Recent Challenges for Mendelian Randomisation Analyses

Vanessa Didelez
Leibniz Institute for Prevention Research and Epidemiology
& Department of Mathematics
University of Bremen, Germany

CRM — Montreal, July 2016

Overview

- Motivation and introduction

  'Mendelian Randomisation': meiosis = natural 'coin-flip'

- Instrumental variable (IV): assumptions and inference

- Multiple IVs (e.g. found from GWAS)

  — all valid

  — unknown subset invalid

- Conclusions

Motivation

Examples of Mendelian Randomisation Analyses

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.)

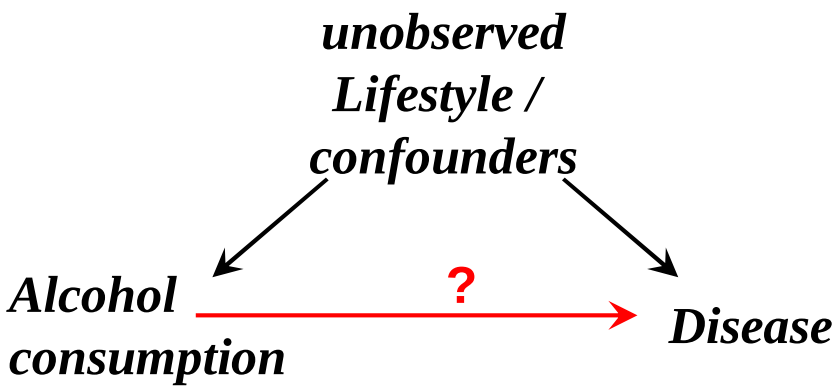
Observational studies are inevitable: preliminary research, but also assessment of effects in general population.

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

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: wildetype *1 and "null" variant *2.

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

$\Rightarrow$ *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!

Mendelian Randomisation: How to Find IVs?

MR: good idea, but still need to find suitable SNPs...

GWAS: genome-wide association studies, 'scan' for associations between SNPs and phenotypes / outcomes of interest, numerous often weak associations found.

Opportunities, but also lots of problems with using SNPs from GWASs as IVs...

GWAS and Multiple IVs

Example 1:
wanted: causal effect of height X on lung function Y,
180 SNPs associated with 'height', i.e. 180 (possible) IVs,
$n = 3631$, ALSPAC data, Davies et al. (2015).

Example 2:
wanted: causal effect of bp X on CAD Y,
29 SNPs associated with bp, i.e. 29 (possible) IVs,
$n = 86000$, *case-control data*, Ehret et al. (2011; data available online).

GWAS and Multiple IVs

Example 3:
wanted: causal effect of cholesterol X on CAD Y,
185 SNPs associated with cholesterol, i.e. 185 (possible) IVs,
2-samples, *log-odds-ratios*, Do et al. (2013).

Example 4:
wanted: causal effect of BMI X on schizophrenia Y,
97 SNPs associated with BMI, i.e. 97 (possible) IVs,
2-samples ($n_1 = 340000$, $n_2 = 9000$), *log-odds-ratios*, Burgess et al. (2016; data publicly available).

IV Assumptions

the basic case

Motivation for IV

Wanted: causal effect of exposure X on Y.

Problem: X not randomised; and strong suspicion of unobserved confounding by U.

Instrumental Variable: 'next best thing' to actual randomisation of $X \Rightarrow$ 'Nature's' randomisation, natural experiments, imperfect experiments, partial compliance, or similar.

Structural Assumptions: those that regard the interventional distribution $p(\cdot \,|\, \mathrm{do}(X = x))$.

Caution

Two types of IV situations:

(1) An experimenter really was in control of randomisation, but 'imperfect experiment', e.g. RCT with partial compliance.
$\Rightarrow$ know for sure that assignment was properly randomised.

(2) 'Natural experiments' etc. where (1) is not guaranteed, e.g. Mendelian Randomisation.

Instrumental Variables: both can yield valid IVs — but literature often assumes (1) in RCT context and hence makes stronger assumptions, even if not required.

Note: some IV methods require (1)!
e.g. all those assuming 'monotonicity' (see Sonja's talk?).

A Set of IV Assumptions

(I consider these as one of the weakest set of assumptions, though even weaker if stated in terms of lack of correlation)

G is an IV for effect of X on Y wrt. confounding by unobserved U under the following **associational** assumptions:

A1: $G \perp\!\!\!\perp U$

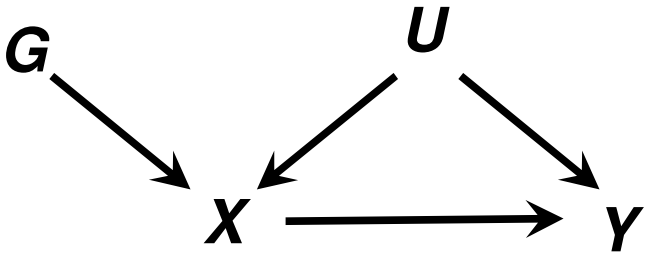
A2: $G \not\perp\!\!\!\perp X$ (2SLS: $Cov(X, G) \neq 0$)

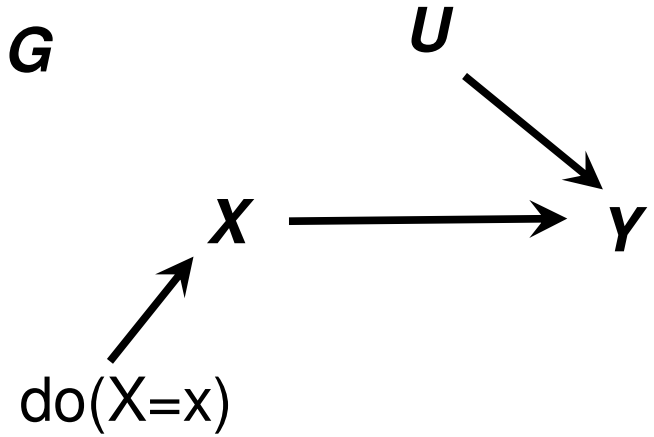
A3: $G \perp\!\!\!\perp Y \mid (X, U)$.

and the following **structural** assumptions:

$$P(Y|U, \mathsf{do}(X)) = P(Y|U, X),\ P(G|\mathsf{do}(X)) = P(G),\ P(U|\mathsf{do}(X)) = P(U).$$

Instrumental Variables with Cond. Indep. DAG





Corresponds to factorisation

$$p(y, x, u, g) = p(y|x, u)p(x|u, g)p(u)p(g)$$

and under intervention

$$p(y, u, g \,|\, \mathsf{do}(X = x')) = p(y|x, u)I\{x = x'\}p(u)p(g).$$

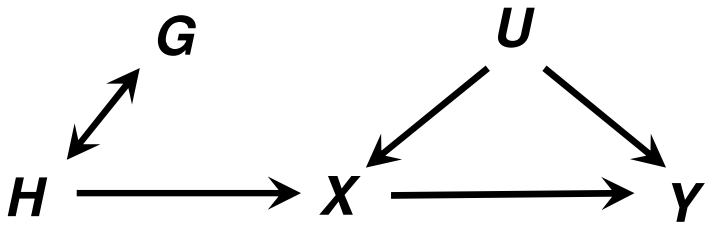
$$\Rightarrow Y \perp\!\!\!\perp G \,|\, \mathsf{do}(X = x) \qquad \approx \text{exclusion restriction.}$$

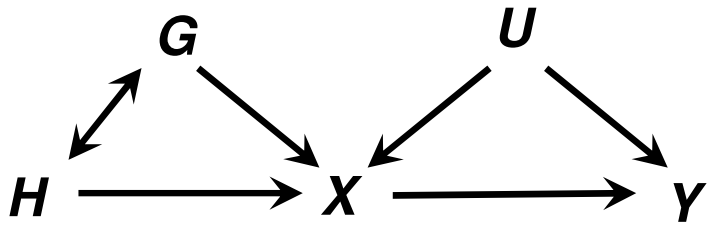
Note on IV Assumptions

Note: $(G\text{-}X)$–association allowed to be 'confounded',

e.g. G not the (only) causal gene but e.g. in linkage disequilibrium.

Let $G =$ IV and $H =$ hidden / unobserved.





In both scenarios, IV and structural assumptions on G, X, U, Y and wrt. do(X) still valid.

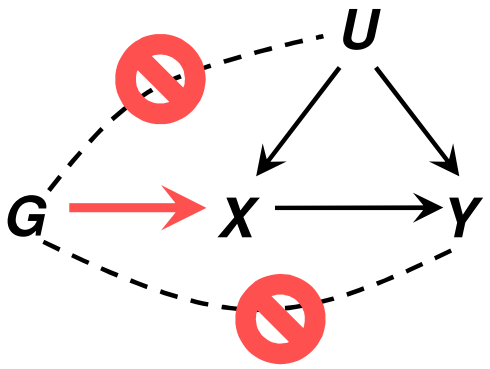
(Particular IV methods require 'causal' gene.)

'Untestable' Assumptions

The assumptions

A1: $G \perp\!\!\!\perp U$

A3: $G \perp\!\!\!\perp Y \mid (X, U)$.



impose inequality constraints when X, Y, G discrete;
otherwise no constraints. (cf. Stata package, Palmer et al., 2011)

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

$\Rightarrow$ Need to be justified based on subject matter background knowledge.

$\Rightarrow$ **Mendelian randomisation** & subject matter knowledge

Sources of Violations of IV Assumptions

Pleiotropy: G affects Y through another pathway than through X.

Linkage Disequilibrium: G associated with H and H affects Y through different pathway.

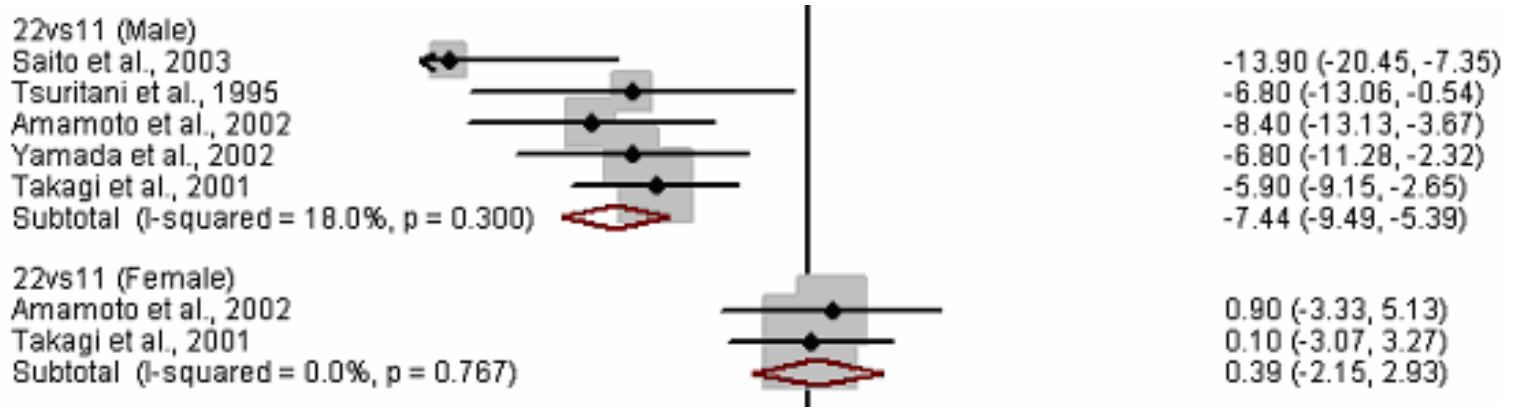
Population stratification: 'confounding' G and Y.

& others...

Example: Alcohol Consumption

Note: if $G \perp\!\!\!\perp X$ in subgroup, other assumptions valid, expect $G \perp\!\!\!\perp Y$!

Is condition $Y \perp\!\!\!\perp G | (X, U)$ satisfied? (Chen et al., 2008)



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.

Inference with IVs

some principles

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. (Null–Preservation)

Estimation:

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

(cf. `Stata` package, Palmer et al., 2011)

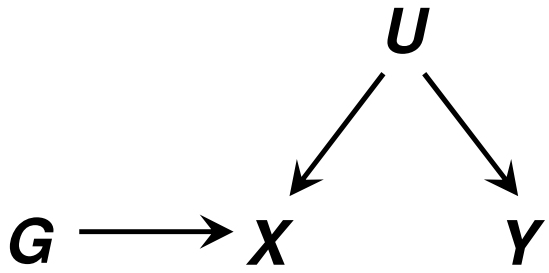
(2) for point estimates need some (semi–)parametric / structural assumptions, as well as clear definition of target causal parameter.

Using IV to Test for Causal Effect

simply test if

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

i.e. independence between instrument and outcome;



if find $Y \not\perp\!\!\!\perp G$ then evidence for causal effect (under IV assumptions)

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 H_0: if U acts as effect modifier in a very specific way.)

Two–Stage–Least–Squares (2SLS)

'Classical' IV estimator: (or ratio / Wald estimator)

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

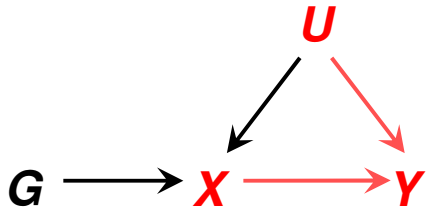
- **simple**;
- only needs pairwise marginal data on (Y, G) and (X, G)
  — could even come from separate data sets; useful for meta-analyses;
- generalises to multivariate X, Y and G:
  predict $\hat{X}$ from G, regress Y on $\hat{X}$;
- consistent for ACE if...

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.

(but weak IV 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 some measurement levels.

Weak Instrument Bias

Clearly: if G-X association 'weak' then G cannot contain much information about the causal effect $\Rightarrow$ unstable.

But also: 2SLS (& Co.) suffer from weak-IV / small-sample bias as $G \perp\!\!\!\perp U$ never exactly true in small samples:

when G-X association weak relative to the sample size, 2SLS IV–estimators based on joint data (G, X, Y) necessarily biased towards naïve (OLS) estimator. (Bound et al., 1995; Stock et al., 2002)

Weakness tested by R^2 or F–statistic in linear case.

Multiple Instruments

$G_1, \ldots, G_K$ candidate instruments

Problems with Multiple IVs

... especially when found from GWAS:

– weak G_k-X associations $\Rightarrow$ low power and potential bias;

$\Rightarrow$ want most efficient but robust methods;

– found by 'data-mining', so more likely some SNPs invalid IVs;

– often no joint data:
2-sample data: G-X and G-Y separately
only summary data on G-X and G-Y associations;

– GWAS often case-control data.

Multiple Independent IVs

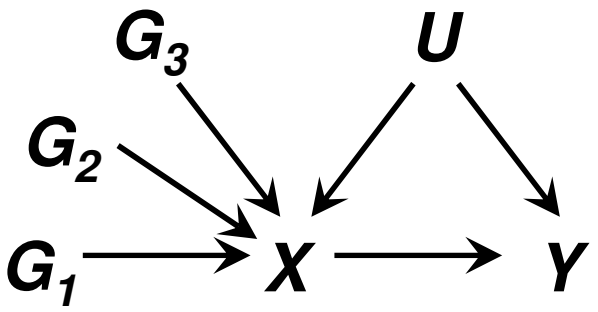
(Obvious generalisation...)

$G = \{G_k\}$ are IVs for effect of X on Y wrt. confounding by unobserved U under the same **associational** assumptions:

A1: $G_k \perp\!\!\!\perp U$

A2: $G_k \not\perp\!\!\!\perp X$

A3: $G_k \perp\!\!\!\perp Y \mid (X, U)$.



and the same **structural** assumptions as before.

Note: here also independent $G_k \Rightarrow$ each G_k individually valid,
e.g. different genetic regions; independence testable.

$\Rightarrow$ just use G as vector-valued IV, 2SLS still works, proceed as usual...?

Multiple Independent IVs

E.g. linear case.

Opportunity:

each SNP G_k yields a separate estimate $\hat{\beta}_k$ of $ACE = \beta$ (e.g. separate 2SLS estimators).

If all assumptions met, all $\hat{\beta}_k$ 'similar' (overall: 2SLS weighted average)

Hence:
large deviations indicate that 'something' is wrong, but not clear what.

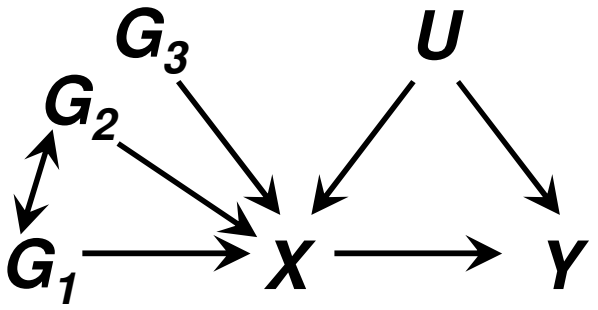
Note: more IVs than causal parameters $\Rightarrow$ 'over-identified'.

Also, $\approx$ Meta-analysis, $\hat{\beta}_k$ effect estimates from different studies.

Basis for 'over-identification' statistical test (Sargan-Hansen; Cochran's Q-statistic).

Multiple Dependent IVs

E.g. due to linkage disequilibrium.



Similar assumptions as before.

But A2: $G_k \not\perp\!\!\!\perp X \mid \{G_{-k}\}$

Spearate IV estimates $\hat{\beta}_k$ 'less independent'.

2SLS on individual data as before;

more generally, estimators require correlation–matrix for weighting — may not be available for summary data.

Weak Multiple IVs

Tempting: select exposure model $X \sim G_1, \ldots, G_K$ in data-driven way to maximise IV-strength. BUT: over-fitting exacerbates weak-IV bias!

(Didelez & Sheehan, 2011)

(To date) Recommendation: split-sample type model selection;

Popular: use of **allele scores**, i.e. predetermined weighted average of SNPs $\Rightarrow$ reduced # parameters in 1st stage, so less over-fitting.

(Burgess & Thompson, 2013)

Alternatively: improvements on 2SLS: adaptive weighting strategies like LIML and CUE.

Davies et al. (2015)

All above assume: all SNPs valid.

Some / All Instruments Invalid

... and we do not know which!

Possibly Relevant Questions

Question 1:

how are invalid IVs invalid?

do they violate A1: $G_k \perp\!\!\!\perp U$ or A3: $G_k \perp\!\!\!\perp Y | (X, U)$, or both?

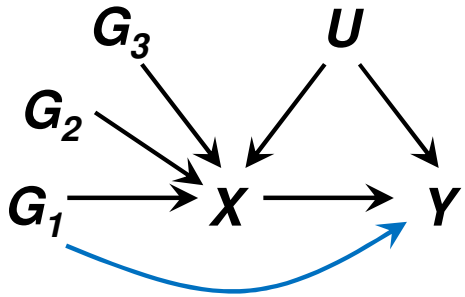
Question 2:

how are valid IVs valid? marginally or conditionally?

Violations of IV Assumptions

Pleiotropy:

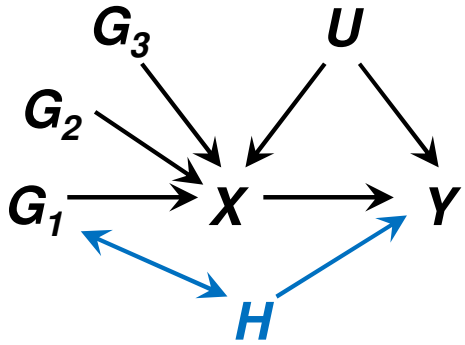
G affects Y through another pathway.



Linkage Disequilibrium:

G associated with H (hidden) and

H affects Y through different pathway.



Population stratification:

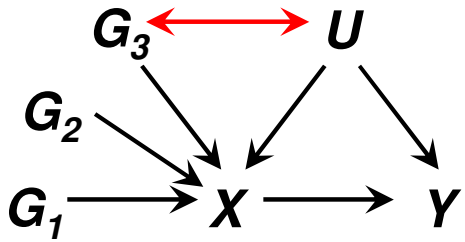
'confounding' between G and Y.

& others...

Note: $G \not\perp\!\!\!\perp U$ 'destroys' orthogonality.

$\Rightarrow$ methods can deal with partial violation of A3: $G \perp\!\!\!\perp Y | (X, U)$,

not so easy for violations of A1: $G \perp\!\!\!\perp U$.



Testability of IV Assumptions?

Remember: IV assumption only testable to very limited extend!

(Inequality constraints for multiple IVs?)

G_k valid IV, no (conditional) independencies among G_k, X, Y;

Ideas:

Multiple (independent) IVs: over-identification test.

Negative controls?

Margially / Individually Valid?

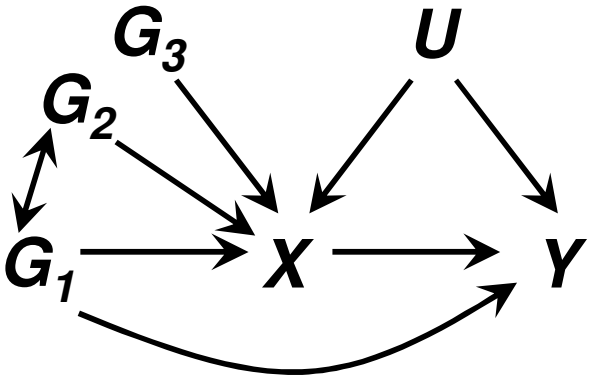
Here:

G_3 marginally valid IV

G_1 invalid (pleiotropy)

G_2 only valid conditional on G_1

$\Rightarrow$ analyse jointly



Note: relevant to methods that require $>$ half of IVs valid...

Some Invalid Multiple IVs

Assume:
some *unknown* subset of IVs is invalid, others are 'valid'.

Inspect set of IV estimates $\hat{\beta}_k$...

if we can assume more than half is (marginally & independently) valid, then non-parametrically identified (Kang et al., 2015)

Some Invalid IVs — Methods based on Combining $\hat{\beta}_k$

Combine $\hat{\beta}_k$'s, e.g. (Bowden & Burgess & Co, 2015, 2016)

- — median (half IVs valid)
- — weighted ($\sim$ strength) median (half weighted IVs valid)
- — other 'robust' weighted averages
- — 2-stage hard-thresholding *(Kang's talk)*.

Note 1: type of violation irrelevant, as long as $>$ half valid.
But when using IV strength as basis for weights, violation of A1: $G_k \perp\!\!\!\perp U$ biases the weights.
Also, conditional IVs not allowed.

Note 2: summary data sufficient; but need correlation if IVs dependent.

Note 3: asymptotics of some of these estimators unclear.

Some Invalid IVs — Methods based on Regression

Implicitly or explicitly selecting valid IVs:

— sisVIVE, Kang et al., 2015, (validity-IV.strength unclear?)

— selection of valid IVs,

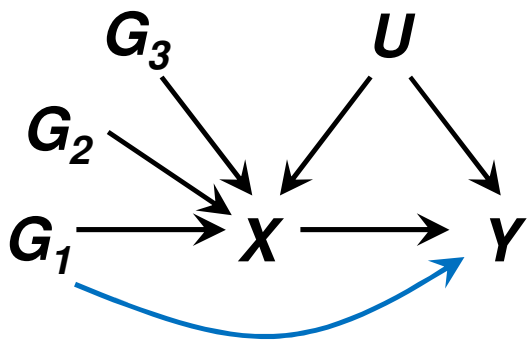
Han (2008); Kolesar et al. (2016); Windmeijer et al. (2015)

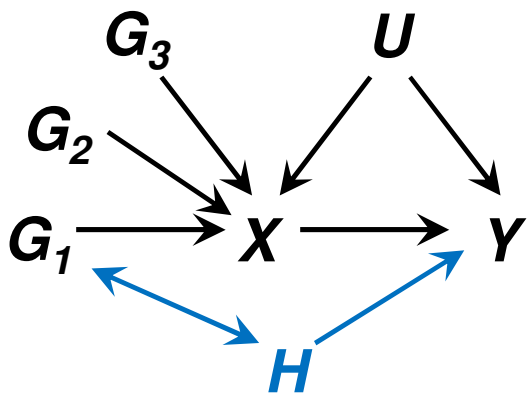
Note 1: A1: $G_k \perp\!\!\!\perp U$ must hold, A3 can be violated for some ($<$ half) IVs; dependent / conditional IVs allowed.

Note 2: require individual-level data.

Possibly all IVs Invalid

Assume: still A1: $\{G_k\} \perp\!\!\!\perp U$, but not A3: $G_k \perp\!\!\!\perp Y \,|\, (X, U)$ for some k.





Possible model

$E(Y|G_1, \ldots, G_K, X, U) = \sum_k \alpha_k G_k + \beta X + h(U)$

Define γ_k from linear projection '$X \sim G_k$' (exposure model needed?)

$\Rightarrow \hat{\beta}_k = \hat{\Gamma}_k / \hat{\gamma}_k$ asy. estimates $\beta + \alpha_k / \gamma_k$

Possibly all IVs Invalid

Still $E(Y|G_1,\ldots,G_K,X,U) = \sum_k \alpha_k G_k + \beta X + h(U)$, and γ_k from linear projection '$X \sim G_k$' so that $\hat{\beta}_k$ asy. estimates $\beta + \alpha_k/\gamma_k$, then

Idea: similar situation to Meta-analyses with small study bias, so...

'Egger' Regression claims: Bowden et al. (2015)

if $\alpha_k \perp \gamma_k$ can recover β even if **all** $\alpha_k \neq 0$

by 'regressing' $\hat{\Gamma}_k$ on $\hat{\gamma}_k$ allowing for *intercept* (= bias term).

Note:

some indeterminism due to choice of scale for G_k's... tbc!

Effect Heterogeneity

Remember: all estimation methods so far assumed

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

i.e. in particular U no effect modifier.

Cannot expect $\hat{\beta}_k$'s to be similar if this is not the case, even if all instruments valid.

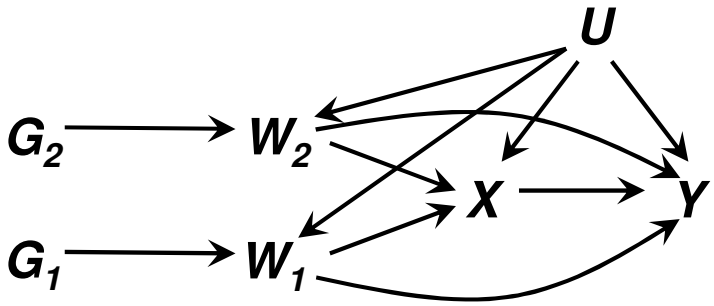
(See also discussion by Hernan & Robins, 2006)

Effect Heterogeneity?

... or violation of consistency?

"Different SNPs affect X in different ways"

hence, effect heterogeneity...?



DAG: different story!

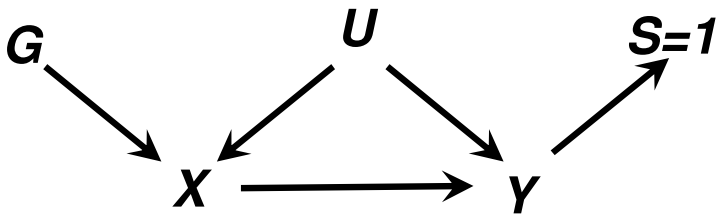
Note: $\tilde{U} = (W_1, W_2, U)$ one big unobserved confounder
$\Rightarrow G_1, G_2$ no valid instruments for X.

But: If such unobserved W_1, W_2 known to exist $\Rightarrow$ 'Null Preservation',
e.g. $G_1 \perp\!\!\!\perp Y$ supports no causal effect of W_1. (Celia's talk?)

IVs and Case-Control Designs

IVs under Selection

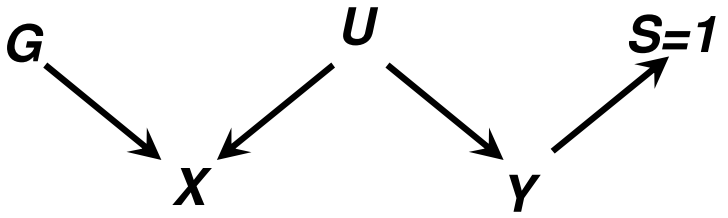
Case-Control Studies:



G violates A1
in *selected* sample: $G \not\perp\!\!\!\perp U | S = 1$.

$\Rightarrow$ typically require outside info on prevalence

Under Null-Hypothesis:



G still valid IV

$\Rightarrow$ can test $Y \perp\!\!\!\perp G \,|\, S = 1$

$\Rightarrow$ valid test of causal null.

(Didelez & Sheehan, 2007)

IV Estimation and Case-Control

Popular estimator: Wald-OR

$$\log \hat{COR}_{Wald} = \frac{\log \hat{OR}_{YG}}{\hat{\beta}_{X|G}}$$

$\hat{OR}_{YG}$ from CC-studies (often meta-analyses),

$\hat{\beta}_{X|G}$ from controls / separate studies.

Straightforward to **extend to multiple IVs using G_k separately**; apply same methods as for linear case...

Justified as approximation under strong assumptions, $X \sim$ Normal;

(Harbord et al., 2013)

Can be severely biased in many realistic situations (e.g. X binary)!

(Didelez et al., 2010)

IV Estimation and Case-Control (ctd.)

Alternative Methods:

Multiplicative-SMM only requires info on $E(G)$ (Bowden & Vansteelandt, 2011)

In MR studies: $E(G)$ simply population allele frequencies.

Logistic-SMM (Bowden & Vansteelandt, 2011) and Princ. Strata (Shinohara et al., 2012) require again some prevalence information for re-weighting.

But: these require individual-level data...

Problems with Other Selection Effects

Self-selection:, e.g. volunteers, maybe healthier etc.

if selection into sample somehow depends on exposure X
$\Rightarrow$ IVs invalidated.

(Summary data:) **reported** G-X **or** G-Y **associations:** may have been adjusted for 'heritable covariates' / post-IV covariates, e.g. WHR and BMI, e.g. to obtain 'direct' effect — can give rise to more substantial collider bias than one might think. (Aschard et al., 2015)

Conclusions

- Would really like a way to verify that a variable is a valid IV...
- ...in the meantime:
  all available methods allowing for some unknown invalid IVs have their pro's and con's — is there anything better we can do than to apply them all and hope they agree?
- Subtle differences in assumptions not always clear...
  'Taxonomy' based on (i) type of violation allowed, (ii) type of validity needed, (iii) # valid IVs required, (iv) further parametric assumptions, (v) types of data required seems useful.

Literature

Aschard, Vilhjalmsson, Joshi, Price, Kraft (2015). Adjusting for Heritable Covariates Can Bias Effect Estimates in Genome-Wide Association Studies. The American Journal of Human Genetics, 96, 329339.

Bound, Jaeger, Baker (1995). Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak. JASA, 90(430), 443-450.

Bowden, Davey Smith, Burgess (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int. J. Epidemiol., 44(2), 512-525.

Bowden, Davey Smith, Haycock, Burgess (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genetic Epidemiology, 40(4), 304314.

Burgess, Thompson (2013). Use of allele scores as instrumental variables for Mendelian randomization. International Journal of Epidemiology, 42(4), 1134-1144.

Bowden, Vansteelandt (2011). Mendelian randomization analysis of case-control data using structural mean models. Statistics in Medicine, 30(6), 678-694.

Burgess, Bowden, Dudbridge, Thompson (2016). Robust instrumental variable methods using multiple candidate instruments with application to Mendelian randomization. arXiv.

Burgess, Bowden, Fall, Ingelsson, Thompson (2016). Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. Epidemiology.

Chen, Davey Smith, Harbord, Lewis (2008). Alcohol Intake and Blood Pressure: A Systematic Review Implementing a Mendelian Randomization Approach. PLoS Med 5(3).

Davies, Von Hinke Kessler Scholder, Farbmacher, Burgess, Windmeijer, Davey Smith, (2015). The many weak instruments problem and Mendelian randomization. Statistics in Medicine, 34(3), 454-468.

Didelez, Meng, Sheehan (2010). Assumptions of IV Methods for Observational Epidemiology. Statistical Science 25, 22-40.

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

Didelez, Sheehan (2011). Commentary: Can many weak instruments ever be strong? International Journal of Epidemiology, 40(3), 752-754.

Do, Willer, Schmidt, Sengupta et al. (2013). Common variants associated with plasma triglycerides and risk for coronary artery disease. Nature Genetetics, 45(11), 1345-52.

Ehret, Munroe, Rice, Bochud et al. (2011) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature, 478(7367), 103-9.

Guo, Kang, Cai, Small, (2016). Confidence Interval for Causal Effects with Possibly Invalid Instruments Even After Controlling for Many Confounders. Technical Report.

Han, (2008). Detecting Invalid Instruments using L1-GMM. Economics Letters, 101, 285-287.

Harbord, Sterne, Didelez, Palmer, Lawlor, Meng, Davey Smith, Sheehan (2013). Severity of bias of a simple estimator of the causal odds ratio in Mendelian randomisation studies. Statistics in Medicine. 32(7), 1246-1258.

Hernán, Robins (2006). Instruments for causal inference: an epidemiologist's dream? Epidemiology, 17(4), 360-372.

Kang, Zhang, Cai, Small, (2016). Instrumental Variables Estimation with Some Invalid Instruments and its Application to Mendelian Randomization. JASA, 111, 132-144.

Kolesar, Chetty, Friedman, Glaeser, Imbens, (2016). Identification and Inference with Many Invalid Instruments. Journal of Business and Economic Statistics, in press.

Palmer, Ramsahai, Sheehan, Didelez (2011). Nonparametric bounds for the causal effect in a binary instrumental variable model. The Stata Journal, 11(3), 345-367.

Shinohara, Frangakis, Platz, Tsilidis (2012). Designs combining instrumental variables with case-control: estimating principal strata causal effects. The International Journal of Biostatistics, 8(1), 1-21.

Stock, Wright, Yogo (2002). A survey of weak instruments and weak identification in generalized method of moments. Journal of Business & Economic Statistics, 20(4), 518-529.

Windmeijer, Farbmacher, Davies, Davey Smith, White (2016). Selecting (In)Valid Instruments for Instrumental Variables Estimation. Technical Report.