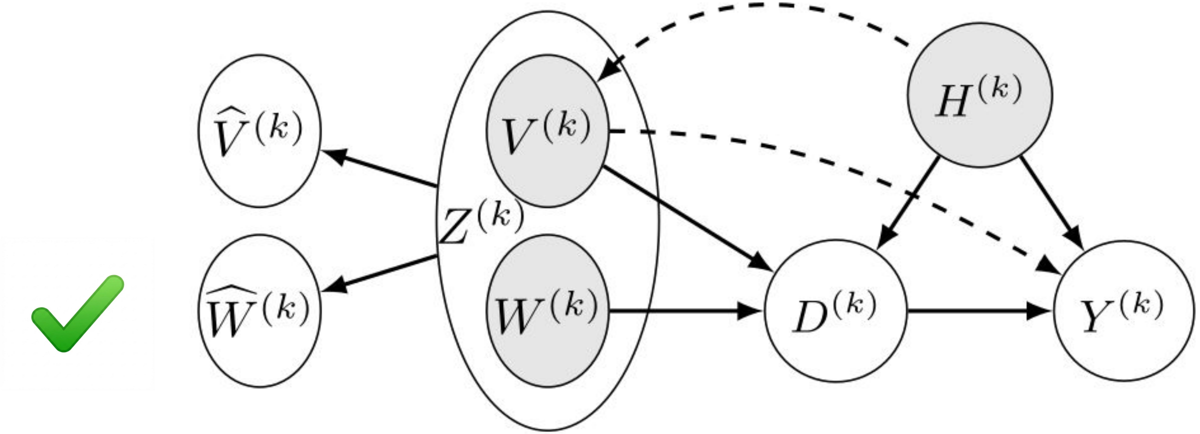


$\forall k \in [K]$

where the confounder only affects V, and we observe data from K populations.

Method: Contrastive Learning using Multi-Population Data

Suppose we can decompose Z into two parts, V and W…



$\forall k \in [K]$

where the confounder only affects V, and we observe data from K populations.

Key Findings

Identifiability of W and V: Under certain assumptions

$$\hat{W} = AW + c$$

can be learned by contrasting data from multiple populations, as well as

$$\hat{V} = BV + d$$

Key Findings

Identifiability of W and V: Under certain assumptions

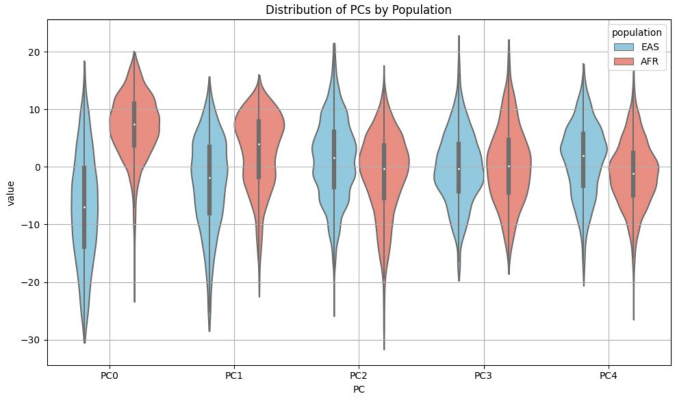
$$\hat{W} = AW + c$$

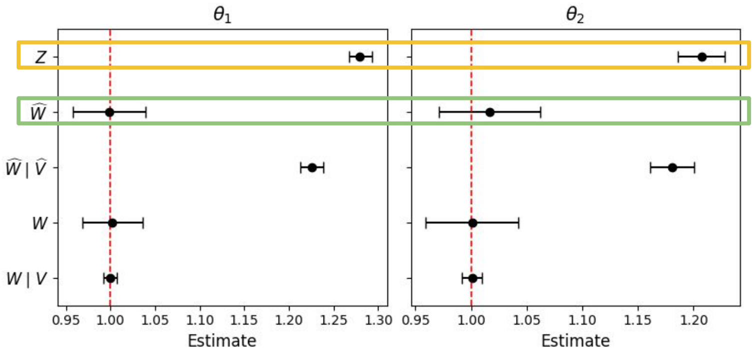
can be learned by contrasting data from multiple populations, as well as

$$\hat{V} = BV + d$$

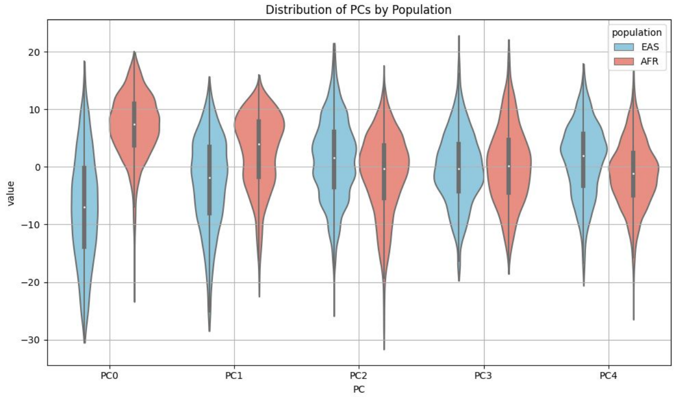
Average causal effect (ACE) estimation: $\hat{W}$ can identify the ACE as long as W can, and it *may* be helpful to control $\hat{V}$.

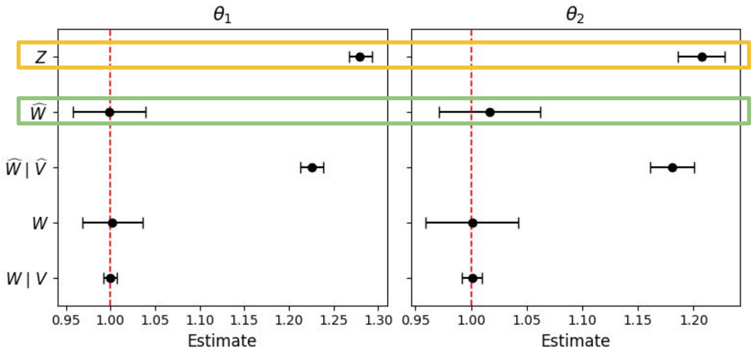
Results and Conclusion





Results and Conclusion





We leverage representation learning in MR under violation of **instrument-outcome confounding**.

We adapt and extend the theoretical results of **contrastive learning** for identifying the invariant component of the instruments.

We show empirically that autoencoders satisfying the **invariance constraint** can identify a valid instrument for MR.

Thank you for your attention!

Please feel free to reach out to shimeng.huang@ista.ac.at

My website