

Mendelian Randomisation as an Instrumental Variable Approach to Causal Inference

Vanessa Didelez
School of Mathematics
University of Bristol

MRC project “Inferring epidemiological causality using Mendelian randomisation ”
With T.Palmer, N.Sheehan, S.Meng, R.Harbord, F.Windmeijer, J.Stern, D.Lawlor,
G.Davey Smith

Bordeaux, June 2011

Based on:

DIDELEZ, SHEEHAN (2007). Mendelian randomisation as an instrumental variable approach to causal inference. *Statistical Methods in Medical Research*, **16**, 309-330.

DIDELEZ, SHEEHAN (2007). Mendelian randomisation: why epidemiology needs a formal language for causality. *In: F.Russo and J.Williamson (eds.), Causality and Probability in the Sciences, College Publications London.*

DIDELEZ, MENG, SHEEHAN (2010). Assumptions of IV Methods for Observational Epidemiology. *Statistical Science* **25**, 22-40.

PALMER, STERNE, HARBORD, LAWLOR, SHEEHAN, MENG, GRANELL, DAVEY SMITH, DIDELEZ (2011). Instrumental variable estimation of the causal risk ratio and causal odds ratio in Mendelian randomization analyses. *American Journal of Epidemiology*, **173**(12), 1392-1403. DOI: 10.1093/aje/kwr026

Overview

- **Part 1:** Mendelian Randomisation and Instrumental Variables
- **Part 2:** A Closer Look at the IV Assumptions
- **Part 3:** Using IV to Estimate a Causal Effect

Part 1:

Mendelian Randomisation and Instrumental Variables

- Motivation and basic idea
- Example: effect of alcohol consumption on health outcome using ALDH2 genotype
- Formal definition of IV

Motivation

Epidemiology interested in effect of interventions ('drink less alcohol', 'eat folic acid' etc.)

Observational studies are inevitable: preliminary research, but also assessment of effects in general population.

Obvious problem is **confounding**: effects of interest are entangled with many other effects — this can never be fully excluded.

Instrumental variables allow *some* inference on effects of interventions in the presence of confounding.

Problem with this is: how to find a suitable instrument? It has recently become popular to look for a genetic variant as IV — **Mendelian randomisation**.

Motivation

Epidemiology interested in effect of interventions ('drink less alcohol', 'eat folic acid' etc.)

Observational studies are inevitable: preliminary research, but also assessment of effects in general population.

Obvious problem is **confounding**: effects of interest are entangled with many other effects — this can never be fully excluded.

Instrumental variables allow *some* inference on effects of interventions in the presence of confounding.

Problem with this is: how to find a suitable instrument? It has recently become popular to look for a genetic variant as IV — **Mendelian randomisation**.

Motivation

Epidemiology interested in effect of interventions ('drink less alcohol', 'eat folic acid' etc.)

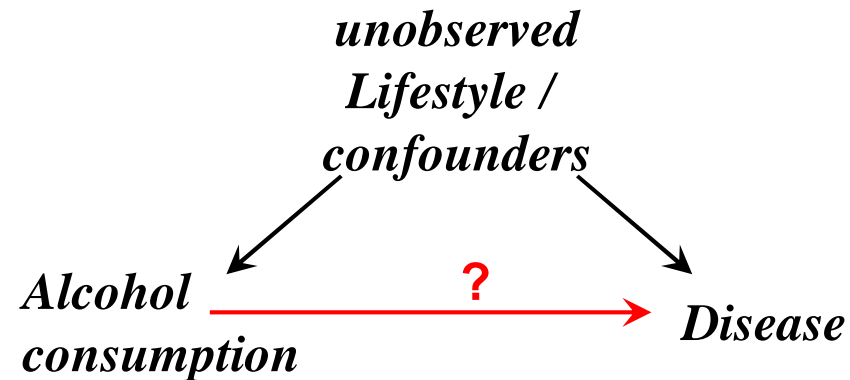
Observational studies are inevitable: preliminary research, but also assessment of effects in general population.

Obvious problem is **confounding**: effects of interest are entangled with many other effects — this can never be fully excluded. **Example**

Instrumental variables allow *some* inference on effects of interventions in the presence of confounding.

Problem with this is: how to find a suitable instrument? It has recently become popular to look for a genetic variant as IV — **Mendelian randomisation**.

Example: Alcohol Consumption



Chen et al. (2008)

Alcohol consumption has been found in observational studies to have positive 'effects' (coronary heart disease) as well as negative 'effects' (liver cirrhosis, some cancers, mental health problems).

But also strongly associated with all kinds of confounders (lifestyle etc.), as well as subject to self-report bias. Hence doubts in causal meaning of above 'effects'.

Motivation ctd.

Epidemiology interested in effect of interventions ('drink less alcohol', 'eat folic acid' etc.)

Observational studies are inevitable: preliminary research, but also assessment of effects in general population.

Obvious problem is **confounding**: effects of interest are entangled with many other effects — this can never be fully excluded.

Instrumental variables allow *some* inference on effects of interventions in the presence of confounding.

Problem with this is: how to find a suitable instrument? It has recently become popular to look for a genetic variant as IV — **Mendelian randomisation**.

Motivation ctd.

Epidemiology interested in effect of interventions ('drink less alcohol', 'eat folic acid' etc.)

Observational studies are inevitable: preliminary research, but also assessment of effects in general population.

Obvious problem is **confounding**: effects of interest are entangled with many other effects — this can never be fully excluded.

Instrumental variables allow *some* inference on effects of interventions in the presence of confounding.

Problem with this is: how to find a suitable instrument? It has recently become popular to look for a genetic variant as IV — **Mendelian randomisation**.

Mendelian Randomisation: Basic Idea

Idea: if we cannot randomise, let's look for instances where 'nature' has randomised, e.g. through genetic variation.

Example: Alcohol Consumption

Genotype: ALDH2 determines blood acetaldehyde, the principal metabolite for alcohol.

Two alleles/variants: wildtype *1 and "null" variant *2.

*2*2 homozygous individuals suffer facial flushing, nausea, drowsiness and headache after alcohol consumption.

⇒ *2*2 homozygous individuals have low alcohol consumption *regardless* of their other lifestyle behaviours

IV-Idea: check if these individuals have a different risk than others for alcohol related health problems!

Effect of Interventions — Formally

Want notation to distinguish between **association** and **causation**.

Intervention: **setting** X to a value x denoted by $\sigma_X = x$.

$p(y; \sigma_X = x)$ not necessarily the same as $p(y|X = x; \sigma_X = \emptyset)$.

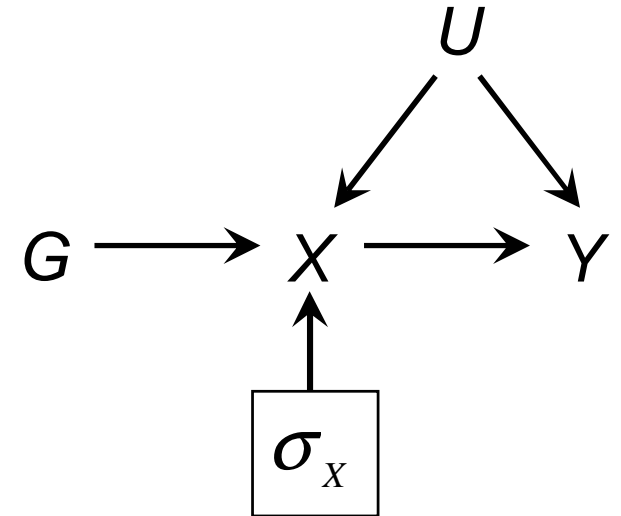
- $p(y; \sigma_X = x)$, if it depends on x then regard X as causal for Y
 \Rightarrow typically observed in a randomised study.
Note: this is a *total* (or population) effect.
- $p(y|X = x; \sigma_X = \emptyset)$ will also depend on x when there is confounding, reverse causation etc.
 \Rightarrow typically observed in an observational study.

Note: for simple intervention we often use **Pearl's** $do(X = x)$ notation.

Definition of IV

Definition of **instrumental variable**

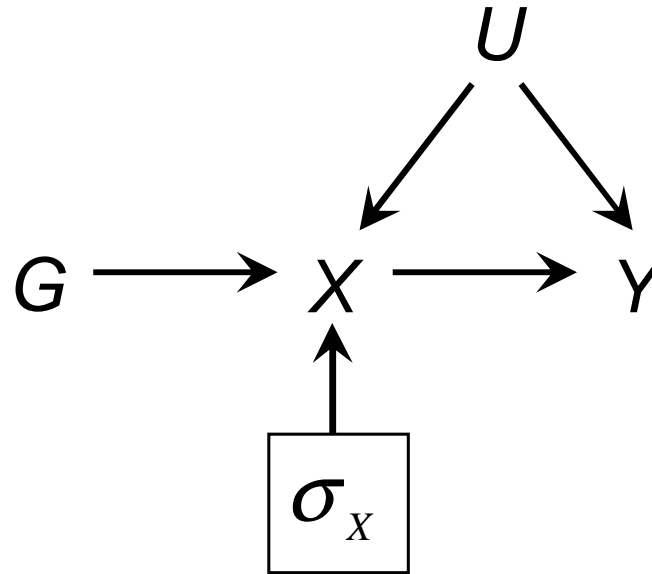
1. $G \perp\!\!\!\perp U \mid \sigma_X = \emptyset$
2. $G \not\perp\!\!\!\perp X \mid \sigma_X = \emptyset$
3. $G \perp\!\!\!\perp Y \mid (X, U, \sigma_X = \emptyset)$.



And if the following **structural** assumptions are valid:

$$Y \perp\!\!\!\perp \sigma_X \mid (X, U), \quad G \perp\!\!\!\perp \sigma_X \quad \text{and} \quad U \perp\!\!\!\perp \sigma_X.$$

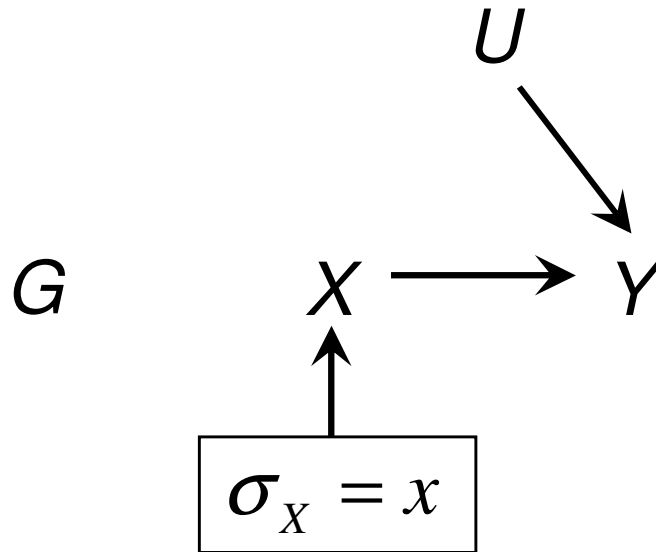
Instrumental Variables — Graphically



Equivalent to factorisation

$$p(y, x, u, g; \sigma_X) = p(y|x, u)p(x|u, g; \sigma_X)p(u)p(g)$$

Instrumental Variables — Graphically



With **structural** assumption: under intervention in X

$$p(y, u, g; \sigma_X = \tilde{x}) = p(y|\tilde{x}, u; \sigma_X = \emptyset)p(u)p(g)$$

Note 1: implies $Y \perp\!\!\!\perp G \mid \sigma_X \neq \emptyset$ – also known as **exclusion restriction**.

Note 2: exclusion restriction does not refer to U .

Why does IV Help with Causal Inference?

Testing:

check if $Y \perp\!\!\!\perp G$ — this is (roughly) testing whether there is a causal effect at all.

Estimation:

(1) when all observable variables are discrete, we can obtain **bounds** on causal effects without further assumptions.

(2) for point estimates need some (semi-)parametric / structural assumptions, as well as clear definition of target causal parameter.

Part 2:

A Closer Look at the IV Assumptions

- How can we justify the assumptions?
- Use IV to test for a causal effect
- Possible violations of the IV

'Untestable' Assumptions

The assumptions

1. $G \perp\!\!\!\perp U$

3. $G \perp\!\!\!\perp Y \mid (X, U)$.

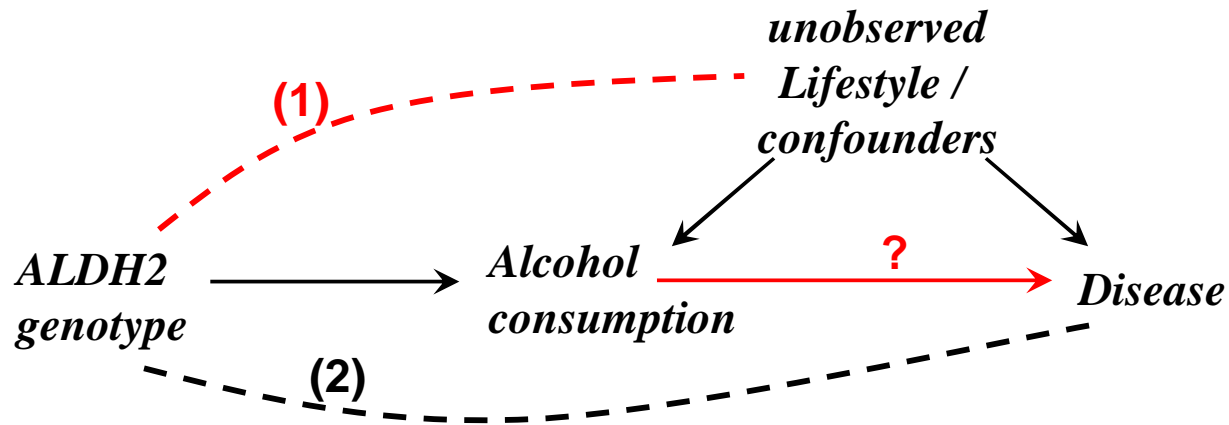
impose inequality constraints when X, Y, G discrete.

They do not imply that $G \perp\!\!\!\perp Y \mid X$ or $G \perp\!\!\!\perp Y$!

⇒ Need to be justified based on subject matter background knowledge.

⇒ **Mendelian randomisation**

Example: Alcohol Consumption

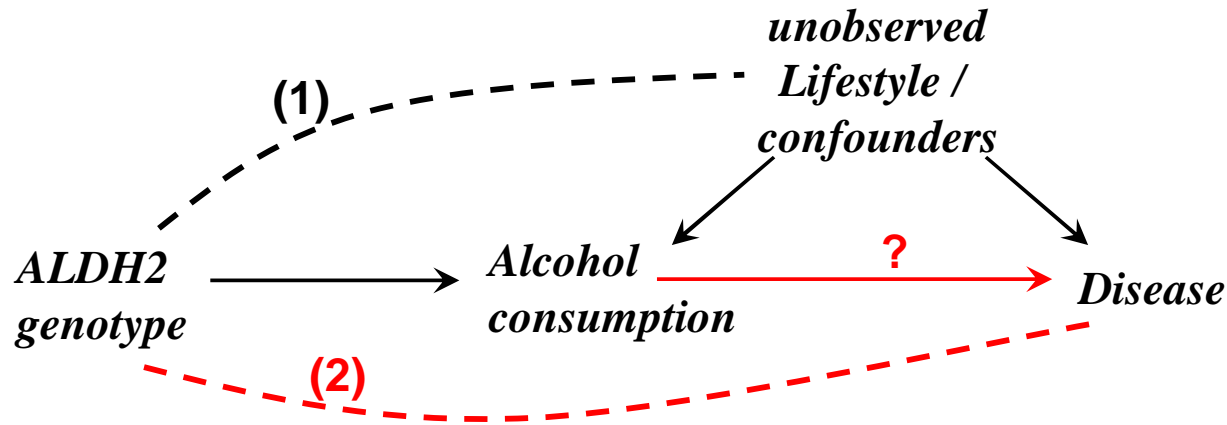


Note 1: due to random allocation of genes at conception, can be fairly confident that genotype is not associated with unobserved confounders (subpopulation structure can be a problem).

Further evidence: in extensive studies no evidence for association with *observed* confounders, e.g. age, smoking, BMI, cholesterol.

(see also Davey Smith et al., 2007)

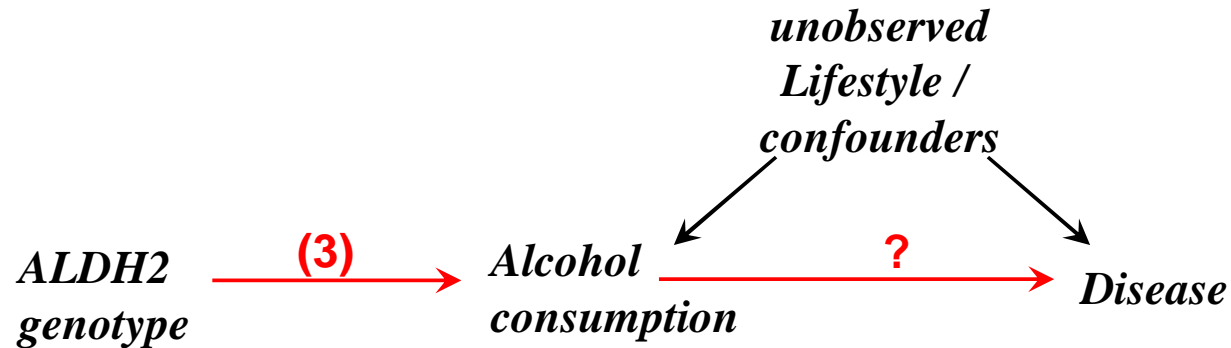
Example: Alcohol Consumption



Note 2: due to known ‘functionality’ of ALDH2 gene, we can exclude that it affects the typical diseases considered by *another* route than through alcohol consumption.

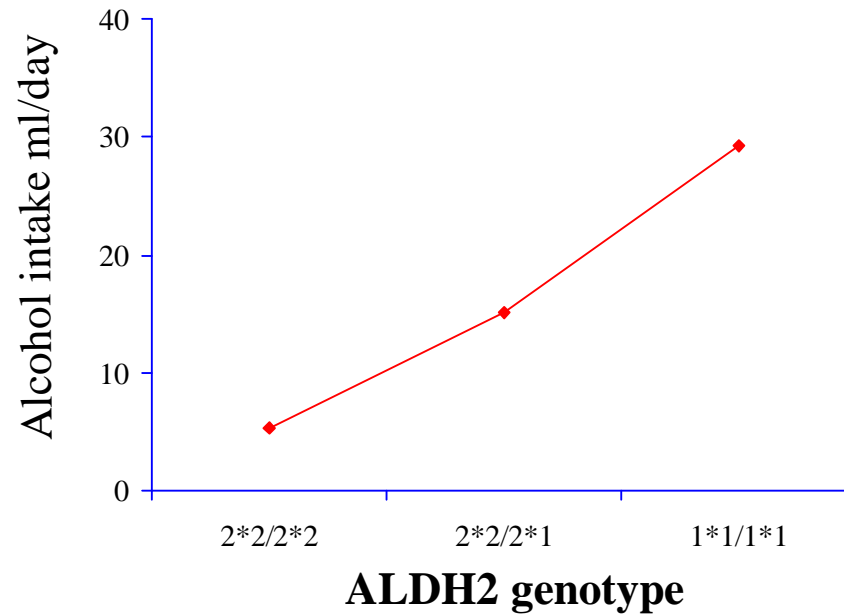
⇒ important to use well studied genes as instruments!

Example: Alcohol Consumption



Note 3: association of ALDH2 with alcohol consumption well established, strong, and underlying biochemistry well understood.

Example: Alcohol Consumption



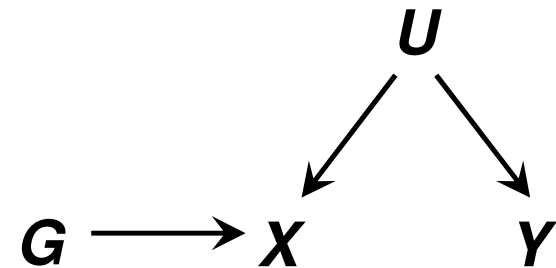
Note 3: association of ALDH2 with alcohol consumption well established, strong, and underlying biology well understood.

Using IV to Test for Causal Effect

simply test if

$$Y \perp\!\!\!\perp G$$

i.e. independence between
instrument and outcome

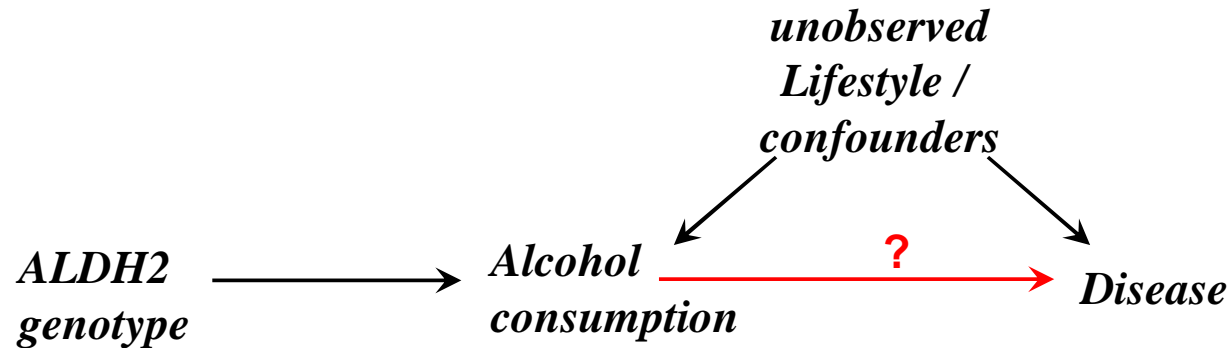


Message: regardless of measurement level, testing $Y \perp\!\!\!\perp G$ is valid test for presence of causal effect of X on Y ;

no further parametric assumptions required!

(Exception to the rule: if U acts as effect modifier in a very specific way.)

Example: Alcohol Consumption



Causal Effect? under IV assumptions, the null-hypothesis of no causal effect of alcohol consumption, should imply no association between *ALDH2* and disease;

While if alcohol consumption has a causal effect we would expect an association between *ALDH2* and disease.

Example: Alcohol Consumption

Findings:

(Meta-analysis by Chen et al., 2008)

Blood pressure on average 7.44mmHg higher and risk of hypertension 2.5 higher for ALDH2*1*1 than for ALDH2*2*2 carriers (only males).
⇒ mimics the effect of *large versus low* alcohol consumption.

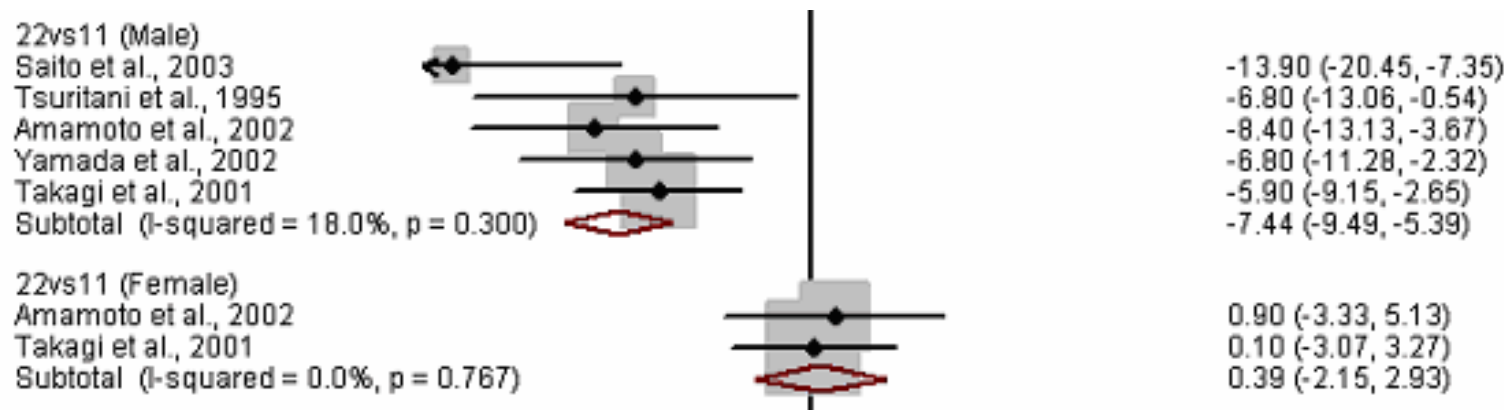
Blood pressure on average 4.24mmHg higher and risk of hypertension 1.7 higher for ALDH2*1*2 than for ALDH2*2*2 carriers (only males).
⇒ mimics the effect of *moderate versus low* alcohol consumption.
⇒ it seems that **even moderate** alcohol consumption is **harmful**.

Note: studies mostly in Japanese populations (where ALDH2*2*2 is common), where women drink only little alcohol in general.

Example: Alcohol Consumption

(Chen et al., 2008)

Is condition $Y \perp\!\!\!\perp G | (X, U)$ satisfied?



Some indication

Women in Japanese study population do not drink. ALDH2 genotype in women not associated with blood pressure \Rightarrow there does not seem to be another pathway creating a $G-Y$ association here.

Violations of Core Conditions

So far have argued that in Mendelian randomisation studies we can reasonably believe that core conditions are *satisfied*.

Now will discuss situations and examples when they are typically *violated*.

(Davey Smith & Ebrahim, 2003)

Use **graphs** to represent assumptions / background knowledge to check if core conditions are violated or not.

Population Stratification

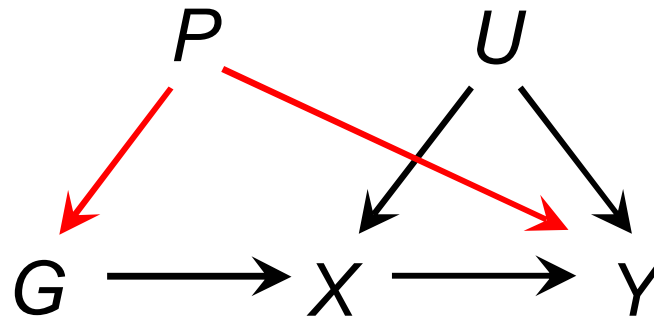
Population stratification occurs when there exist population subgroups that experience both, different disease rates (or different distributions of phenotypes) and have different frequencies of alleles of interest.

⇒ might violate condition $Y \perp\!\!\!\perp G | (X, U)$.

Population Stratification (1)

Graphical representation:

variable P = population predicts G as well as Y



Here $Y \not\perp\!\!\!\perp G | (X, U)$!

\Rightarrow can be avoided by sensible study design.

Population Stratification (1)

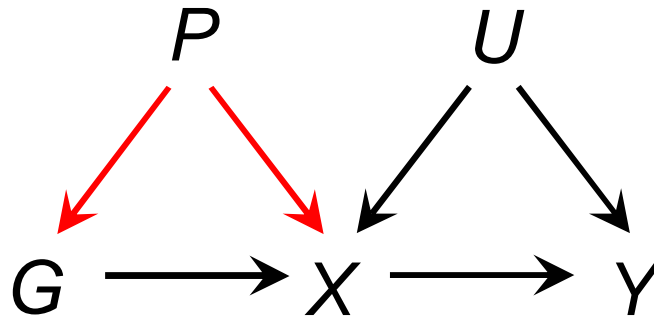
Example: Study of native Americans from Pima and Papago tribe.

(Knowler et al, 1988)

- Strong inverse association between a HLA–haplotype and type 2 diabetes;
 - Individuals with full American Indian heritage: haplotype prevalence 1% and type 2 diabetes prevalence 40%;
 - In Caucasian population: haplotype prevalence 66% and type 2 diabetes prevalence 15%.
- ⇒ Solution: carry out analyses *within population strata!*
- **Note:** need to be aware of existence of such population subgroups.

Population Stratification (2)

Also possible: phenotype distribution different in subgroups



Core conditions still satisfied.

Strength of IV could be affected in positive or negative way.

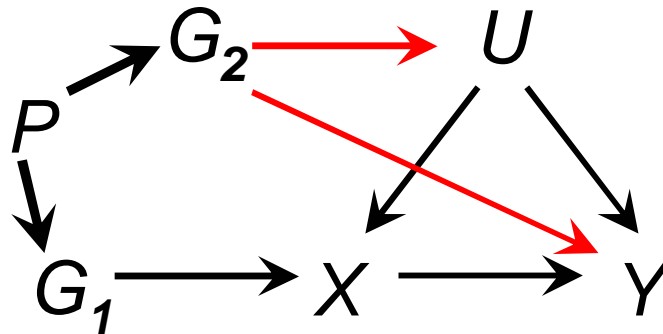
Linkage Disequilibrium (1)

Linkage disequilibrium (LD): is the correlation between allelic states at different loci, traditionally regarded as stemming from close proximity to each other on chromosome.

If genetic variant chosen as instrument is in LD with another gene that is in turn associated with some of the unobserved confounders or even predicts the disease, then condition $Y \perp\!\!\!\perp G | (X, U)$ might again be violated.

Linkage Disequilibrium (1)

Graphical representation: $G_1 =$ chosen instrument, in LD with other gene G_2 through parental genes P .



Here $Y \not\perp\!\!\!\perp G | (X, U)$ or $G \not\perp\!\!\!\perp U$!

Linkage Disequilibrium (2)

'Right' gene?

Often it is plausible that the gene chosen as instrument is not *the* causal gene for the phenotype of interest, but instead it is in LD with the causal gene.

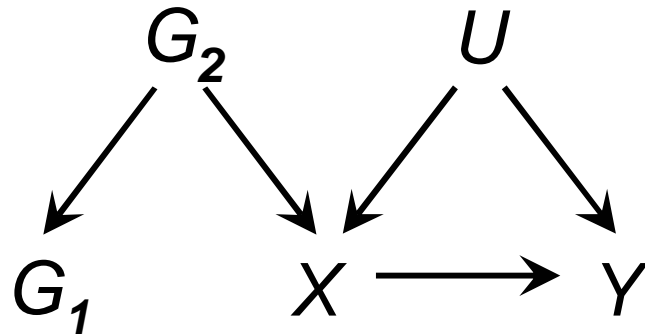
We could also regard this as **measurement error** when assessing the gene.

This does not necessarily imply any violations of the core IV conditions.

⇒ we do not need to use the causal gene for Mendelian randomisation, just one that is correlated with the phenotype, as long as core conditions satisfied.

Linkage Disequilibrium (2)

Chosen gene G_1 is not 'causal'



Core conditions still satisfied.

Again, strength of IV could be affected.

Pleiotropy

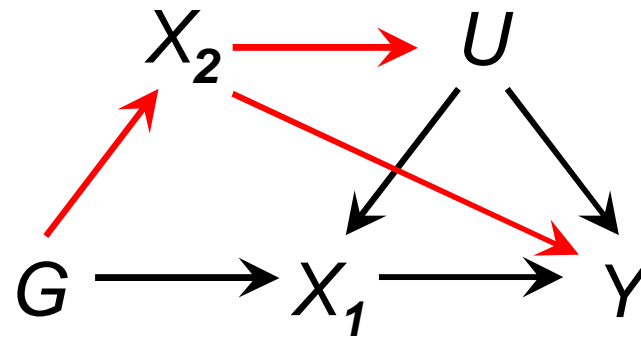
Pleiotropy refers to a genetic variant having multiple functions, i.e. the chosen gene / instrument might not only affect the phenotype of interest but also other traits.

If the pleiotropic effects influence or predict the outcome through other pathways, then the IV core conditions might again be violated.

Pleiotropy

Graphical representation:

G affects another phenotype X_2 (phenotype of interest is X_1)



Here, again, $Y \not\perp\!\!\!\perp G \mid (X, U)$ or $G \not\perp\!\!\!\perp U$.

Pleiotropy

Example: use of APOE genotype as IV for causal effect of both types of cholesterol (HDLc and LDLc) on myocardial infarction risk.

- Causal effect of HDLc and LDLc well known from RCTs.
 - APOE strongly associated with HDLc and LDLc, but surprisingly APOE not associated with myocardial infarction risk.
 - Explanation: the $\epsilon 2$ variant of APOE is also related to less efficient transfer of very low-density lipoproteins and chylomicrons from the blood to the liver, greater postprandial lipaemia, and an increased risk of type III hyperlipoproteinaemia → all of which increase myocardial infarction risk.
- ⇒ Due to multiple effects of APOE, all affecting myocardial infarction risk, this is an unsuitable IV for a Mendelian randomisation study.

Genetic Heterogeneity

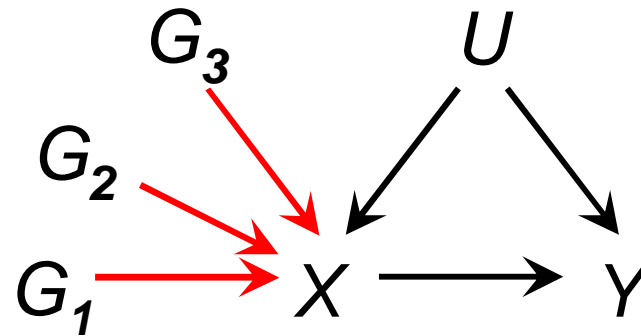
Genetic heterogeneity means that more than one genotype affect or determine the phenotype of interest, possibly via different biochemical routes.

To decide if this is a problem we need to establish whether these are in LD and whether any of them have pleiotropic effects.

Genetic Heterogeneity

Graphical representation:

Independent genes affecting the phenotype.



All core IV conditions still satisfied for each (or any subset of) genes.

Genetic Heterogeneity

Potential **Benefits**:

- Multiple instruments can be used to examine if the core conditions are violated. If different genes affect the phenotype via different pathways, it is unlikely that both are affected in the same way by population stratification, LD or pleiotropy.
- ⇒ compare results obtained for each genotype separately.
- CRP example: circulating CRP is related to *CRP*-gene but also to other locus, *IL6*, via different pathways.
 - Can also combine multiple instruments for *one* IV analysis (common in econometrics).

Measurement Errors

All measurements in a Mendelian randomisation study are prone to measurement error.

In particular G : we often do not have the 'right' causal gene, but just a strongly linked locus (cf. LD).

But G might also be mis-measured for other reasons \Rightarrow "in principle" not a problem if measurement error not differential.

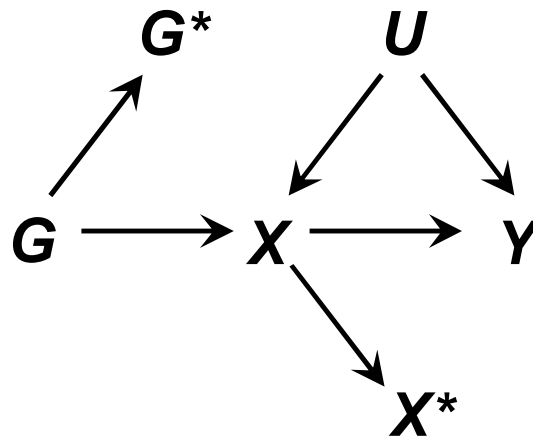
But might be a problem if $G-X$ association is obtained from a different study than $G-Y$ association (meta analysis).

In practice, X is also typically measured with error. The 'true' phenotype through which G acts is e.g. lifelong exposure to high or low CRP levels, while what is measured is only the value at a particular time.

Measurement Errors

Graphical representation:

G^* and X^* are the actual measurements of G and X with possible measurement errors.



Can use G^* instead of G but not X^* instead of X as $G^* \not\perp\!\!\!\perp Y | (X^*, U)$.

Note: H_0 'no causal effect' can be tested by $Y \perp\!\!\!\perp G$ regardless of X

Finding Genetic Instrumental Variable

Main limitation:

finding genetic variant that is suitable as instrumental variable!

Not many are known yet for typical exposures of interest in epidemiology.

Optimism: rapid expansion of knowledge in functional genomics!

Genome wide association studies:

gene–phenotype associations often weak, low power, not reproducible.

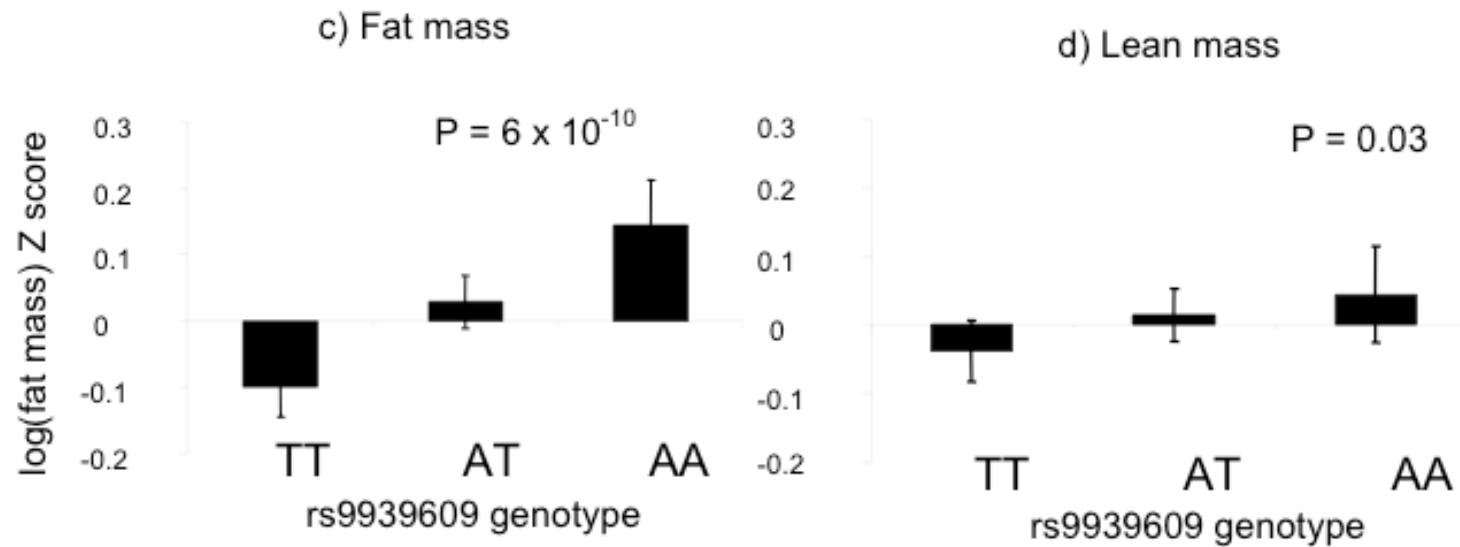
But even when strong/reproducible association found, **functionality of genes not well understood** if only based on association studies.

⇒ cannot be as confident that core conditions satisfied.

Example: *FTO* genotype associated with BMI / fat mass but functionality not (yet?) understood. (Frayling et al., 2007)

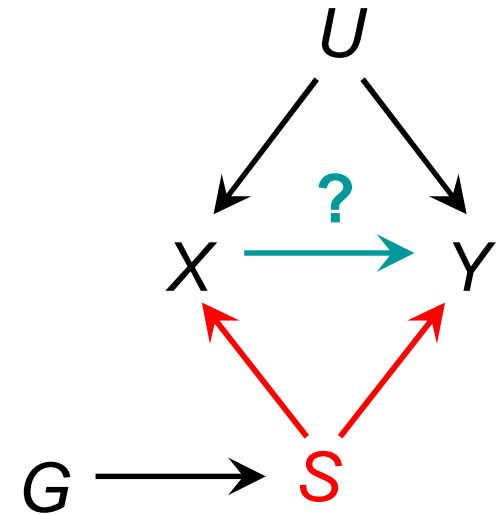
Example: FTO — fat/lean mass

From genome wide association studies:



Finding Genetic IV — Association Studies

Alternative explanation: G causes a condition S , which in turn causes the phenotype and disease of interest.



Here $G \not\perp\!\!\!\perp X$ (genotype and phenotype are associated), $G \perp\!\!\!\perp U$.

But $G \not\perp\!\!\!\perp Y | (X, U)$ — and cannot check this from data.

Part 3:

Using IV to Estimate Causal Parameters

- Target parameters
- Classical 2-stage-least-squares
- Methods for binary outcomes
- Application: effect of BMI on asthma in children

Estimating Causal Parameters with IVs

There is a *multitude* of methods using an IV to obtain an estimate for the causal effect of interest.

IV-based estimation methods differ in

- target causal parameter,
- model assumptions.

Will only consider a subjective selection of methods here.

Target Causal Parameters (1)

Population causal parameters:

Contrast between $p(y; \sigma_X = x)$ for different values of x , e.g.

ACE average causal effect $E(Y; \sigma_X = x_1) - E(Y; \sigma_X = x_0)$

CRR causal risk ratio $E(Y; \sigma_X = x_1) / E(Y; \sigma_X = x_0)$

or COR or...

Example: folic acid, hormone replacement therapy etc.

Target Causal Parameters (2)

Effect of **‘treatment on the treated’** or **‘exposure on exposed’**

Let X = ‘natural’ value of exposure (e.g. alcohol consumption)
and X^* = exposure due to manipulation.

⇒ Contrast $P(Y|X = x)$ ‘no intervention’
with $P(Y|X = x; \sigma_{X^*} = 0)$ ‘enforcing baseline exposure’.

Example: make treatment available for those who want it, reduce alcohol intake, etc.

⇒ target parameter of **Structural Mean Models**

Target Causal Parameters (3)

Complier causal effect:

Target: $E(Y_1 - Y_0 | X_1 > X_0)$.

Definition 'complier' ($X_1 > X_0$):

if 'set' $G = 0$ then $X = 0$ **and** if 'set' $G = 1$ then $X = 1$.

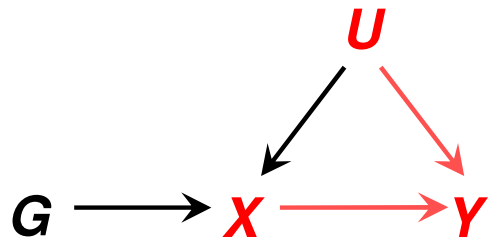
Example: randomised trial with partial compliance, etc.

(Note: more complicated when X continuous.)

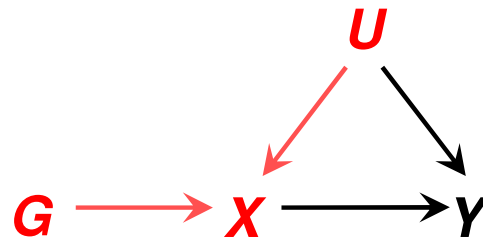
Typical Parametric Assumptions with IVs

Estimating a Causal Effect:

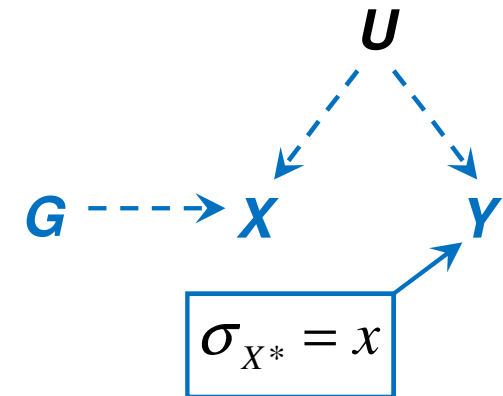
requires *some* additional parametric assumptions.



Outcome models
 $E(Y|X, U)$



Exposure models
 $E(X|G, U)$



Structural Mean models
 $E(Y|X, G; \sigma_{X^*} = x)$

$X \rightarrow$ natural exp.

$\tilde{X} \rightarrow$ manip. exp.

'Classical' IV Estimator

2–Stage–Least–Squares (2SLS): (or ratio / Wald estimator)

$$\hat{\beta}_{IV} = \frac{\hat{\beta}_{Y|G}}{\hat{\beta}_{X|G}}$$

- **simple;**
- only needs pairwise marginal data on (Y, G) and (X, G)
— could even come from separate data sets;
- generalises to multivariate X , Y and G :
predict \hat{X} from G , regress Y on \hat{X} ;
- **consistent for ACE** if...

Consistency of $\hat{\beta}_{IV}$

... linearity and additivity
in outcome model

$$E(Y|X = x, U = u) = \beta x + h(u)$$

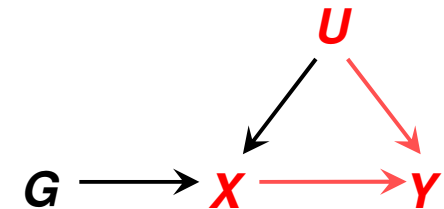
then $\hat{\beta}_{IV}$ consistent for $\beta = ACE$ for a unit difference in X .

(Unfortunately, there is small sample bias!)

Note: we only need an assumption for outcome model $E(Y|X, U)$.

No assumption about exposure model other than $Cov(X, G) \neq 0$.

But: implausible for binary Y and continuous X .



Something Similar for Binary Outcome Y ?

Suggestions for estimating CRR and COR ('Wald-type' estimators)

$$\hat{\gamma}_{IV.RR} = \frac{\log RR(Y|G)}{\hat{\beta}_{X|G}} \qquad \hat{\eta}_{IV.OR} = \frac{\log OR(Y|G)}{\hat{\beta}_{X|G}}$$

But: need additional assumption about **exposure model** $E(X|G, U)$.

Note: again, only needs pairwise marginal data on (Y, G) and (X, G) — could come from separate data sets.

And, $\log OR(Y|G)$ could come from case-control study.

Consistency of $\hat{\gamma}_{IV.RR}$

Assume outcome model

$$\log E(Y|X = x, U = u) = \gamma x + h(u)$$

$$\Rightarrow (\text{population}) \text{ } CRR = \exp(\gamma).$$

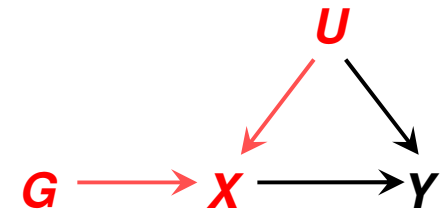
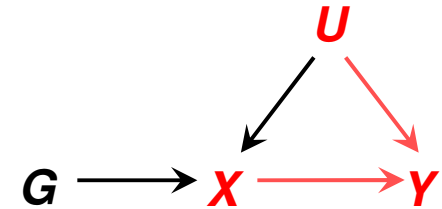
Assume also **exposure model**:

X normally* distributed, mean linear & additive

$$E(X|G = g, U = u) = \delta g + k(u).$$

$$\Rightarrow \hat{\gamma}_{IV.RR} \text{ consistent for } \gamma$$

$\hat{\eta}_{IV.OR}$ as approximation when $P(Y = 1)$ is small.



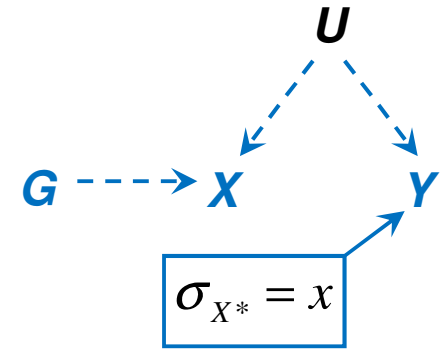
Comments on $\hat{\gamma}_{IV.RR}$ and $\hat{\eta}_{IV.OR}$

- Both are consistent under the null–hypothesis of no causal effect.
 - Both are sensitive to misspecification of **exposure model**.
 - When **no confounding** present, outcome model correct, but exposure model wrong, they will be **biased** (while naive regression of Y on X not biased).
 - Even **with confounding**, outcome model correct, but exposure model seriously wrong, they will be **more biased** than naive regression of Y on X .
 - The bias increases with size of true causal parameter.
- ⇒ ...better be confident of exposure model if you use these estimators!

Multiplicative Structural Mean Model (MSMM)

Assume

$$\frac{E(Y|X = x, G = g)}{E(Y|X = x, G = g; \sigma_{X^*} = x_0)} = \exp(\tilde{\gamma}(x - x_0))$$



⇒ exploit exclusion restriction to obtain estimating equations

⇒ solution $\hat{\gamma}_{MSMM}$ consistent for $\tilde{\gamma}$.

Note: various choices can make estimating equations more efficient.

Comments on MSMM

- MSMM needs **joint data** (X, Y, G) — not from separate data sets.
- If outcome model for $\hat{\gamma}_{IV.RR}$ valid, then MSMM satisfied and $\hat{\gamma}_{MSMM}$ consistent for *population CRR*.
 \Rightarrow if very different, then assumptions of $\hat{\gamma}_{IV.RR}$ possibly violated.
- MSMM estimator solves the same equations as **multiplicative Generalised Method of Moment (MGMM)** estimator — the latter is motivated by structural equations approach; see **ivpois** in Stata.

What about Logistic Models?

Might want to assume **logistic outcome model**

$$\text{logit } P(Y = 1|X = x, U = u) = \eta^* x + h(u)$$

Interpretation of η^* as **causal parameter difficult** (non-collapsibility).

Some IV methods target η^* (e.g. logistic control function approach); these also rely of specific parametric assumption of exposure model and distribution of U .

Or: **Logistic SMM**

$\text{logit } P(Y = 1|X = x, G = g)$ –

$$\text{logit } P(Y = 1|X = x, G = g; \sigma_{X^*} = x_0) = \tilde{\eta} (x - x_0)$$

⇒ estimating equations more complicated (again: non-collapsibility).

Application: Effect of BMI on Asthma in Children

Data:

Avon Longitudinal Study of Parents and Children (ALSPAC), $N = 4647$

Outcome: presence of asthma (binary) at age 7, 14.5% cases

Exposure: BMI at age 7, $\bar{x} = 16.22$, $s = 2.06$

Covariates: sex, maternal smoking (pre- and post-natal) etc.

Instrument: FTO with alleles T and A (risk allele), coded as 1=TT, 2=AT, 3=AA.

Note: average BMI only slightly higher for FTO=AA than others — this is a very weak instrumental variable!

Estimates of BMI Effect on Asthma

| Method | Target | Estimate | 95% CI |
|-----------------------------|---------------|----------|-------------|
| <i>Naive</i> | <i>RR</i> | 1.05 | 1.02 – 1.09 |
| <i>Naive</i> | <i>OR</i> | 1.06 | 1.02 – 1.11 |
| $\exp \hat{\gamma}_{IV.RR}$ | <i>CRR</i> | 1.37 | 0.68 – 2.78 |
| $\exp \hat{\gamma}_{MSMM}$ | \tilde{CRR} | 0.81 | 0.44 – 1.48 |
| $\exp \hat{\gamma}_{MGMM}$ | <i>CRR</i> | 0.81 | 0.44 – 1.48 |
| Logist.ControlF | <i>COR*</i> | 1.44 | 0.63 – 3.28 |
| Logist.SMM | \tilde{COR} | 1.64 | 0.29 – 9.31 |
| Probit (transf.) | <i>CRR*</i> | 1.36 | 0.75 – 1.97 |

Interactions / Effect Modification

All IV estimators assume

- either no effect modification by U in outcome model $E(Y|X, U)$ on relevant scale,
- or (weaker) no effect modification by G in SMM on relevant scale.

Note: in simulations, interaction in exposure model $E(X|G, U)$ leads to negative correlation of $\hat{\gamma}_{IV.RR}$ and $\hat{\gamma}_{MSMM}$

Conclusions

- Mendelian randomisation often provides plausible IV — but the defining assumptions of an IV are not automatically valid! still need to put careful thought into possible violations.
- Remember that testing null-hypothesis of no causal effect by checking $Y \perp\!\!\!\perp G$, is possible without specific parametric assumptions — don't even need to measure X !
- In view of the multitude of untestable assumptions (involving unobserved U), sensitivity analyses are recommended.
- Here, focus was entirely on consistency — maybe not the most important property of an estimator.