Mendelian Randomization
(Using genes to tell us about the environment)

David Evans
University of Queensland
University of Bristol
Some Criticisms of DOC Modelling in Twins

• Measurement error

• Power (both variables need to have radically different aetiologies)

• Only useful for testing strong hypotheses about causation

• We now have genotypes for use in Mendelian randomization studies
This Session

• The problem with observational studies

• What is Mendelian Randomization (MR)?

• Examples of MR in research
  – Opportunities for SEM (mediation/network models)
  – Pharmacogenomics

• Using R to perform MR
RCTs are the Gold Standard in Inferring Causality

- **Randomised Controlled Trial**
  - **Randomisation Method**
    - Exposed: Intervention
    - Control: No Intervention
      - Confounders equal between groups
        - Outcomes compared between groups
Observational Studies

• RCTs are expensive and not always ethical or practically feasible

• Association between environmental exposures and disease can be assessed by observational epidemiological studies like case-control studies or cohort studies

• The interpretation of these studies in terms of causality is problematic
CHD risk according to duration of current Vitamin E supplement use compared to no use

Rimm et al NEJM 1993; 328: 1450-6
May 20, 1993

Vitamin E Greatly Reduces Risk Of Heart Disease, Studies Suggest

By JANE E. BRODY

Two new studies of more than 120,000 men and women strongly suggest that supplements of vitamin E can significantly reduce the risk of disease. Researchers and other experts cautioned against rushing out to buy the vitamin supplements before further clinical trials confirm that they are beneficial.

The studies, by researchers at the Harvard School of Public Health and Brigham and Women's Hospital in Boston, showed that initially healthy people were at a rate about 40 percent lower than comparable men and women whose intake of this vitamin was lowest. The preventive was blood levels of cholesterol.

The greatest protection was found at levels of about 100 international units of vitamin E a day for more than two years. The Federal recommender consume fewer than 25 units from foods like vegetable oils, wheat germ, seeds, whole grains and nuts.

The researchers said vitamin E, as an antioxidant, might reduce heart disease by having an effect on low-density lipoprotein cholesterol, or LDL type of cholesterol damages arteries primarily after it has been oxidized.

The new findings, which appear today in The New England Journal of Medicine, are some of the first to find health benefits from taking large- "megadoses" of vitamins as a popular remedy whose value is unproven. Expert Urge Caution

While a person might conclude from the findings that it would be wise to take large doses of vitamin E supplements daily, their long-term safety has not been established. Future trials are planned.
The average American lifespan has increased nearly 3 years over the last 2 decades.*

We’ve been selling vitamins at a discount since 1977.

Coincidence? We don’t think so.

At VitaminShoppe.com we see vitamins as an essential part of a healthy life – not a luxury. And our pricing reflects that philosophy. Right now we are taking 40% off every item we stock. After 23 years in the vitamin business, we’ve learned how to assemble the finest vitamins, minerals, and supplements at the lowest prices...all 18,000 of them.
Use of vitamin supplements by US adults, 1987-2000

Vitamin E levels and risk factors: Women’s Heart Health Study

Childhood SES ↓
Manual social class ↓
No car access ↓
State pension only ↓
Smoker ↓
Daily alcohol ↑
Exercise ↑
Low fat diet ↑
Obese ↓
Height ↑
Leg length ↑

Lawlor et al, Lancet 2004
Vitamin E supplement use and risk of Coronary Heart Disease

Stampfer et al NEJM 1993; 328: 144-9; Rimm et al NEJM 1993; 328: 1450-6; Eidelman et al Arch Intern Med 2004; 164:1552-6
"Well, so much for antioxidants."
Classic limitations to “observational” science

- Confounding
- Reverse Causation
- Bias
An Alternative to RCTs: Mendelian randomization

In genetic association studies the laws of Mendelian genetics imply that comparison of groups of individuals defined by genotype should only differ with respect to the locus under study (and closely related loci in linkage disequilibrium with the locus under study).

Genotypes can proxy for some modifiable risk factors, and there should be no confounding of genotype by behavioural, socioeconomic or physiological factors (excepting those influenced by alleles at closely proximate loci or due to population stratification).

Mendel in 1862
Fisher and Confounding

“Generally speaking the geneticist, even if he foolishly wanted to, could not introduce systematic errors into comparison of genotypes, because for most of the relevant time he has not yet recognized them”

Fisher (1952)
Mendelian randomisation and RCTs

Mendelian Randomisation

Random Segregation of Alleles

Exposed: Functional Alleles

Control: Null Alleles

Confounders Equal Between Groups

Outcomes Compared Between Groups

Randomised Controlled Trial

Randomisation Method

Exposed: Intervention

Control: No Intervention

Confounders Equal Between Groups

Outcomes Compared Between Groups
Mendelian Randomization- Core Assumptions

(1) SNP is associated with the exposure

(2) SNP is not associated with confounding variables

(3) SNP only associated with outcome through the exposure
Equivalence Between Directed Acyclic Graphs and SEMs

SNP → Exposure → Outcome

Confounders

\[ \xi_1 \quad \xi_2 \]

SNP → Exposure → Outcome

\[ \beta_x \quad \beta_Y \]
Why Perform MR?

• Assess causal relationship between two variables

• Estimate magnitude of causal effect
Calculating Causal Effect Estimates

SNP → Exposure → Outcome

\[ \beta_{\text{SNP-OUTCOME}} \]

\[ \beta_{\text{SNP-EXPOSURE}} \]

\[ \beta_{\text{EXP-OUTCOME}} \]

2SLS:

1. Regress exposure on SNP
2. Regress outcome on predicted exposure from 1st stage regression
3. Adjust standard errors

Wald Test:

\[ \frac{\beta_{\text{SNP-OUTCOME}}}{\beta_{\text{SNP-EXPOSURE}}} \]

*Can be used in different samples (“Two sample MR”)
Examples – *using* instruments for adiposity

If adiposity DOES NOT causally affect metabolic traits, then the FTO variant should NOT be related to these metabolic traits

If adiposity causally affects metabolic traits, then the FTO variant should also be related to these metabolic traits to the extent to which it affects adiposity
Do intermediate metabolic traits differ as one would expect given a \textit{FTO}-BMI effect?

Given the per allele \textit{FTO} effect of \(~0.1\)SD and known observational estimates one can derive an expected, per allele, effect on metabolic traits.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Expected Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting insulin</td>
<td>0.038 (0.033, 0.043)</td>
</tr>
<tr>
<td>Fasting Glucose</td>
<td>0.018 (0.014, 0.021)</td>
</tr>
<tr>
<td>Fasting HDL</td>
<td>-0.026 (-0.029, -0.023)</td>
</tr>
<tr>
<td></td>
<td>N~12,000 samples of European ancestry</td>
</tr>
</tbody>
</table>
Bidirectional MR
CRP and BMI

• C-Reactive Protein (CRP) is a biomarker of inflammation

• It is associated with BMI, metabolic syndrome, CHD and a number of other diseases

• It is unclear whether these observational relationships are causal or due to confounding or reverse causality

• This question is important from the perspective of drug development
“Bi-directional Mendelian Randomization”
<table>
<thead>
<tr>
<th>Outcome/explanatory variable</th>
<th>Observational</th>
<th>Instrumental variable</th>
<th>$P_{IV}$</th>
<th>$P_{diff}$</th>
<th>$F_{first}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRP/BMI</td>
<td>1.075</td>
<td>1.06</td>
<td>0.002</td>
<td>0.6</td>
<td>50.2</td>
</tr>
<tr>
<td></td>
<td>(1.073, 1.077)</td>
<td>(1.02, 1.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Using MR to Inform Mother-Offspring Associations
Light drinking in pregnancy may be good for baby boys, says study
Researchers find fewer behavioural problems and higher test scores at age 3

Sarah Boseley, health editor
The Guardian, Friday 31 October 2008
Article history

Light drinking during pregnancy has no ill effects and may have benefits for boys
Mom's light drinking doesn't harm baby

REUTERS, Oct 6, 2010, 10.15am IST

**Pregnant drinkers**: drinkaware.co.uk - Want to know how alcohol may be affecting your baby? Find out today

**Tags**: women, pregnancy, drinking, child, alcohol

*Women who have one or two alcoholic drinks a week during pregnancy do not harm their children's behavioural or intellectual development, according to a study by British scientists.*

The researchers found that pregnant women who drank up to a glass (175 millilitres) of wine, up to 50 ml of spirits or just under a pint of beer a week did not affect their children. But children whose mothers were heavy drinkers were more likely to be hyperactive and have behavioural and emotional problems than those whose mothers did not drink during pregnancy, the scientists said.
Total difficulties in top 10% of scores by mother’s drinking category

% of mothers who smoked during pregnancy by drinking category

% of Mothers who never worked, long-term unemployed etc, by drinking category

### Table 2. Results for adjusted model including 4 child variants.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Maternal drinking during pregnancy</th>
<th></th>
<th></th>
<th>Non-drinkers N = 1375</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&lt;1–6 units per week N = 2792</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Per allele effect on WISC score &amp; 95% confidence intervals</td>
<td>P-value</td>
<td>Per allele effect on WISC score &amp; 95% confidence intervals</td>
<td>P-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADH1A</td>
<td>rs2866151</td>
<td>-1.95 (-3.29 to -0.61)</td>
<td>0.004</td>
<td>-0.38 (-2.47 to 1.71)</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADH1A</td>
<td>rs975833</td>
<td>-1.72 (-3.23 to -0.21)</td>
<td>0.03</td>
<td>-0.66 (-2.90 to 1.59)</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADH1B</td>
<td>rs4147536</td>
<td>-1.47 (-2.97 to 0.02)</td>
<td>0.05</td>
<td>-0.71 (-2.92 to 1.50)</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADH7</td>
<td>rs284779</td>
<td>-1.27 (-2.10 to -0.44)</td>
<td>0.003</td>
<td>-0.11 (-1.12 to 1.35)</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Alcohol dehydrogenase (ADH) risk allele score in offspring and offspring IQ, stratified by maternal alcohol intake during pregnancy

P value for interaction of risk allele score and drinking during pregnancy equals 0.009

Any wine and kid's a plonker

Wine worry ... a mum-to-be (posed by model)

MUMS WARNED

By EMMA LITTLE, Health and Science Editor

MUMS-to-be who drink just ONE GLASS of wine give birth to kids with a lower IQ, researchers have claimed.

A study found any amount of alcohol during pregnancy can hit a baby's developing brain.

Doctors at Oxford and Bristol universities tracked
Multiple Variants, Multivariable MR, Two-Step MR and Network Models
Using Multiple Genetic Variants as Instruments

- Allelic scores

![Diagram showing Mendelian randomisation analysis using genetic variants as instrumental variables for fat mass on bone mineral density](image)

Figure 1. DAG for a Mendelian randomisation analysis using four genetic variants as instrumental variables for the effect of fat mass on bone mineral density.


- Testing multiple variants individually
Two Step Mendelian Randomization (Mediation Analysis)

- How much of the effect of BMI on CHD is mediated through LDL?
- Can be done using two sample MR
- Model can be fitted using SEM
Multivariable MR

SNP1
SNP2
SNP3
SNP4
SNP5

Fat Mass
Lean Mass
Bone Mineral Density (Osteoporosis)
Multivariable MR

SNP1
SNP2
SNP3
SNP4
SNP5

Fat Mass

Lean Mass

Bone Mineral Density (Osteoporosis)

$\xi_1$

$\xi_2$

$\xi_3$
Multivariable MR

SNP1
SNP2
SNP3
SNP4
SNP5

Fat Mass

Lean Mass

Bone Mineral Density (Osteoporosis)*

*Difficult to fit in 2SLS framework
Multivariable MR

*This model is not identified
Mendelian Randomization and Drug Targets
Late Stage Failure in Drug Trials

Cook et al. (2014) *Nat Rev Drug Disc*
Association of LDL-C, HDL-C, and risk for coronary heart disease (CHD)

302K participants in 68 prospective studies

Emerging Risk Factors Collaboration, JAMA 2009
LDL and CHD Risk

Ference et al, JACC 2012
HDL: endothelial lipase Asn396Ser

Loss-of-function variants in endothelial lipase are a cause of elevated HDL cholesterol in humans

Andrew C. Edmondson, Robert J. Brown, Sekar Kathiresan, L. Adrienne Cupples, Serkalem Demissie, Alisa Knodle Manning, Majken K. Jensen, Eric B. Rimm,
Jian Wang, Amrith Rodrigues, Vaneeta Bamba, Sumeet A. Khetarpal, Megan L. Wolfe, Stephanie DerOhannessian, Mingyao Li, Muredach P. Reilly, Jens Aberle,
David Evans, Robert A. Hegele, and Daniel J. Rader

- 2.6% of population carry Serine allele
- higher HDL-C
- No effect on other lipid fractions
- No effect on other MI risk factors

Edmondson, J Clin Invest 2009
LIPG N396S and plasma HDL-C

396S carriers have 5.5 mg/dl higher HDL-C
P<10^{-8}
After testing in 116,320 people, summary OR for LIPG Asn396Ser is 0.99
Individuals who carry the HDL-boosting variant have the same risk for heart attack as those who do not carry the variant.
Effects of Torcetrapib in Patients at High Risk for Coronary Events

Philip J. Barter, M.D., Ph.D., Mark Caulfield, M.D., M.B., B.S., Mats Eriksson, M.D., Ph.D., Scott M. Grundy, M.D., Ph.D., John J.P. Kastelein, M.D., Ph.D., Michel Komajda, M.D., Jose Lopez-Sendon, M.D., Ph.D., Lori Mosca, M.D., M.P.H., Ph.D., Jean-Claude Tardif, M.D., David D. Waters, M.D., Charles L. Shear, Dr.P.H., James H. Revkin, M.D., Kevin A. Buhr, Ph.D., Marian R. Fisher, Ph.D., Alan R. Tall, M.B., B.S., and Bryan Brewer, M.D., Ph.D., for the ILLUMINATE Investigators*

RESULTS

At 12 months in patients who received torcetrapib, there was an increase of 72.1% in high-density lipoprotein cholesterol and a decrease of 24.9% in low-density lipoprotein cholesterol, as compared with baseline (P<0.001 for both comparisons), in addition to an increase of 5.4 mm Hg in systolic blood pressure, a decrease in serum potassium, and increases in serum sodium, bicarbonate, and aldosterone (P<0.001 for all comparisons). There was also an increased risk of cardiovascular events (hazard ratio, 1.25; 95% confidence interval [CI], 1.09 to 1.44; P=0.001) and death from any cause (hazard ratio, 1.58; 95% CI, 1.14 to 2.19; P=0.006). Post hoc analyses showed an increased risk of death in patients treated with torcetrapib whose reduction in potassium or increase in bicarbonate was greater than the median change.
Effects of Dalcetrapib in Patients with a Recent Acute Coronary Syndrome

Gregory G. Schwartz, M.D., Ph.D., Anders G. Olsson, M.D., Ph.D., Markus Abt, Ph.D., Christie M. Ballantyne, M.D., Philip J. Barter, M.D., Ph.D., Jochen Brumm, Ph.D., Bernard R. Chaitman, M.D., Ingar M. Holme, Ph.D., David Kallend, M.B., B.S., Lawrence A. Leiter, M.D., Eran Leitersdorf, M.D., John J.V. McMurray, M.D., Hardi Mundl, M.D., Stephen J. Nicholls, M.B., B.S., Ph.D., Prediman K. Shah, M.D., Jean-Claude Tardif, M.D., and R. Scott Wright, M.D.,

for the dal-OUTCOMES Investigators*

RESULTS

At the time of randomization, the mean HDL cholesterol level was 42 mg per deciliter (1.1 mmol per liter), and the mean low-density lipoprotein (LDL) cholesterol level was 76 mg per deciliter (2.0 mmol per liter). Over the course of the trial, HDL cholesterol levels increased from baseline by 4 to 11% in the placebo group and by 31 to 40% in the dalcetrapib group. Dalcetrapib had a minimal effect on LDL cholesterol levels. Patients were followed for a median of 31 months. At a prespecified interim analysis that included 1135 primary end-point events (71% of the projected total number), the independent data and safety monitoring board recommended termination of the trial for futility. As compared with placebo, dalcetrapib did not alter the risk of the primary end point (cumulative event rate, 8.0% and 8.3%, respectively; hazard ratio with dalcetrapib, 1.04; 95% confidence interval, 0.93 to 1.16; \( P=0.52 \)) and did not have a significant effect on any component of the primary end point or total mortality. The median C-reactive protein level was 0.2 mg per liter higher and the mean systolic blood pressure was 0.6 mm Hg higher with dalcetrapib as compared with placebo (\( P<0.001 \) for both comparisons).
OPINION

HDL—is it too big to fail?

Dominic S. Ng, Norman C. W. Wong and Robert A. Hegele

Abstract | The HDL hypothesis has suffered damage in the past few years. Clinical trials have shown that raising HDL cholesterol levels does not improve cardiovascular disease (CVD) outcomes. In addition, Mendelian randomization studies have shown that DNA variants that alter HDL cholesterol levels in populations are unrelated to incident CVD events. Balancing this deluge of negative data are substantial basic science data supporting the concept that raising HDL cholesterol levels reduces CVD risk. Also, functionally relevant HDL subfractions might be more important determinants of risk than overall HDL cholesterol levels. But, while wobbly, the HDL hypothesis is still standing, seemingly too big to fail owing to past intellectual, economic and psychological investments in the idea.

Ng, D. S. et al. Nat. Rev. Endocrinol. 9, 308–312 (2013); published online 15 January 2013; doi:10.1038/nrendo.2012.238
Limitations to Mendelian Randomisation

1- Pleiotropy

2- Population stratification

3- Canalisation

4- Power (also “weak instrument bias”)

5- The existence of instruments
30TH THOMAS FRANCIS JR MEMORIAL LECTURE

‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease?*

George Davey Smith and Shah Ebrahim

Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality

David M. Evans1,2 and George Davey Smith3

1University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland 4102, Australia. email: d.evans@uq.edu.au
2MRC Integrative Epidemiology Unit and School of Social and Community Medicine, University of Bristol, Bristol BS8 1SN, United Kingdom. email: julia.mackay@bristol.ac.uk

Keywords
causal analysis, structural equation modeling, genetic epidemiology, mining the phenome, pharmacogenomics
Fitting Mendelian Randomization Models in R
Equivalence Between Directed Acyclic Graphs and SEMs

SNP → Exposure → Outcome

Confounders

ξ₁

SNP → Exposure

βₓ

Exposure → Outcome

βᵧ

ξ₂
Simulation

\[ \beta_{ZX} \Rightarrow \beta_{UX} \Rightarrow \beta_{UY} = 0.1 \]

\[ \xi_1 \Rightarrow \xi_2 \]
Practical
Acknowledgments

- George Davey Smith
- Nic Timpson
- Sek Kathiresan