**Twin-based direction of causation (DOC) modeling**

Nathan Gillespie

Virginia Institute for Psychiatric and Behavior Genetics (VIPBG)

Virginia Commonwealth University

nathan.gillespie@vcuhealth.org

When making a causal inference e.g. the impact of cannabis (CAN) use on SZ, the **primary difficulty** is to determine the degree to which the CAN-SZ association stems from either known or unknown confounding that influences the both the RISK and an OUTCOME.

**Gold Standard** = randomized clinical trial (RCT):

1. The response of experimental participants 'assigned' to exposure is compared with responses of participants in a non-exposed control group
2. The 'assignment' of participants to exposure and control groups is 'random'
3. The manipulation of the exposure is controlled by the researcher

Application of RCTs to psychiatric outcomes is invariably impractical and unethical.

Need **alternative** means of inferring causality

Natural Experiments:

1. **Pre-Post Designs**
2. **Co-relative designs**

Compares the decline in exposure-outcome associations with increasing control of genetic and shared environmental influences by using relatives with varying degrees of genetic relatedness e.g. pairs of first cousins, sib pairs including DZ and MZ twin pairs

- Nicotine use causes major depression (Kendler, 1993)

- CSA causes adult psychiatric disorders (Kendler, 2000)

- Alcohol abuse-dependence causes lifetime adult major depression (Prescott, 2000)

- Cannabis use causes psychosis (McGrath, 2010)

In co-relative designs, if risk factor-outcome association is:

**Causal**

- Controlling for background & genetic effects makes no difference

- Estimates (ORs) are same

**Partly due to G factors influencing risk & outcome**

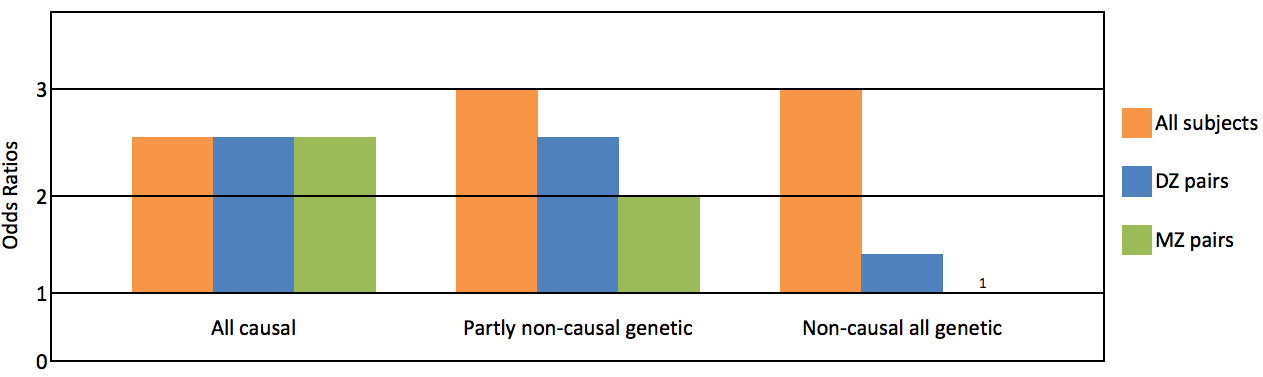
- Association strongest in entire sample (no control),

- Intermediate for DZs (full E, part G control)

- Lowest for MZ (full G & E control)

**Entirely genetic**

* MZ ORs approach = 1, DZ ORs are midway



1. **Twin-based SEM DOC modeling**

- simple bivariate DOC between measured phenotypes

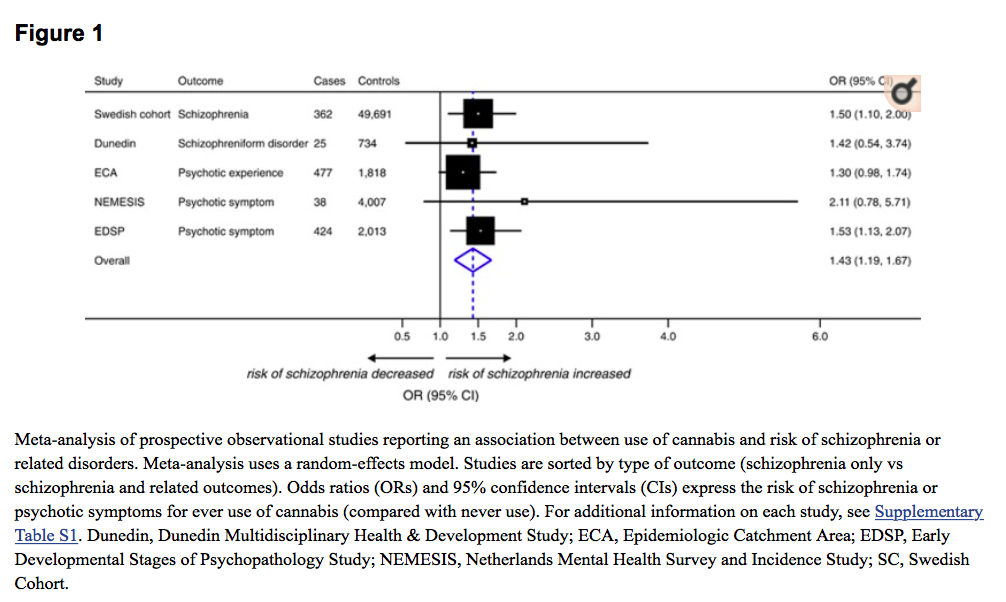
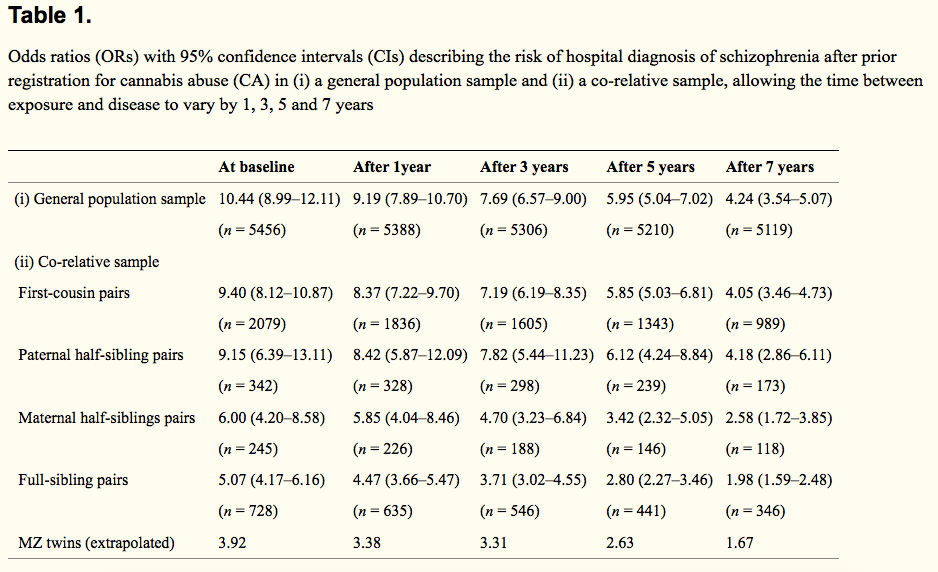
- causal-common-contingency (CCC) modeling

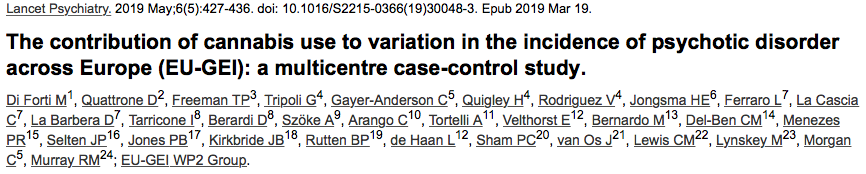
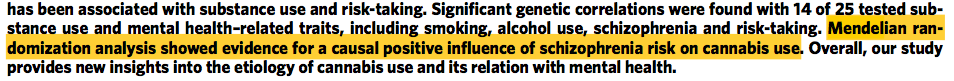
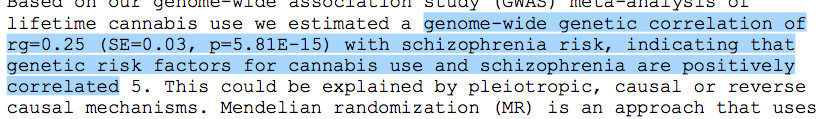
- multiple indicator DOC modeling between latent phenotypes

- longitudinal DOC modeling & cross-panel designs

1. **Instrumental Variable Analyses**

- Mendelian Randomization

**Warning:** All causal inferences invariably rely on methodological assumptions that rarely produce unequivocal results e.g. **cannabis** & **schizophrenia**

****



Solution? Triangulation

The confidence in a causal assumption should be based not on a single but multiple divergent methods each differing in their theoretical assumptions. This ‘**triangulation’** approach (Munafo, 2018) ought to provide stronger converging lines of evidence compared to any well-replicated single approach that ignores unacknowledged biases.

**Twin-based SEM DOC modeling**

Best natural experimental means of modeling DOC while controlling sources of genetic & environmental confounding



**Twin-based SEM DOC modeling**

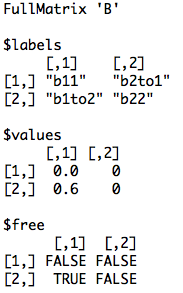
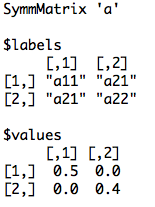
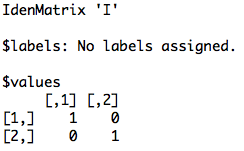
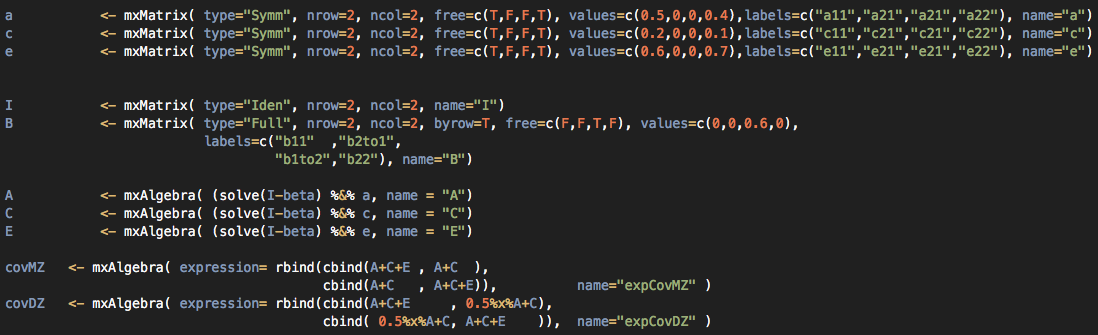
The ‘magic’ or power of twin-based SEM DOC modeling is based on the information from the cross-twin cross-trait covariance

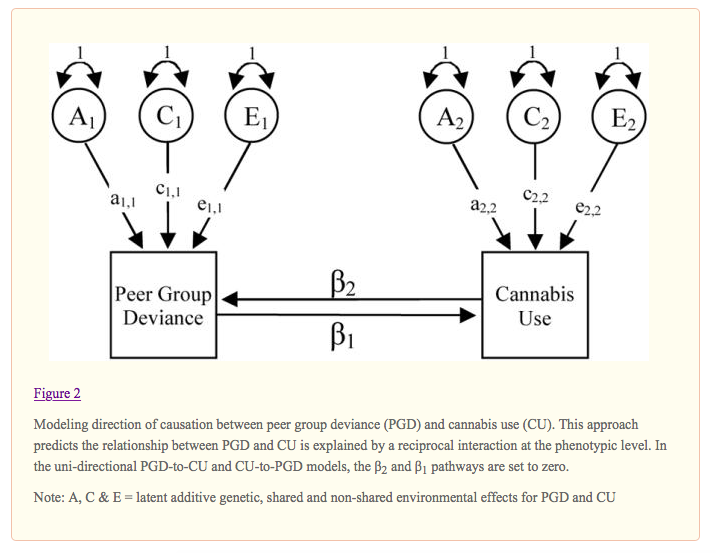
****

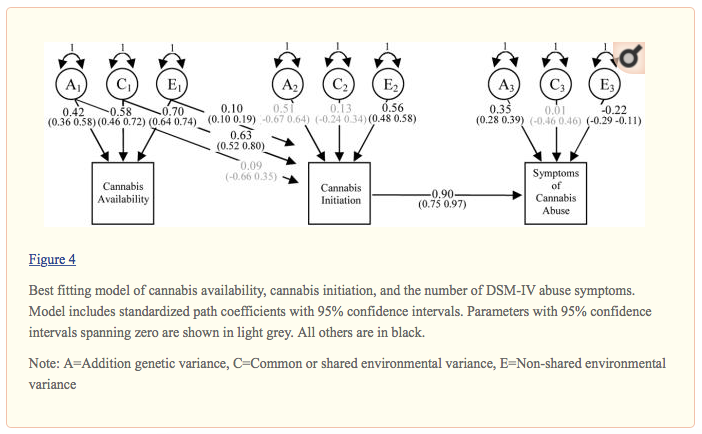
1. Members of a twin pair are not having any mutual effect on one another, i.e., sibling cooperation/rivalry, either within or across variables
2. Relationship between the two target variables is equivalent for twin 1 and twin 2
3. **Twin pair correlations are different between target variables**
4. There are no unmeasured variables influencing both measures and thereby inflate the correlations arising through the causal influence of one variable on the other

**Twin-based SEM DOC modeling**

Modeling causality at phenotypic level







**DOC modeling**

Problems with modeling DOC at the phenotypic level

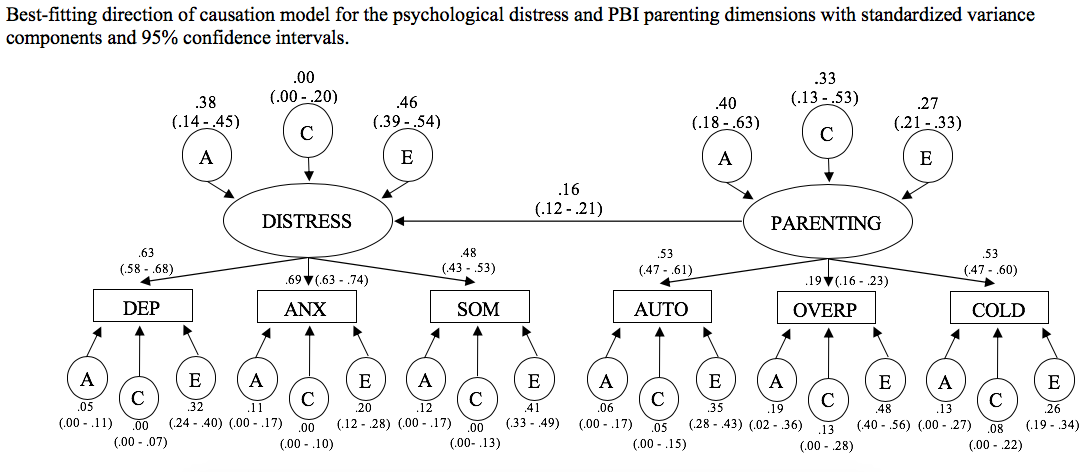
Omission measurement error known to produce biased estimates of the causal parameters (Neale & Cardon, 1992)

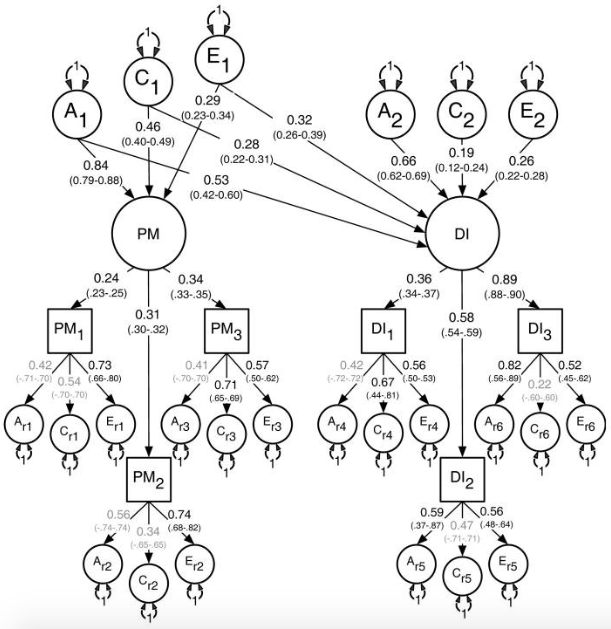


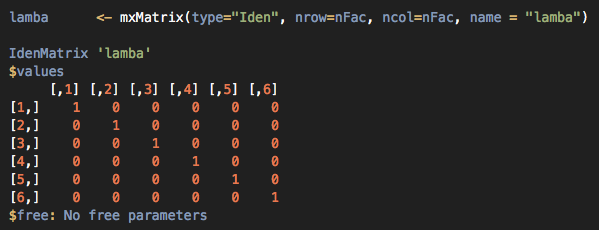


**DOC modeling**

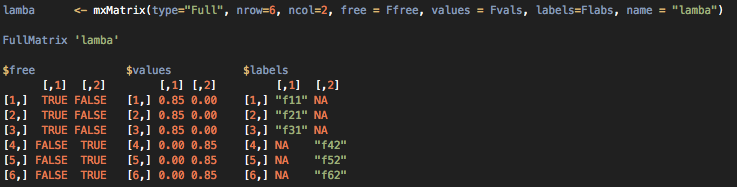
Modeling DOC at the latent factor level

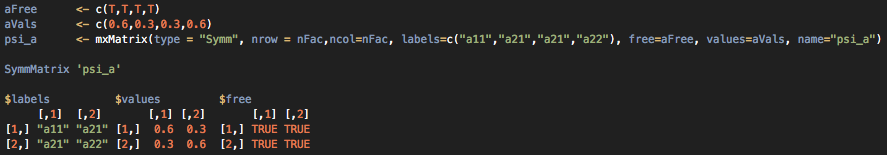


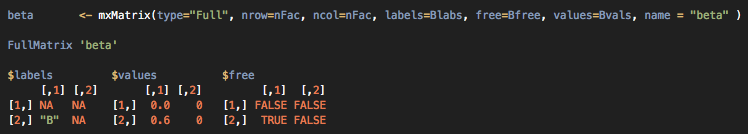
****

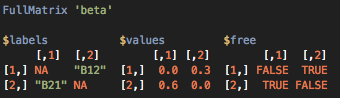
**Specifying the model**

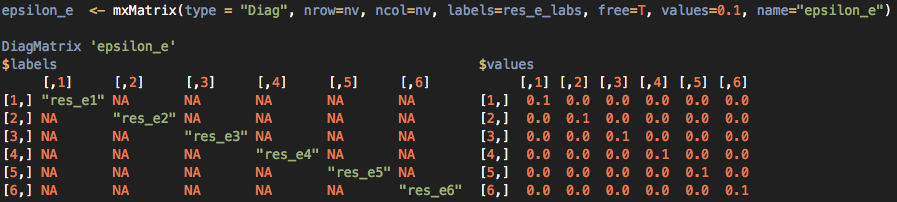
 

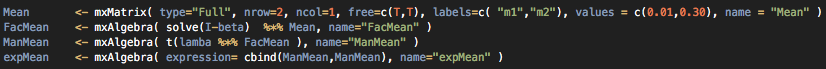
 









Varieties of causal and non-causal models

Hybrid model



**Tutorial**

1. Run ACE variance components (non-causal) model = null hypothesis
2. Specify / test AE variance components (non-causal) model.
3. Specify / test CE variance components (non-causal) model
4. Compare ACE, AE & CE variance components (non-causal) models

Extra credit: Re-run ACE model to estimate Cis on the A,C,E parameters. Which are non-significant?

1. Run the ‘CP1 to CP2’ causal model (no correlated A, C or E effects)
2. Run the ‘CP2 to CP1’ causal model (no correlated A, C or E effects)
3. Run the ‘CP1 to CP2’ & ‘CP2 to CP1’ causal model (no correlated A, C or E effects) i.e. reciprocal interaction.
4. Compare the 3 causal models to the null hypothesis



