Discovering Mechanistic Heterogeneity using Mendelian Randomization

Qingyuan Zhao

Statistical Laboratory, University of Cambridge

Joint work with Daniel Iong (who made most of the slides) and Yang Chen

September 26, 2020 @ PCIC
Outline

1 Motivation

2 Mechanistic Heterogeneity in MR

3 MR-PATH
   • Model Assumptions
   • Statistical inference

4 Results
   • HDL-CHD
   • BMI-T2D

5 Conclusion
Mendelian randomization (MR)

- MR = Using genetic variation as instrumental variables.
- Surging interest in epidemiology and genetics.

Number of publications in MR by year (Source: Web of Science).
Example: Causal effect of the LDL-cholesterol

Basic idea: People who inherited certain alleles of \textit{rs17238484} and \textit{rs12916} have \textbf{naturally} higher concentration of LDL cholesterol.
Example: Causal effect of the LDL-cholesterol

Basic idea: People who inherited certain alleles of rs17238484 and rs12916 have naturally higher concentration of LDL cholesterol.
Motivation

Motivation for this work

- **Exclusion restriction**: Instruments (genetic variants) can only affect the outcome through the risk exposure.
  - In MR, this assumption may be violated due to pleiotropy.
  - Many pleiotropy-robust MR methods (e.g. MR-RAPS) have been developed.
- Most robust MR methods rely on the “effect homogeneity” assumption: the risk exposure has the same causal effect for every individual.

Our contributions

1. A novel concept—**Mechanistic heterogeneity**.
2. A transparent mixture model—**MR-PATH**.
For exposure $X$, outcome $Y$, unobserved confounding variables $U$, and SNPs $Z_1, \ldots, Z_p$, the commonly assumed linear structural equation model is given by

$$X = \sum_{i=1}^{p} \theta_{X_i} Z_i + \eta_X U + E_X,$$

$$Y = \beta X + \sum_{i=1}^{p} \alpha_i Z_i + \eta_Y U + E_Y.$$
**Review: Linear structural equation model for MR**

\[ X = \sum_{i=1}^{p} \theta X_i Z_i + \eta X U + E_X, \]

\[ Y = \beta X + \sum_{i=1}^{p} \alpha_i Z_i + \eta Y U + E_Y \]

- If \( Z_i \) is a valid instrument, \( \theta X_i \neq 0 \), \( Z_i \perp \{U, E_X, E_Y\} \), and \( \alpha_i = 0 \).
- However, it is often the case that \( \alpha_i \neq 0 \) due to pleiotropy and multiple causal pathways.
- If \( \alpha_i \neq 0 \) for some SNPs, then the causal effect \( \beta \) cannot be estimated consistently without further assumptions on \( \alpha_i \).
  - e.g. \( \alpha_i \sim N(0, \tau^2) \) for most SNPs.
Two scenarios of mechanistic heterogeneity

(a) Scenario 1: Multiple pathways of horizontal pleiotropy.

(b) Scenario 2: Multiple mechanisms for the exposure $X$. 
Two scenarios of mechanistic heterogeneity

If we interpret the diagrams in the previous slide as linear structural equations as before, we can derive the Wald estimands for each pathway.

<table>
<thead>
<tr>
<th>Instruments</th>
<th>Pathway</th>
<th>Effect of M on X</th>
<th>Effect of M on Y</th>
<th>Wald estimand</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z_{1,1}, \ldots, Z_{1,p_1}$</td>
<td>$M_1$</td>
<td>$\theta_1$</td>
<td>$\theta_1 \beta$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>$Z_{2,1}, \ldots, Z_{2,p_2}$</td>
<td>$M_2$</td>
<td>$\theta_2$</td>
<td>$\theta_2 \beta + \alpha_2$</td>
<td>$\beta + \alpha_2/\theta_2$</td>
</tr>
<tr>
<td>$Z_{3,1}, \ldots, Z_{3,p_3}$</td>
<td>$M_3$</td>
<td>$\theta_3$</td>
<td>$\theta_3 \beta + \alpha_3$</td>
<td>$\beta + \alpha_3/\theta_3$</td>
</tr>
</tbody>
</table>

- SNPs on the same pathway have the same Wald estimand, while SNPs across different pathways generally have different estimands.
- Mechanistic heterogeneity can arise even when all SNPs are valid instruments (Scenario 2).
Mechanism-specific causal effect

The same clustering phenomenon also occurs in nonlinear models.

- It is well known that assuming **monotonicity**, IV nonparametrically estimates the **complier average treatment effect** (Angrist et al., *JASA*, 1996).

- By assuming **monotonicity** and Pearl’s nonparametric structural equation model with independent errors (**NPSEM-IE**), our paper showed that (if $X, Z, M$ are all binary variables)

\[
\begin{align*}
\mathbb{E}[Y(X = 1) - Y(X = 0) \mid X(Z_{kj} = 1) > X(Z_{kj} = 0)] = & \\
= \mathbb{E}[Y(X = 1) - Y(X = 0) \mid X(M_k = 1) > X(M_k = 0)],
\end{align*}
\]

where $k$ indexes the mechanism and $j$ indexes the gene within.
**Assumption** (Error-in-variables regression)

The observed SNP-exposure and SNP-outcome associations are distributed as

\[
\begin{pmatrix}
\hat{\theta}_X_i \\
\hat{\theta}_Y_i
\end{pmatrix}
\overset{\text{indep.}}{\sim} N\left(
\begin{pmatrix}
\theta X_i \\
\beta_i \theta X_i
\end{pmatrix},
\begin{pmatrix}
\sigma^2 X_i & 0 \\
0 & \sigma^2 Y_i
\end{pmatrix}
\right), \quad i = 1, \ldots, p,
\]

where \(\sigma X_i, \sigma Y_i\) are (fixed) measurement errors.

**Assumption** (Mixture model for mechanistic heterogeneity)

\[
Z_i \sim \text{Categorical} \left(\pi_1, \ldots, \pi_K\right),
\]

\[
\beta_i | Z_i = k \sim N(\mu_k, \sigma^2_k), \quad k = 1, \ldots, K.
\]
MR-PATH: Statistical Inference

1. Monte-Carlo EM algorithm for obtaining model parameter estimates
2. Approximate confidence intervals for quantifying uncertainty of the estimates
3. Modified Bayesian Information criterion (BIC) for selecting number of clusters

- We perform simulation studies to verify the efficacy of these inference procedures.
- See paper for implementation details.
**Example:** HDL-CHD

**Data (Three-sample MR design)**

- **Selection dataset:** Teslovich et al. 2010<sup>1</sup>
- **Exposure dataset:** Kettunen et al. 2016<sup>2</sup>
- **Outcome dataset:** Nikpay et al. 2015<sup>3</sup>

---


**Example: HDL-CHD**

Results of MR-RAPS.

Qingyuan Zhao (Cambridge)
Example: HDL-CHD

Results of MR-PATH (http://danieliong.me/mr-path/.)
Example: HDL-CHD

Results

HDL-CHD

95% Posterior Credible Interval

Cluster membership prob.

Qingyuan Zhao (Cambridge)

MR-PATH

September 26, 2020 @ PCIC
Example: HDL-CHD
**Example: BMI-T2D**

**Data (Three-sample MR design)**

- **Selection dataset**: Akiyama et al. 2017\(^1\)
- **Exposure dataset**: Locke et al. 2015\(^2\)
- **Outcome dataset**: Mahajan et al. 2018\(^3\)

---


Example: BMI-T2D

Results of MR-RAPS.
Example: BMI-T2D

Results of MR-PATH.
Example: BMI-T2D
Example: BMI-T2D

SNP association with peak blood insulin

SNP-specific slope

rs9068222
rs7903146
rs6444082
rs2237892
rs7020996
rs7923837
rs7020996
rs10906111

1 2
Concluding remarks

- A few other related methods:
  - MR-Clust: Constructs mixture model based on SNP-specific Wald estimators.
  - GRAPPLE: A visualization tool that does not attempt to model different mechanisms explicitly.
  - BESIDE-MR: A Bayesian model averaging approach extends the profile likelihood used in MR RAPS.

- Advantages of MR-PATH:
  - Does not require individually strong instruments.
  - Accounts for measurement error in the summary data.
  - An interpretable generative model for multiple causal mechanisms.
  - Potential extensions to multivariable MR with correlated SNPs.

- Further information: http://danieliong.me/mr-path/.