What is Mendelian Randomization and How Can it be Used as a Tool for Medicine and Public Health? Opportunities and Challenges

George Davey Smith
MRC Integrative Epidemiology Unit
University of Bristol

(@mendel_random)

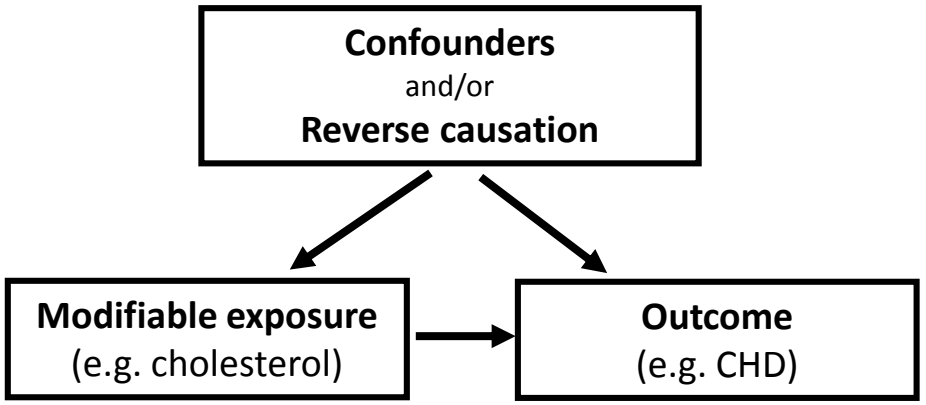
University of BRISTOL

MRC Integrative Epidemiology Unit

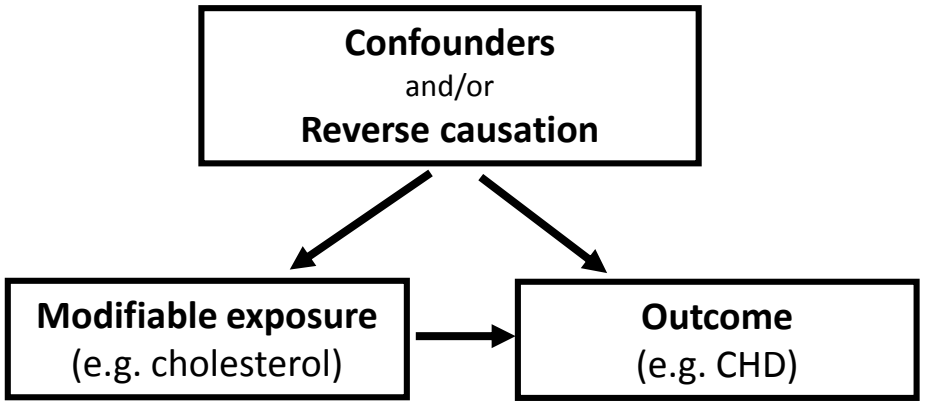
Outline

- Mendelian randomization outline
- Instrumental variables assumptions
- Heterogeneity and sensitivity analyses for assumption violation, including horizontal pleiotropy
- MR with interactions for estimation and checking violtations
- MR and disease liability

Conventional observational epidemiology

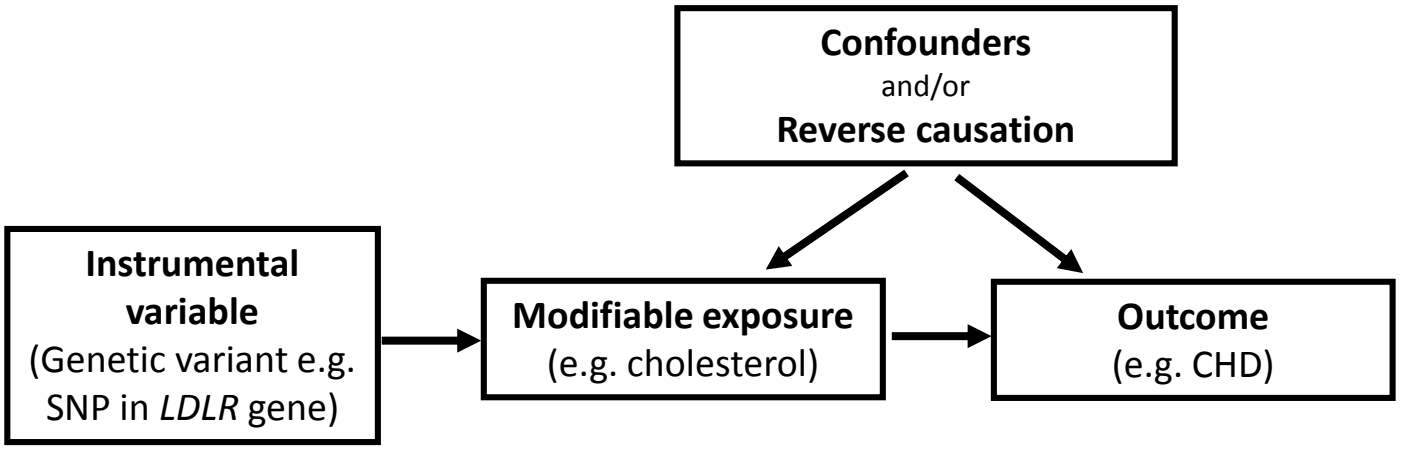


Conventional observational epidemiology



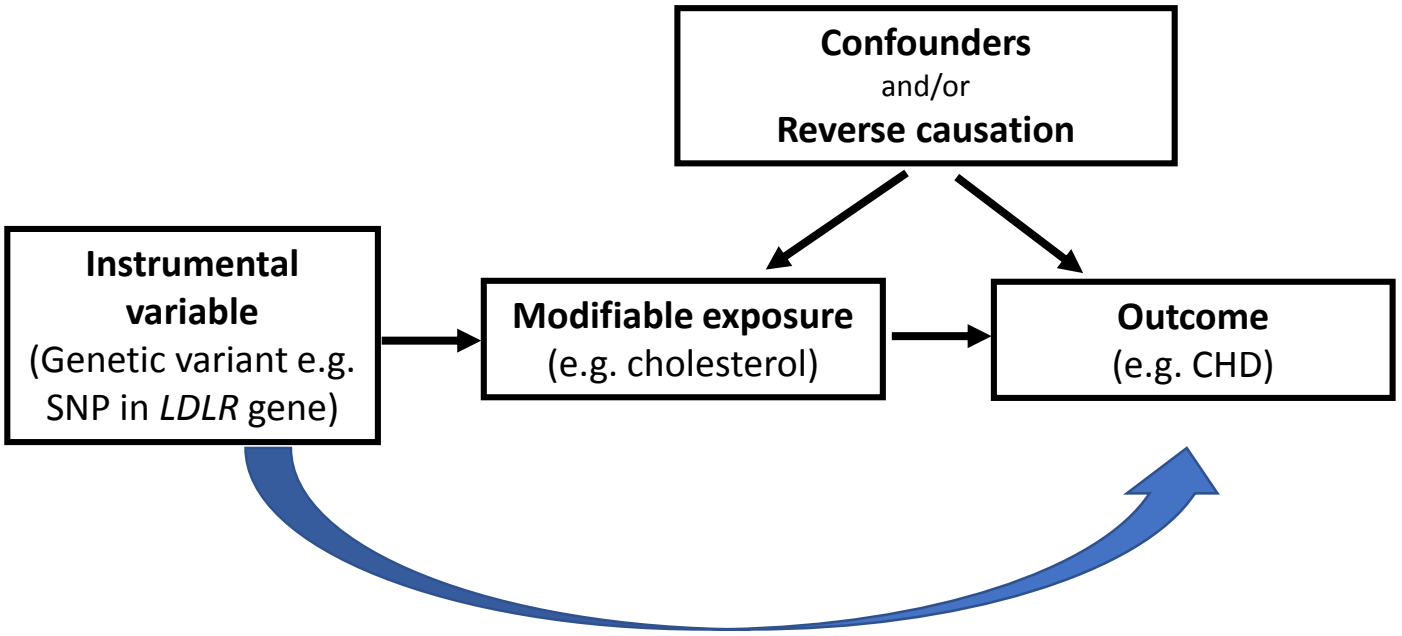
It is often impossible to exclude confounding and /or reverse causation as an explanation for observed exposure/outcome associations

Mendelian randomization approach

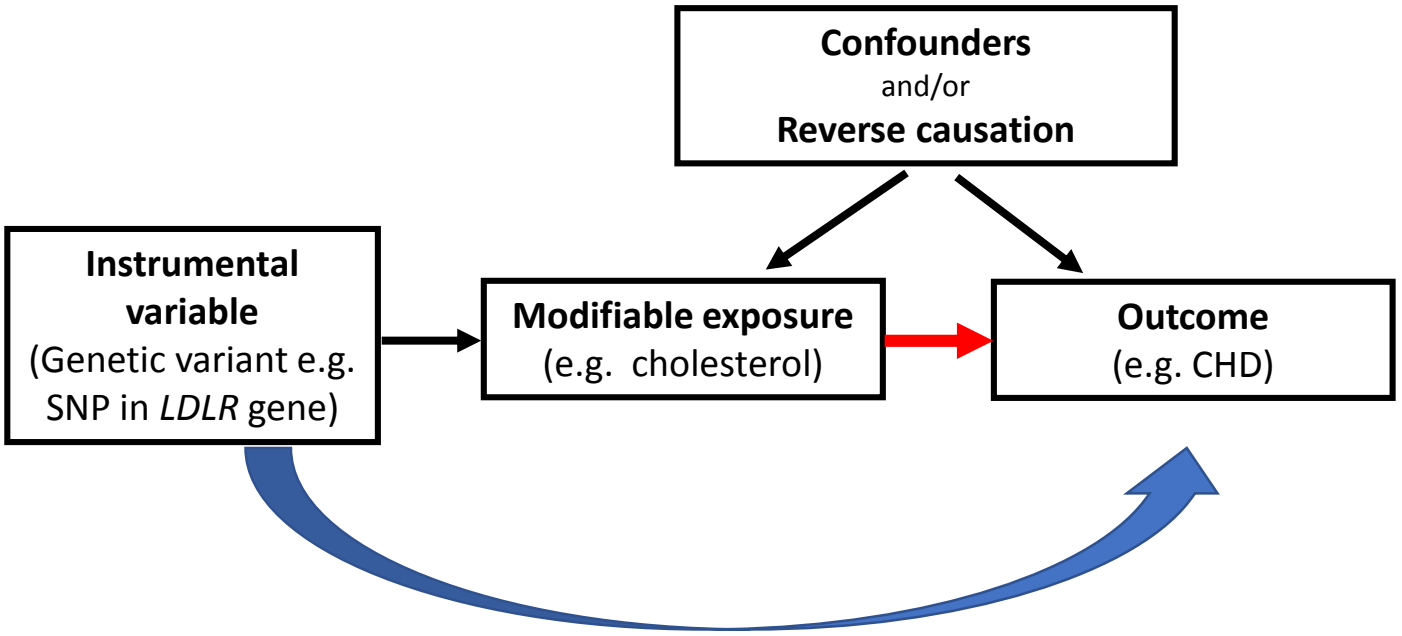


It is often impossible to exclude confounding and /or reverse causation as an explanation for observed exposure/outcome associations

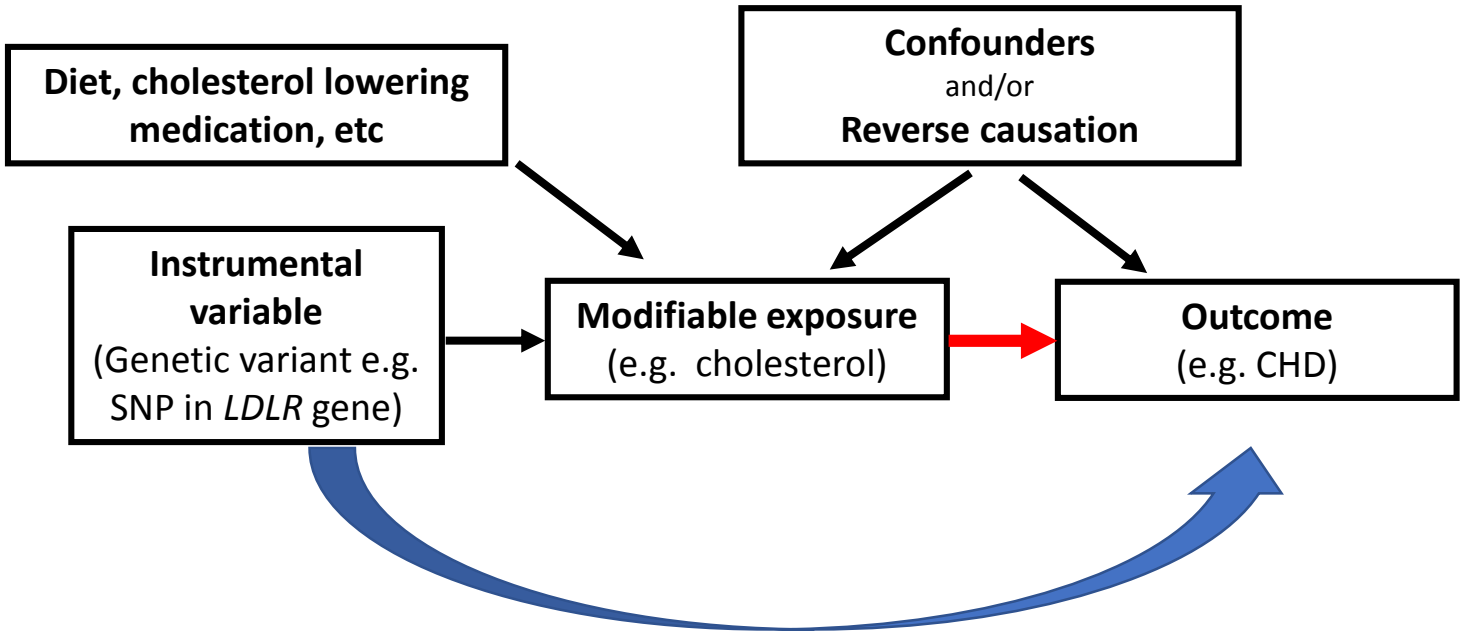
Mendelian randomization approach



Mendelian randomization approach



Mendelian randomization approach



Mendelian randomisation reflects the phenocopy / genocopy dialectic

e.g. Hartnup's syndrome and pellagra

"no doubt all environmental effects can be mimicked by one or several mutations"

(Zuckerkandl and Villet, PNAS 1988)

"Gene / environment equivalence" & "gene – environment interchangability" (West Eberhard 2003)

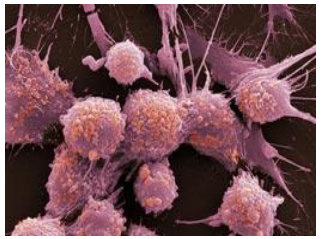
@BrisCancerEpi

Mendelian randomization:

and improving causal inference regarding prevention or treatment of disease

Traits inherited independent of each other & future environmental factors

Selenium



Instrumental variable, G → **Modifiable exposure, X** → **Outcome, Y**

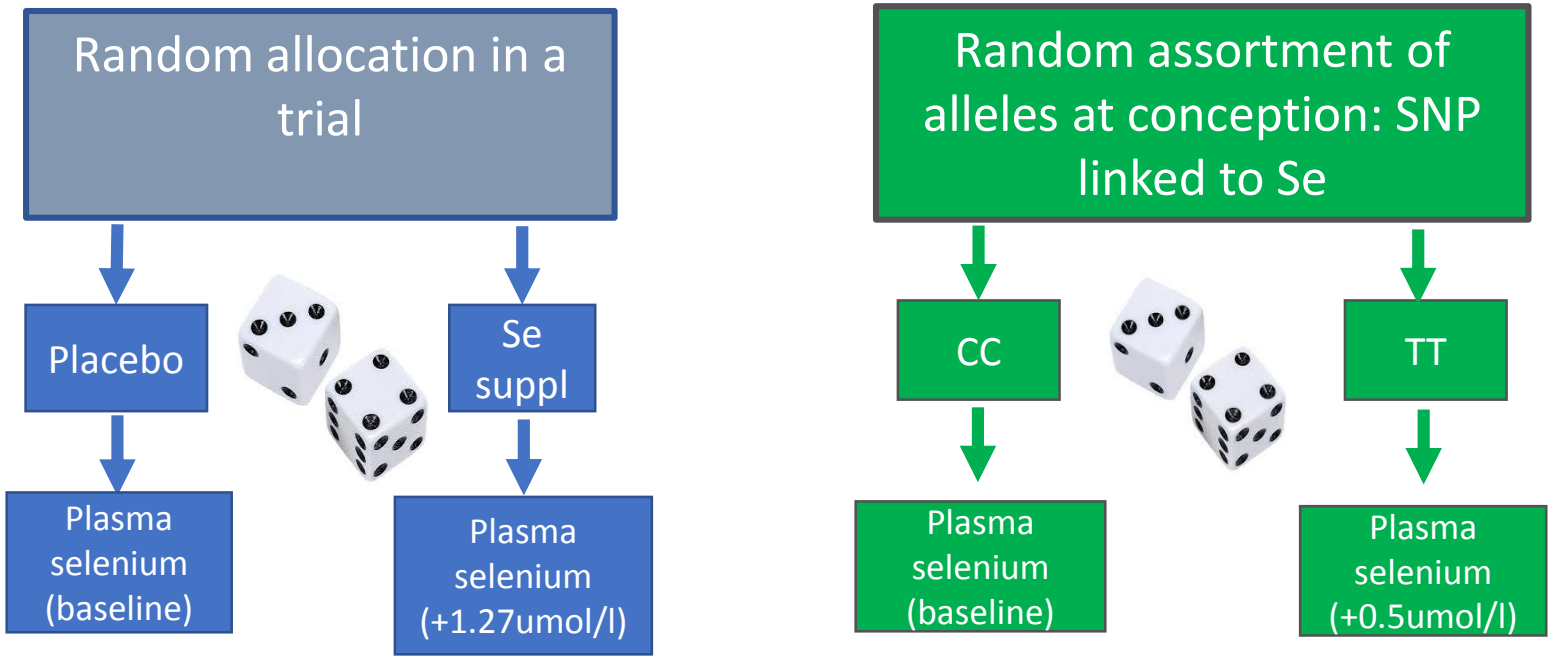
Confounders, U
Reverse causation

Nature's RCT

Free from confounding & reverse causation

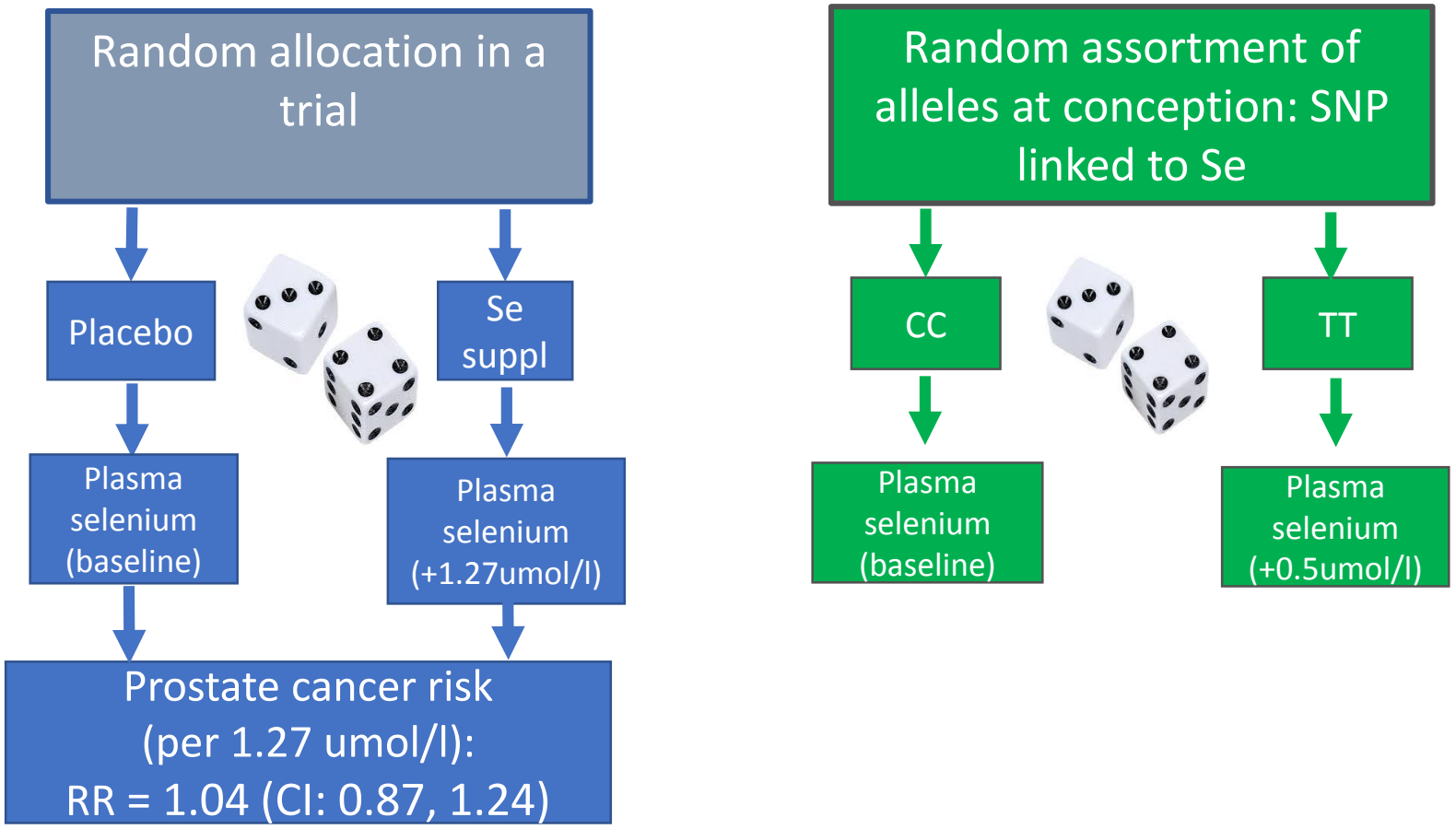
→ Selecting genetically supported targets doubles success at 25% lower costs

Mendelian randomization estimate of the effect of raising selenium on prostate cancer risk



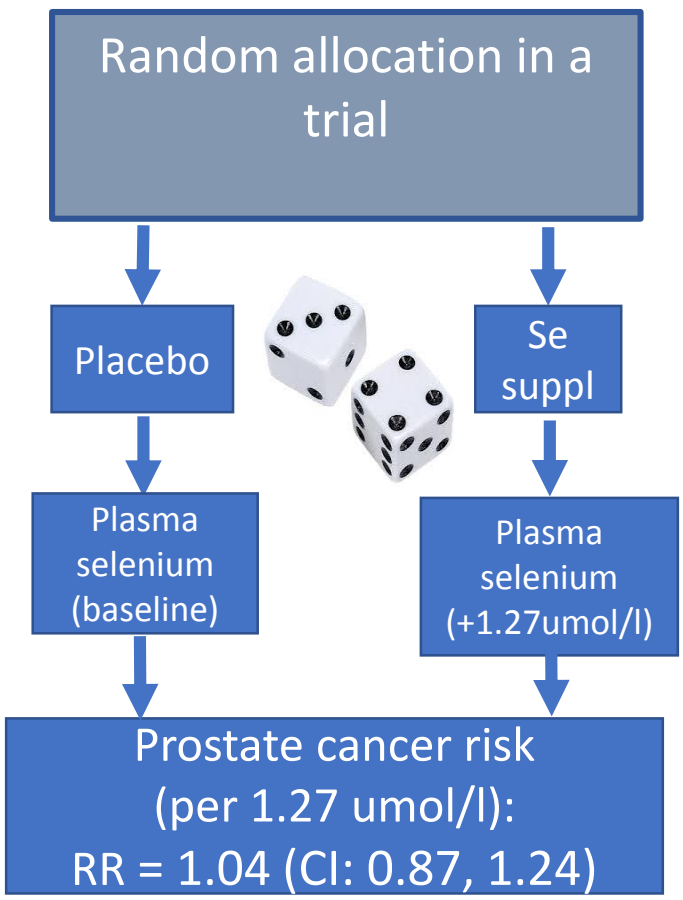
Yarmolinsky J et al. Circulating Selenium and Prostate Cancer Risk: A Mendelian Randomization Analysis. JNCI 2018;110:1035-1038

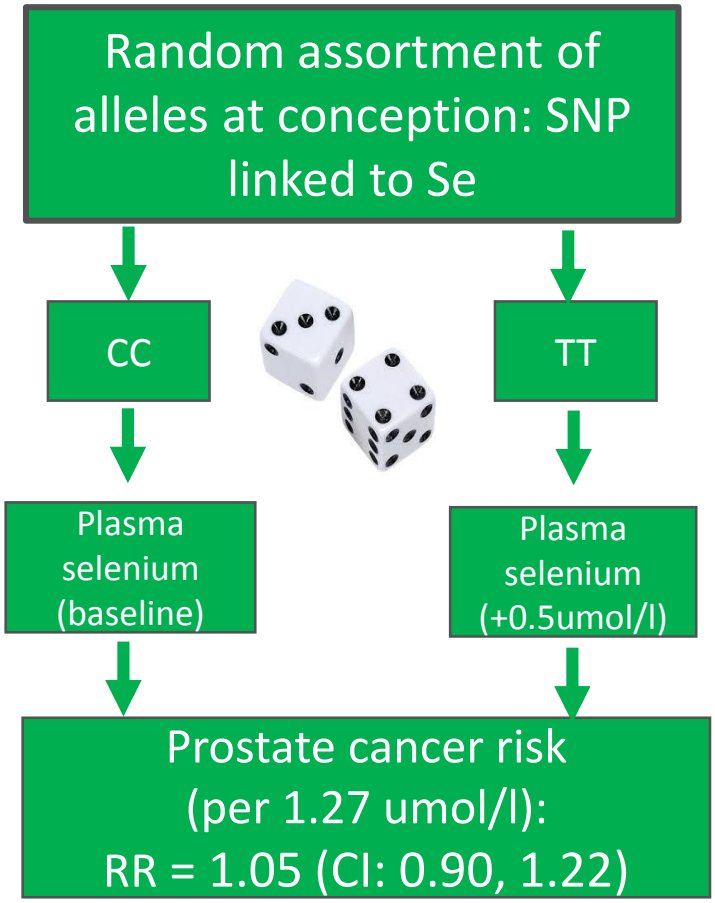
Mendelian randomization estimate of the effect of raising selenium on prostate cancer risk



Effect of Selenium and Vitamin E on risk of prostate cancer and other cancers; The Selenium and Vitamin E Cancer Prevention Trial (SELECT) *JAMA* 2009;301(1):39-51. **35,533 men**

Mendelian randomization estimate of the effect of raising selenium on prostate cancer risk

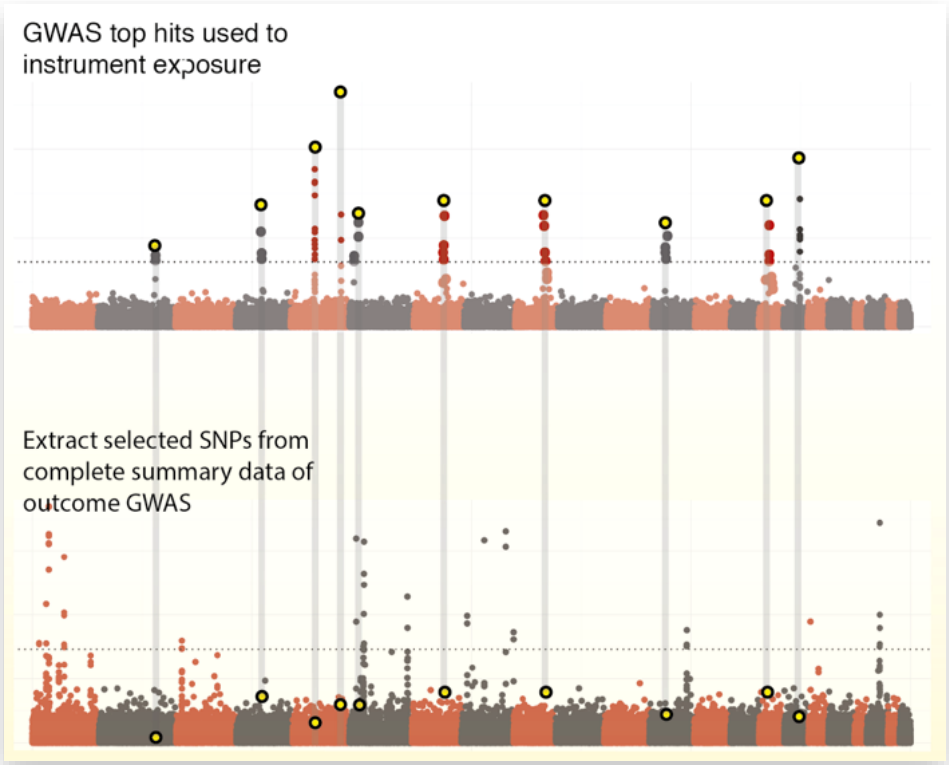




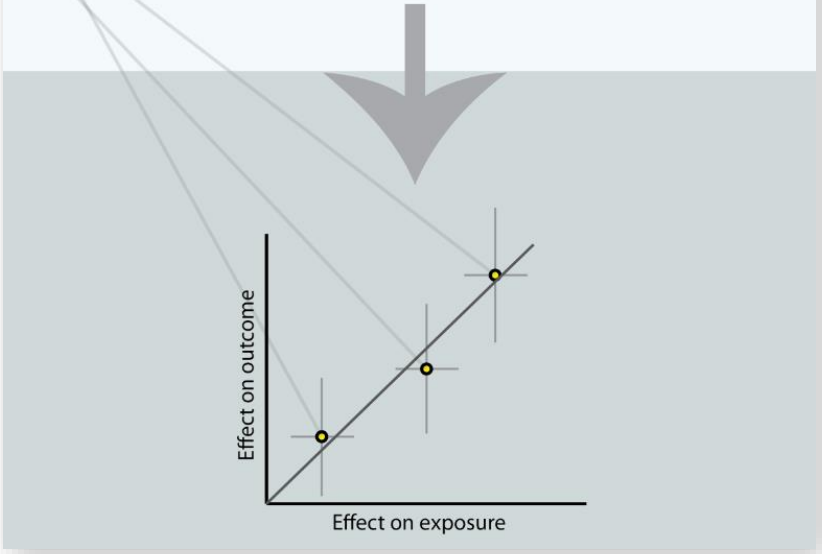
22,000 cases and 22,000 controls from 22 studies in PRACTICAL


Inheritance of Hemophilia
Equal Chance with Each Pregnancy
?
XX
XX
Carrier
XY
Hemophiliac
XY
Carrier mother (XX)
X
X
X
Y

2 sample method

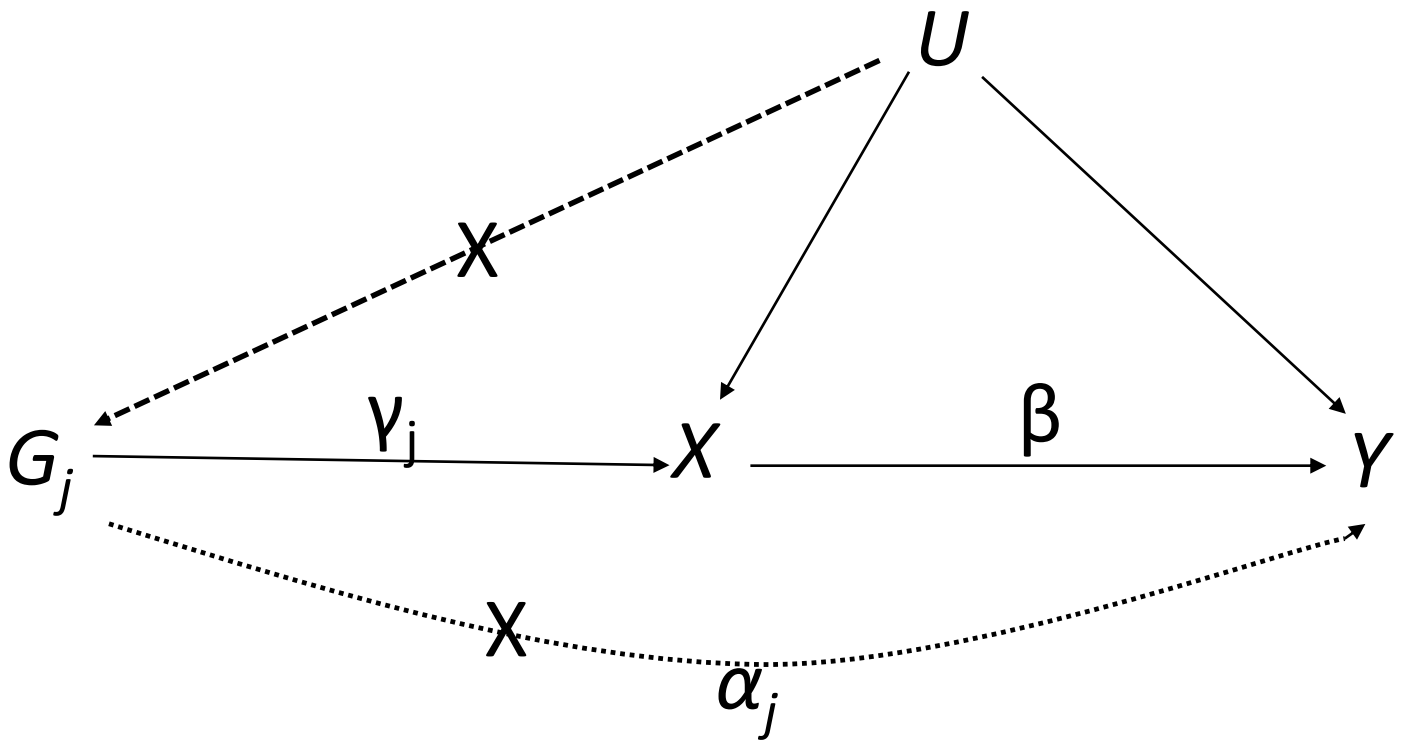


	Exposure GWAS				Outcome GWAS			
SNP	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*
rs123456	0.132	A	G	0.28	0.022	A	G	0.26
rs234567	-0.485	G	T	0.41	**-0.056**	**G**	**T**	**0.39**
rs345678	0.203	G	C	0.11	**0.046**	**G**	**C**	**0.12**



Instrumental variables

- **A genetic variant is a valid instrumental variable if the following assumptions hold:**
- **IV1**: The genetic variant is associated with the exposure X (“relevance”)
- **IV2**: The genetic variant is independent of confounders U (“exchangeability”)
- **IV3**: The genetic variant is independent of the outcome Y conditional on the exposure X (and confounders U) (“exclusion restriction”)



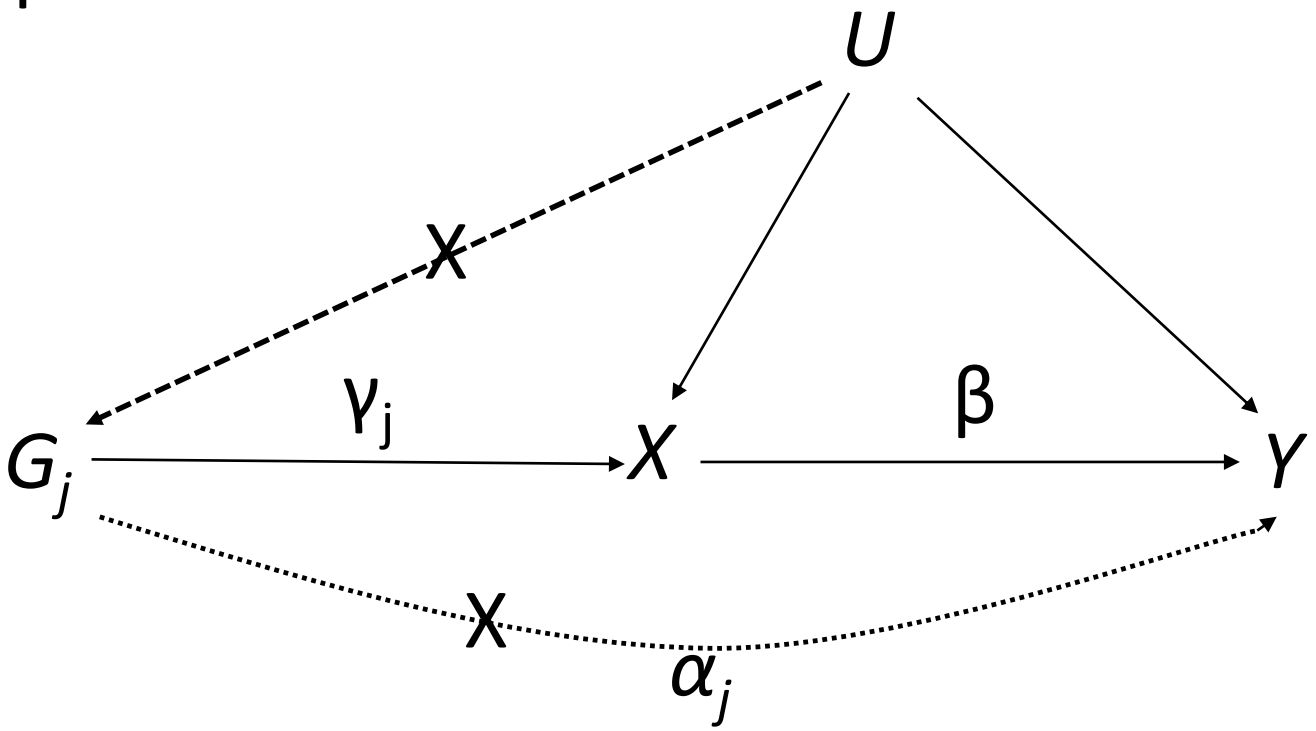
Do not condition on the intermediate phenotype

DO NOT CONDITION ON THE INTERMEDIATE PHENOTYPE

PLEASE DO NOT CONDITION ON THE INTERMEDIATE PHENOTYPE

MR and mediation

The IV assumptions necessary for MR are essentially the same as saying there is complete mediation of the effect of the genetic variant on the outcome by the intermediate phenotype / exposure of interest



In generating IV effect estimates

IV assumption 4 (in one of its varieties)

- Homogeneity of exposure effect on outcome

(allowing population average exposure effect to be estimated)

OR

- Monotonicity of instrument effects on exposure

(allowing complier average exposure effect to be estimated)

OR

- No interaction between G and U with respect to X

Can we interrogate IV4(h), IV4(m) and IV4(NI)?

Examine the variance of exposure and outcome by genotype (for continuous exposures or outcomes). If there is non-homogeneity or non-monotonicity violation there should be increased variance in the "treatment" genotype (whatever you consider the treatment allele: the basic inference is to there being some difference between genotypes)

The point has, I think, received the rather large amount of theoretical attention that it has chiefly through lack of contact with the practical experimental situation

RA Fisher to HE Daniels, 18th February 1938

nature genetics

TECHNICAL REPORT

https://doi.org/10.1038/s41588-018-0225-6

Identifying loci affecting trait variability and detecting interactions in genome-wide association studies

Alexander I. Young[1,2*], Fabian L. Wauthier[1,3] and Peter Donnelly[1,3*]

Identification of genetic variants with effects on trait variability can provide insights into the biological mechanisms that control variation and can identify potential interactions. We propose a two-degree-of-freedom test for jointly testing mean and variance effects to identify such variants. We implement the test in a linear mixed model, for which we provide an efficient algorithm and software. To focus on biologically interesting settings, we develop a test for dispersion effects, that is, variance effects not driven solely by mean effects when the trait distribution is non-normal. We apply our approach to body mass index in the subsample of the UK Biobank population with British ancestry (n ~408,000) and show that our approach can increase the power to detect associated loci. We identify and replicate novel associations with significant variance effects that cannot be explained by the non-normality of body mass index, and we provide suggestive evidence for a connection between leptin levels and body mass index variability.

Nearly all genome-wide association study (GWAS) associations have been discovered by testing the simplest additive model[1]. A long-standing controversy exists about the importance of departures from the additive model in human genetics[2]. In addition, the extent and nature of interactions between genetic variants and environmental factors remains poorly characterized. The advent of large population-based cohorts, such as the UK Biobank[3], provides large samples of genotyped individuals along with rich lifestyle and

algorithms that scale cubically with sample size[14,15], making them impractical to apply to the large sample sizes needed to detect variance effects on complex human traits.

Here we introduce a two-degree-of-freedom test for jointly testing mean and variance effects on quantitative traits. If the trait distribution is non-normal, then additive effects at a locus will induce variance effects that are unlikely to be of interest. We show how to account for this and thereby detect variance effects not driven by

Young AI et al. Identifying loci affecting trait variability and detecting interactions in genome-wide association studies. Nature Genetics 2018.

Interpreting IV estimates

- All of the IV assumptions …

Holmes et al. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nature Reviews Cardiology 2017;14:577-590

Interpreting IV estimates

- All of the IV assumptions …
- May indicate lifetime exposure effect (need measures of SNP-exposure across life)

Holmes et al. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nature Reviews Cardiology 2017;14:577-590

Interpreting IV estimates

- All of the IV assumptions …
- May indicate lifetime exposure effect (need measures of SNP-exposure across life)
- May relate only to exposure during critical period

Holmes et al. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nature Reviews Cardiology 2017;14:577-590

Interpreting IV estimates

- All of the IV assumptions …
- May indicate lifetime exposure effect (need measures of SNP-exposure across life)
- May relate only to exposure during critical period
- May relate to unmeasured upstream phenotype – e.g. in the case of enzyme activity

Holmes et al. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nature Reviews Cardiology 2017;14:577-590

Interpreting IV estimates

- All of the IV assumptions …
- May indicate lifetime exposure effect (need measures of SNP-exposure across life)
- May relate only to exposure during critical period
- May relate to unmeasured upstream phenotype – e.g. in the case of enzyme activity
- May only be interpretable in terms of liability to the "exposure", not the exposure itself

Holmes et al. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nature Reviews Cardiology 2017;14:577-590

Interpreting IV estimates

- All of the IV assumptions …
- May indicate lifetime exposure effect (need measures of SNP-exposure across life)
- May relate only to exposure during critical period
- May relate to unmeasured upstream phenotype – e.g. in the case of enzyme activity
- May only be interpretable in terms of liability to the "exposure", not the exposure itself
- Generally relates to disease occurrence, not outcome

Holmes et al. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nature Reviews Cardiology 2017;14:577-590

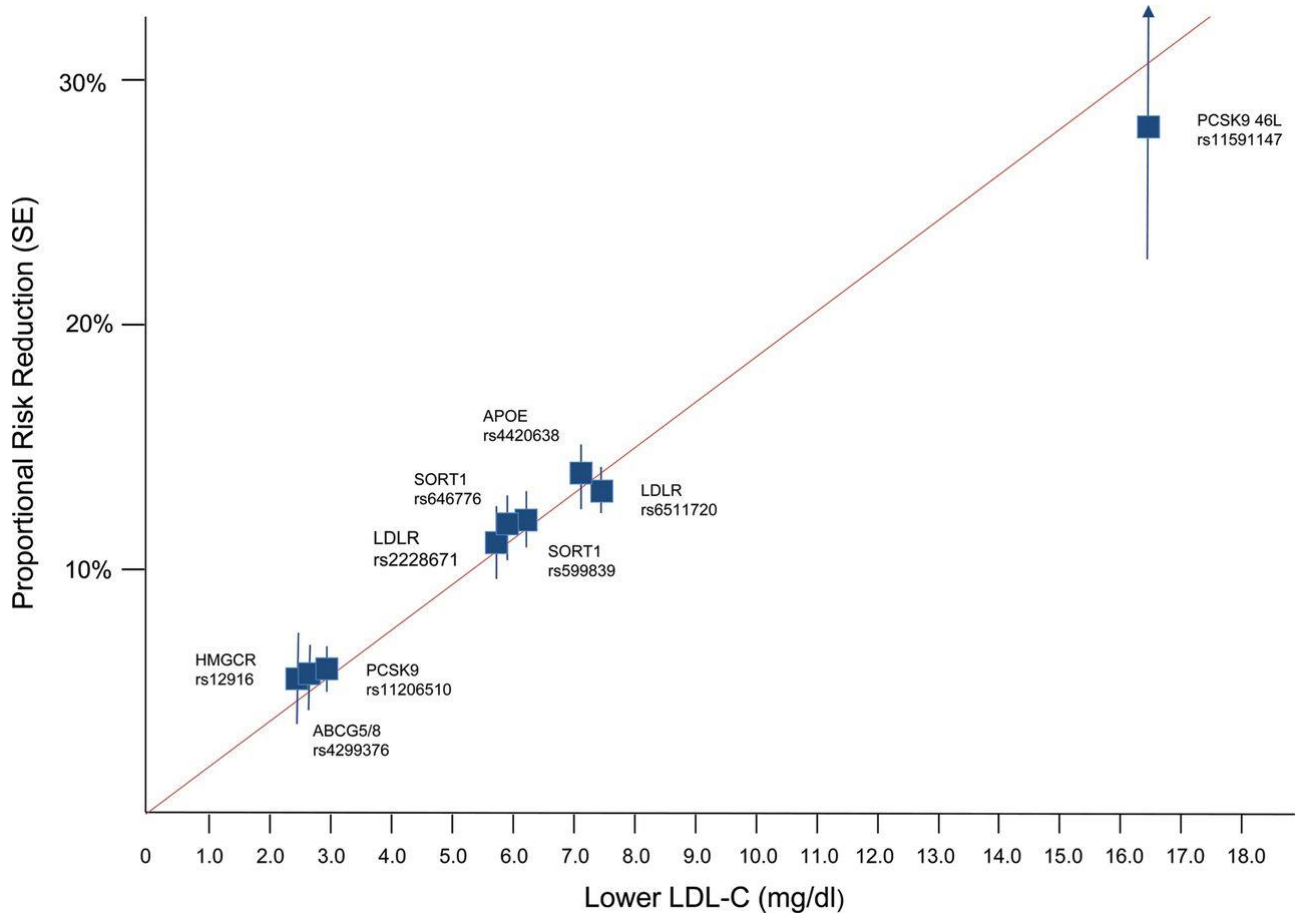
So .. why generate IV estimates?

- Necessary for most of the sensitivity analyses / extensions of MR analyses
- Comparison of IV estimates and what is observed in trials can help inform about relative immediacy of treatment effects
- In some cases (e.g. smoking during pregnancy) can be interpreted as the effect of an exposure in a particular period

Limitations of Mendelian Randomization

- Reintroduced confounding through horizontal pleiotropy

Effect of nine SNPs from six genes on LDL cholesterol and on CHD risk



Adapted from Ference B et al. J. Am. Coll. Cardiol., 2012;60:2631–2639 in Davey Smith G, Hemani G. Human Molecular Genetics 2014;r89-98

American Journal of Epidemiology

DOI: 10.1093/aje/kwy185

Invited Commentary

Invited Commentary: Detecting Individual and Global Horizontal Pleiotropy in Mendelian Randomization—A Job for the Humble Heterogeneity Statistic?

Jack Bowden*, Gibran Hemani, and George Davey Smith

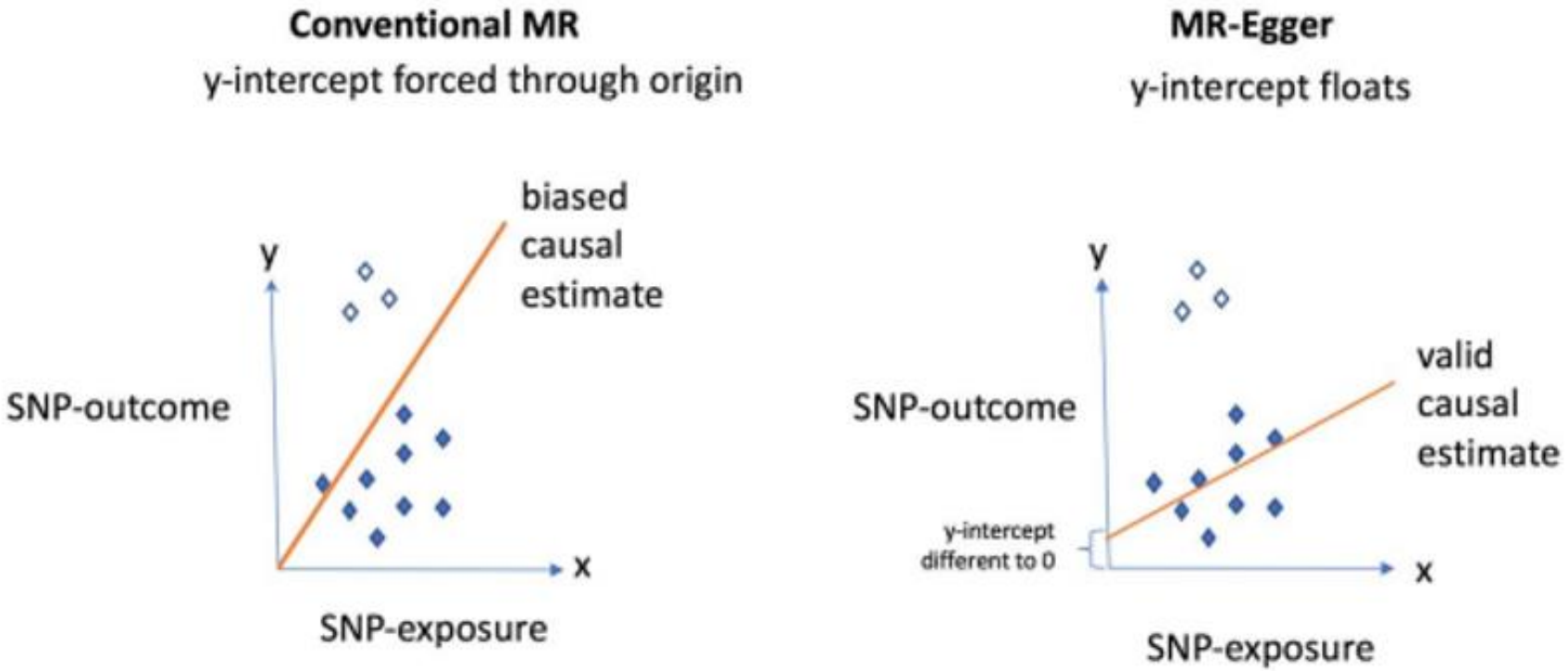
* Correspondence to Dr. Jack Bowden, MRC Integrative Epidemiology Unit, University of Bristol, Oakfield House, Bristol BS8 2BN, United Kingdom (e-mail: jack.bowden@bristol.ac.uk).

Initially submitted June 22, 2018; accepted for publication July 25, 2018.

Mendelian randomization (MR) is gaining in recognition and popularity as a method for strengthening causal inference in epidemiology by utilizing genetic variants as instrumental variables. Concurrently with the explosion in empirical MR studies, there has been the steady production of new approaches for MR analysis. The recently proposed "global and individual tests for direct effects" (GLIDE) approach fits into a family of methods that aim to detect horizontal pleiotropy—at the individual single nucleotide polymorphism level and at the global level—and to adjust the analysis by removing outlying single nucleotide polymorphisms. In this commentary, we explain how existing methods can (and indeed are) being used to detect pleiotropy at the individual and global levels, although not explicitly using this terminology. By doing so, we show that the true comparator for GLIDE is not MR-Egger regression (as Dai et al., the authors of the accompanying article (*Am J Epidemiol*. 2018;000(00):000–000), claim) but rather the humble heterogeneity statistic.

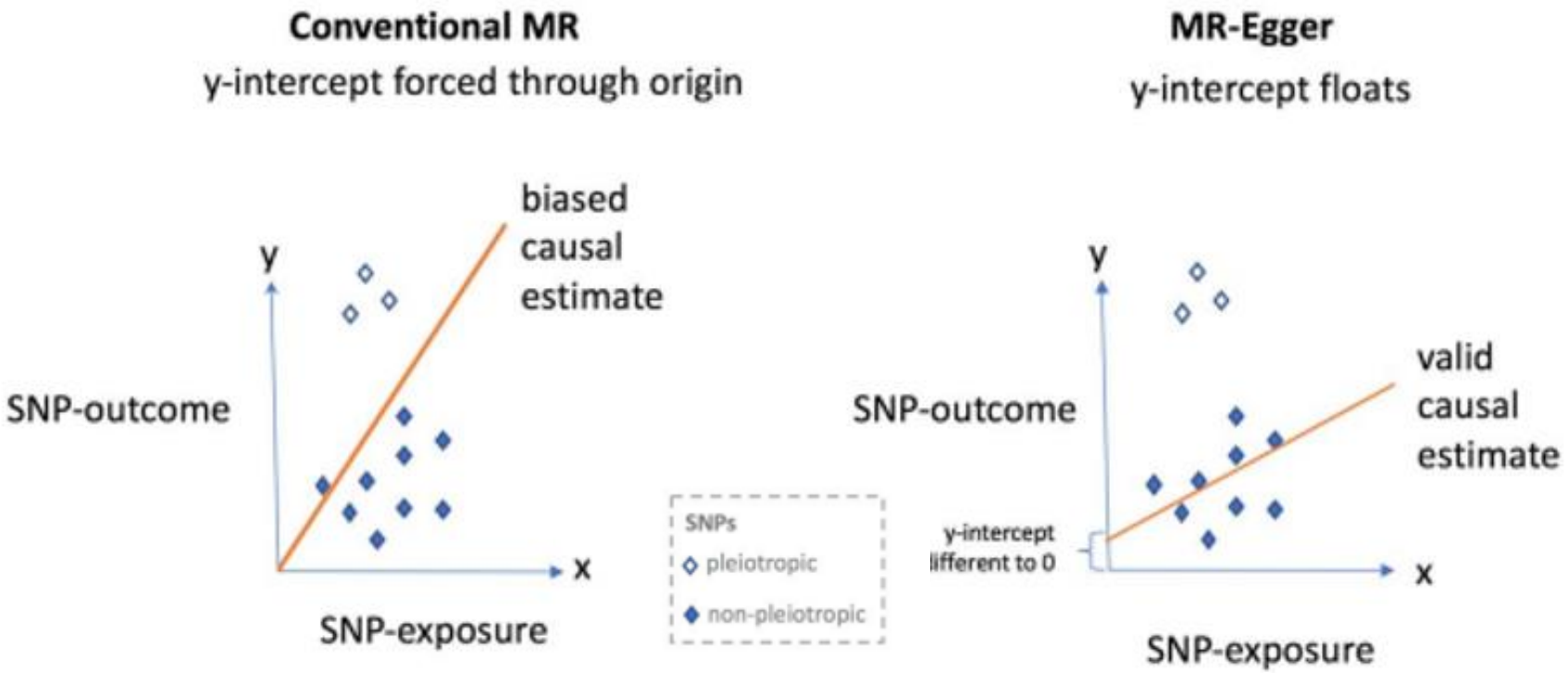
heterogeneity statistic; horizontal pleiotropy; Mendelian randomization; MR-Egger regression; outlier detection

MR-Egger



Bowden J et al. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol, 2015;44:512-525

MR-Egger



Requires Instrument Strength Independent of Direct Effect (InSIDE) assumption

Bowden J et al. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol, 2015;44:512-525

More arriving all the time ..

- Weighted median (Bowden et al, Genet Epidemiol 2016)
- Weighted mode (Hartwig et al, Int J Epidemiol 2017)
- MR-MoE (Hemani et al, bioRxiv 2017)
- Constrained instrumental variables (Jiang, bioRxiv 2017)
- MR-PRESSO (Verbanck et al, Nature Genetics, 2018)
- GLIDE (Dai et al, Am J Epidemiol 2018)
- MR-RAPS (Zhao et al, arXiv, 2018)
- Adaptive Lasso (Windmeijer et al, JASA 2018)
- MR-MIX (Qi et al, bioRxiv 2018)
- MRGxE (Spiller et al Int J Epidemiol 2018, from Chen et al PLoS Med 2008 etc)
- Etc
- Etc

Hemani G et al. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Human Molecular Genetics 2018;27:R195–R208,

GxE in an exposure propensity example: how does alcohol intake influence the risk of disease?

Association between clinically recorded alcohol consumption and initial presentation of 12 cardiovascular diseases: population based cohort study using linked health records

Steven Bell,[1,2] Marina Daskalopoulou,[3] Eleni Rapsomaniki,[4] Julie George,[4] Annie Britton,[2] Martin Bobak,[2] Juan P Casas,[4] Caroline E Dale,[4] Spiros Denaxas,[4] Anoop D Shah,[4] Harry Hemingway[4]

[1]Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge CB1 8RN, UK
[2]Research Department of Epidemiology and Public Health, University College London, London WC1E 7HB, UK
[3]Department of Infection and Population Health, University College London, Royal Free Hospital, London NW3 2PF, UK
[4]Farr Institute of Health Informatics Research (London), University College London, London NW1 2DA, UK
Correspondence to: S Bell
scb81@medschl.cam.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2017;356:j909
http://dx.doi.org/10.1136/bmj.j909

Accepted: 1 February 2017

ABSTRACT

OBJECTIVES

To investigate the association between alcohol consumption and cardiovascular disease at higher resolution by examining the initial lifetime presentation of 12 cardiac, cerebrovascular, abdominal, or peripheral vascular diseases among five categories of consumption.

DESIGN

Population based cohort study of linked electronic health records covering primary care, hospital admissions, and mortality in 1997-2010 (median follow-up six years).

SETTING

CALIBER (ClinicAl research using LInked Bespoke studies and Electronic health Records).

PARTICIPANTS

1937 360 adults (51% women), aged ≥30 who were free from cardiovascular disease at baseline.

MAIN OUTCOME MEASURES

12 common symptomatic manifestations of cardiovascular disease, including chronic stable angina, unstable angina, acute myocardial infarction, unheralded coronary heart disease death, heart failure, sudden coronary death/cardiac arrest, transient ischaemic attack, ischaemic stroke, intracerebral and subarachnoid haemorrhage, peripheral arterial disease, and abdominal aortic aneurysm.

RESULTS

114 859 individuals received an incident cardiovascular diagnosis during follow-up. Non-drinking was associated with an increased risk of unstable angina (hazard ratio 1.33, 95% confidence interval 1.21 to 1.45), myocardial infarction (1.32, 1.24 to1.41), unheralded coronary death (1.56, 1.38 to 1.76), heart failure (1.24, 1.11 to 1.38), ischaemic stroke (1.12, 1.01 to 1.24), peripheral arterial disease (1.22, 1.13 to 1.32), and abdominal aortic aneurysm (1.32, 1.17 to 1.49) compared with moderate drinking (consumption within contemporaneous UK weekly/daily guidelines of 21/3 and 14/2 units for men and women, respectively). Heavy drinking (exceeding guidelines) conferred an increased risk of presenting with unheralded coronary death (1.21, 1.08 to 1.35), heart failure (1.22, 1.08 to 1.37), cardiac arrest (1.50, 1.26 to 1.77), transient ischaemic attack (1.11, 1.02 to 1.37), ischaemic stroke (1.33, 1.09 to 1.63), intracerebral

Moderate drinking can lower risk of heart attack, says study

Drinking in moderation helps protect heart, with study finding it lowers risk of many conditions compared with not drinking

Authors described study as most comprehensive to date on relationship between alcohol consumption and heart health. Photograph: Hero Images/Getty Images

Moderate drinking can lower the risk of several heart conditions, according to a study that will further fuel the debate about the health implications of alcohol consumption.

INDEPENDENT

Drinking pint of beer a day linked to reduced risk of heart attack

Moderate drinkers less likely than teetotallers and heavy drinkers to see doctor for heart conditions including angina and strokes caused by blood clots

Katie Forster | @katieforster | 11 hours ago | 20 comments

A pint a day may keep the doctor away when it comes to some cardiovascular diseases *Getty*

Cheers! Drinkers who have one glass of wine a night 'are at less risk of heart failure than teetotallers'

- **Moderate drinking equates to within 14 units of alcohol a week**
- **This is equivalent to a pint of very weak beer, or two measures of spirits**
- **Cambridge University researchers believe moderate amounts of alcohol may boost levels of good cholesterol in the blood**

Moderate drinking slashes the risk of a heart attack or stroke, a major study reveals today.

Research from 1.9million adults appears to show that the occasional glass of wine or pint of beer may provide a protective effect.

Men and women who drank moderately – within 14 units of alcohol a week – were found to be less at risk of common heart problems than teetotallers.

Drinking alcohol slashes risk of heart problems – if you drink this much per week

BOOZERS rejoice because drinking alcohol can boost your heart health.

By Sarah Buchanan / Published 22nd March 2017

HEART HEALTH: Guzzling a moderate amount of booze can have a beneficial effect

Getty Images

HEART DISEASE

Alcohol Is Good for Your Heart—Most of the Time

Alice Park
Mar 22, 2017

TIME Health

For more, visit TIME Health.

THE IRISH TIMES

Thu, Mar 23, 2017

Moderate drinking may cut risk of heart disease

Research links light alcohol consumption with less risk of heart and blood vessel illness

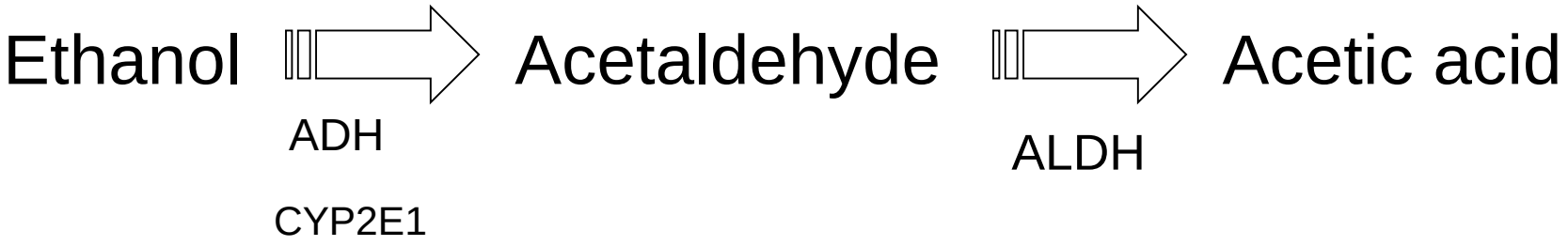
about 12 hours ago

Paul Cullen

A study published in the British Medical Journal looked at the health records of 1.93 million UK adults who were all

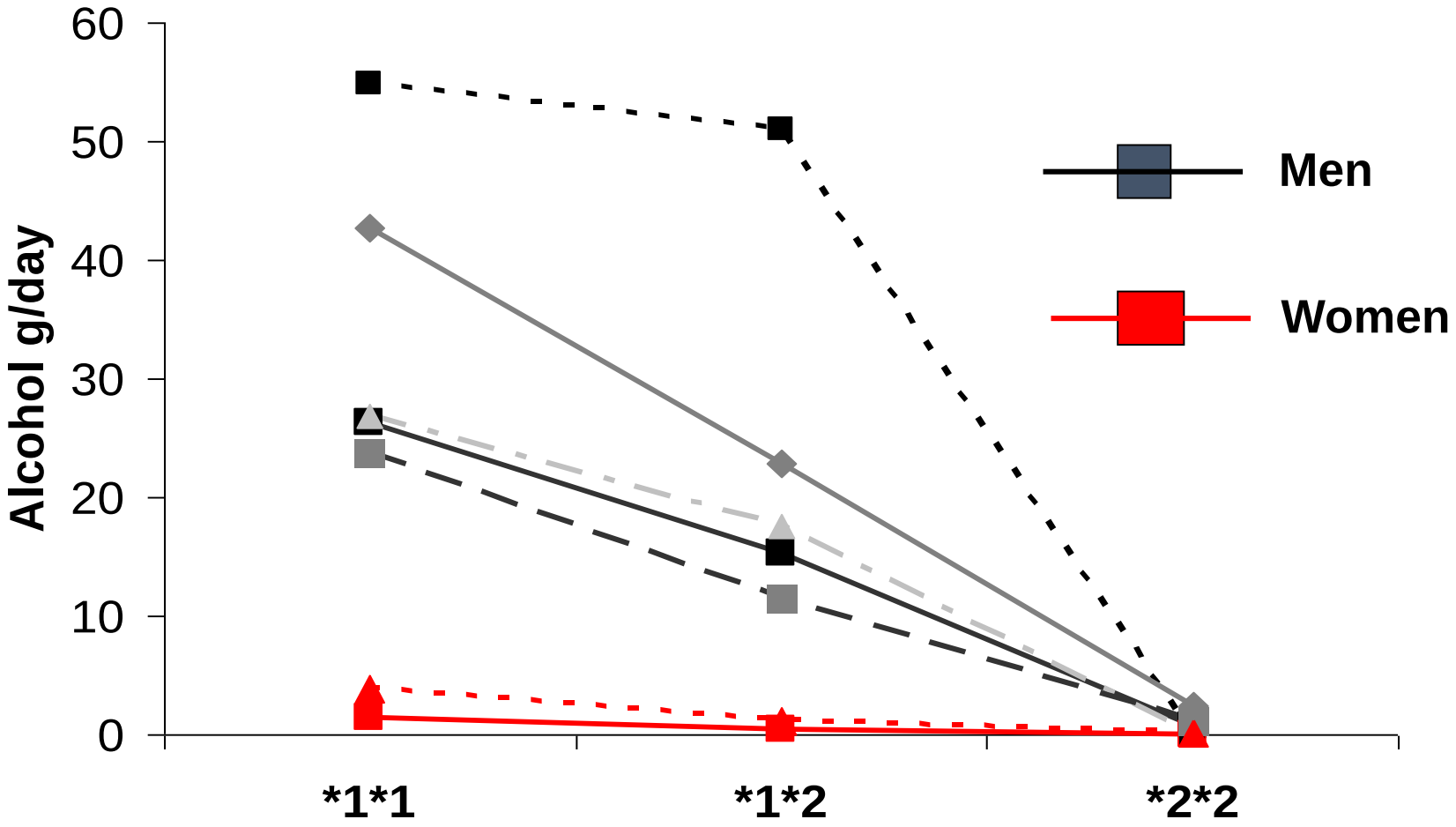
Moderate drinking may be good for your heart and even heavy drinking may lower your risk of heart attack, a new study indicates.

Metabolism of alcohol



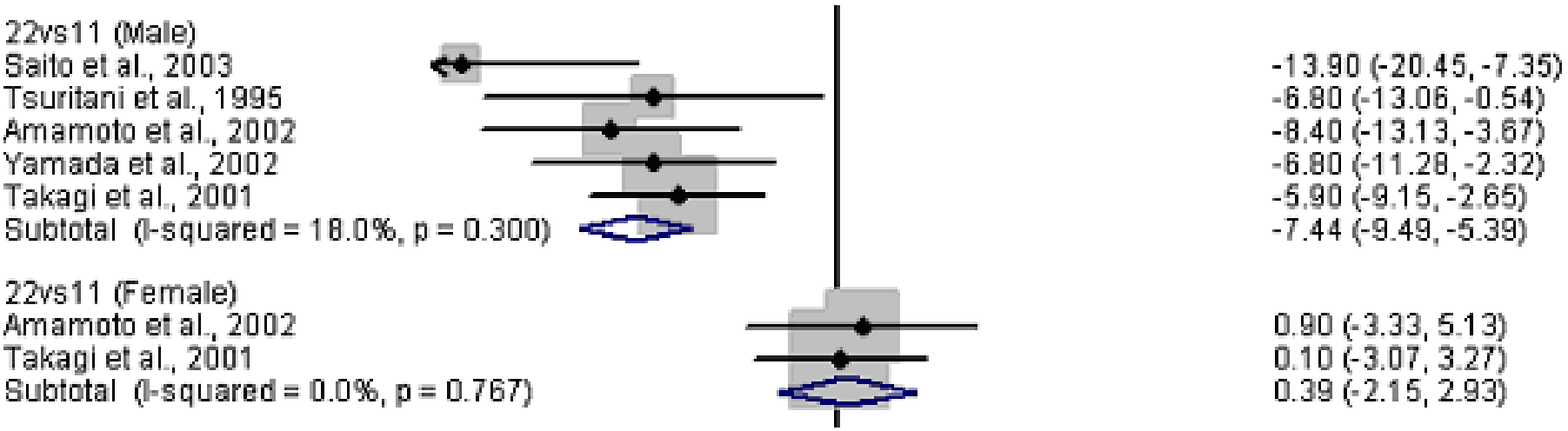
* Mainly occurs in the liver, but some activity is also present in the oral cavity and digestive tract

ALDH2 genotype by alcohol consumption, g/day: 5 studies, n=6815



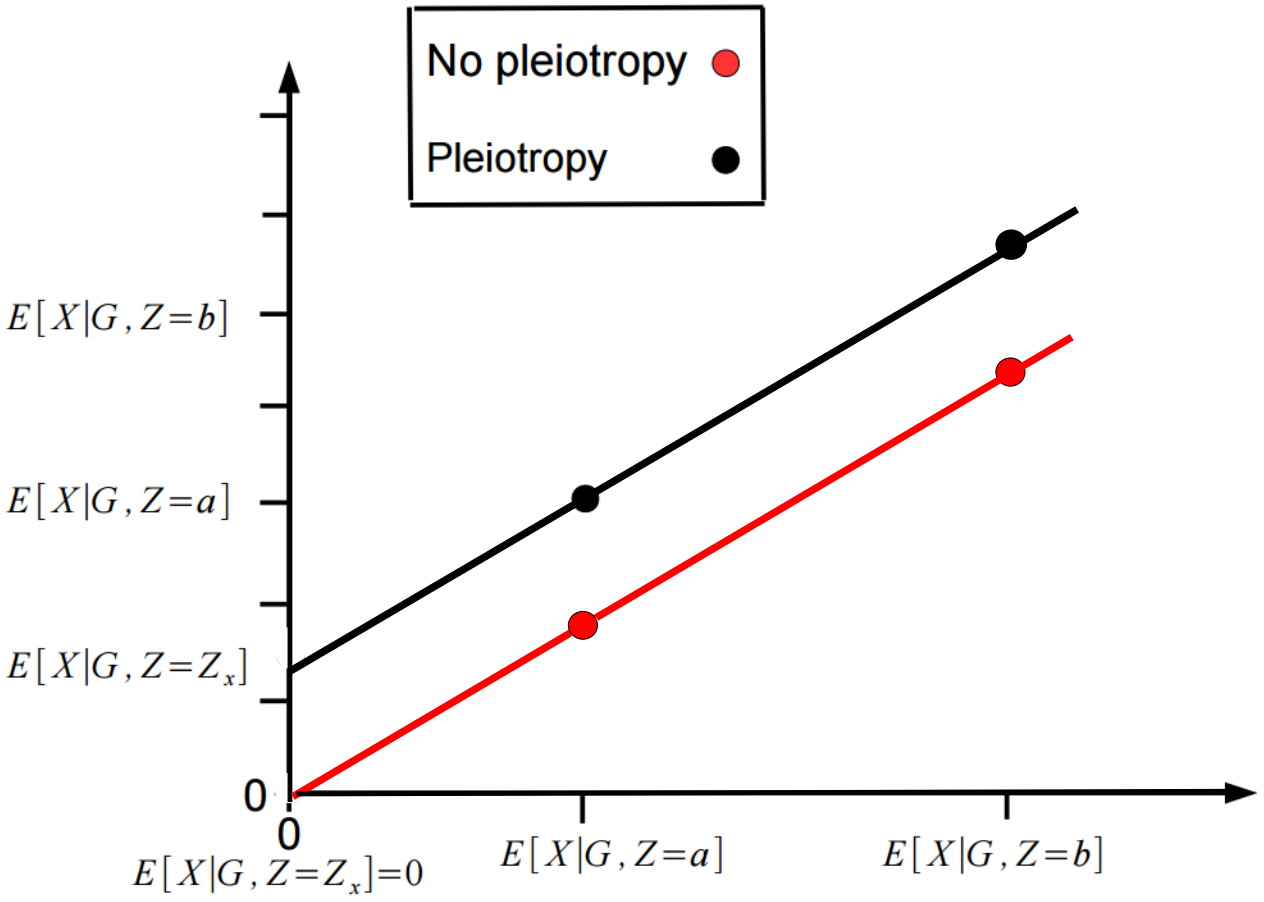
Chen, Davey Smith et al, PLoS Med 2008

ALDH2 genotype and systolic blood pressure



Chen et al, PLoS Medicine 2008

Effect estimation using interactions



Spiller et al, Int J Epidemiol 2018

Instrumental variable estimates of alcohol intake (g/day) on CVD risk factors, based on ALDH2 by sex interactions

	Beta coefficient (95% CI) by IV estimation*	P-value
Systolic blood pressure (mmHg)	0.202 (0.087, 0.317)	0.001
Diastolic blood pressure (mmHg)	0.100 (0.025, 0.175)	0.009
Total cholesterol (mg/dL)	0.038 (-0.197, 0.273)	0.750
HDL cholesterol (mg/dL)	0.166 (0.098, 0.233)	<0.0001
LDL cholesterol (mg/dL)	-0.362 (-0.578, -0.145)	0.001
Log-transformed triglycerides (log(mg/dL))	0.003 (0.002, 0.005)	<0.0001

Cho Y, Shin S, Won S, Relton CL, Davey Smith G, Shin M. Alcohol intake and cardiovascular risk factors: A Mendelian randomisation study. Scientific Reports 2015;5:18422: doi: 10.1038/srep18422

nature neuroscience

ARTICLES

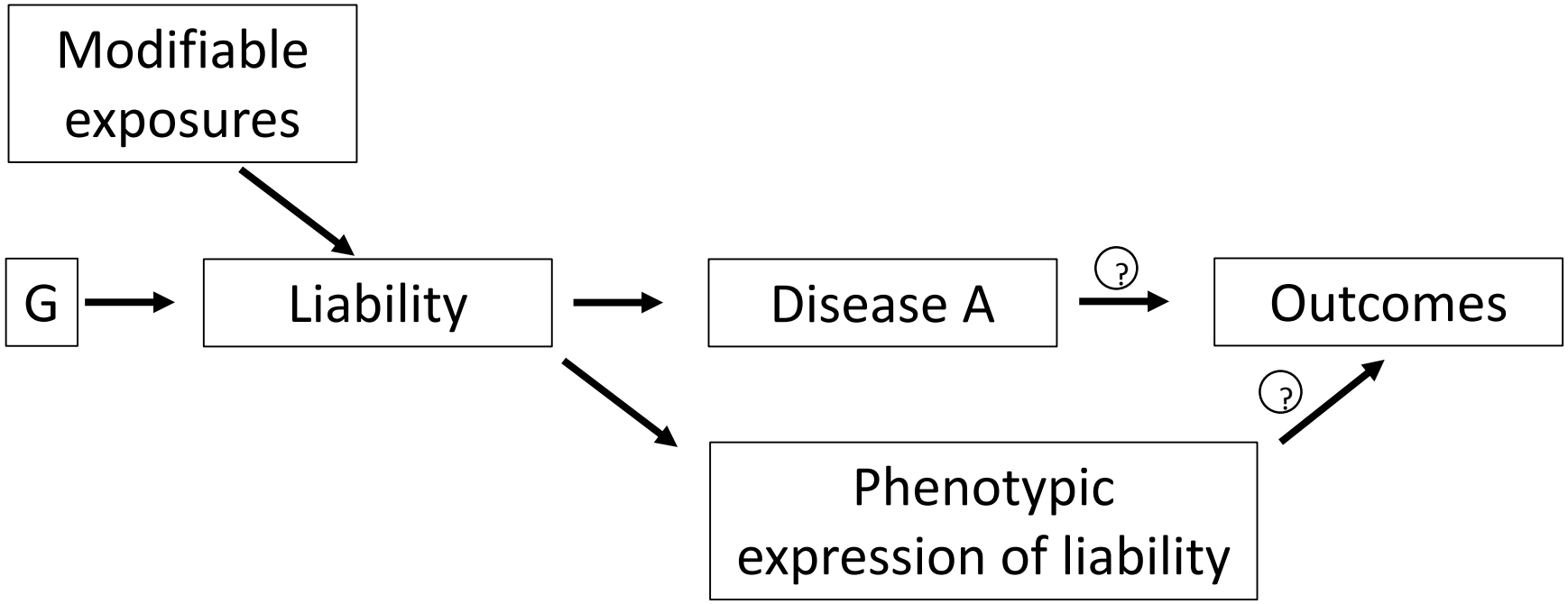
https://doi.org/10.1038/s41593-018-0206-1

GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia

Joëlle A. Pasman[1,37], Karin J. H. Verweij[1,2,37], Zachary Gerring[3], Sven Stringer[4], Sandra Sanchez-Roige[5], Jorien L. Treur[6], Abdel Abdellaoui[2], Michel G. Nivard[7], Bart M. L. Baselmans[7], Jue-Sheng Ong[3], Hill F. Ip[7], Matthijs D. van der Zee[7], Meike Bartels[7], Felix R. Day[8], Pierre Fontanillas[9], Sarah L. Elson[9], the 23andMe Research Team[10], Harriet de Wit[11], Lea K. Davis[12], James MacKillop[13], The Substance Use Disorders Working Group of the Psychiatric Genomics Consortium[14], International Cannabis Consortium[15], Jaime L. Derringer[16], Susan J. T. Branje[17], Catharina A. Hartman[18], Andrew C. Heath[19], Pol A. C. van Lier[20], Pamela A. F. Madden[19], Reedik Mägi[21], Wim Meeus[17], Grant W. Montgomery[22], A. J. Oldehinkel[18], Zdenka Pausova[23], Josep A. Ramos-Quiroga[24,25,26,27], Tomas Paus[28,29], Marta Ribases[24,25,26], Jaakko Kaprio[30], Marco P. M. Boks[31], Jordana T. Bell[32], Tim D. Spector[32], Joel Gelernter[33], Dorret I. Boomsma[7], Nicholas G. Martin[3], Stuart MacGregor[3], John R. B. Perry[8], Abraham A. Palmer[5,34], Danielle Posthuma[4], Marcus R. Munafò[6,35], Nathan A. Gillespie[3,36,38], Eske M. Derks[3,38] and Jacqueline M. Vink[1,38*]

Cannabis use is a heritable trait that has been associated with adverse mental health outcomes. In the largest genome-wide association study (GWAS) for lifetime cannabis use to date ($N = 184,765$), we identified eight genome-wide significant independent single nucleotide polymorphisms in six regions. All measured genetic variants combined explained 11% of the variance. Gene-based tests revealed 35 significant genes in 16 regions, and S-PrediXcan analyses showed that 21 genes had different expression levels for cannabis users versus nonusers. The strongest finding across the different analyses was CADM2, which has been associated with substance use and risk-taking. Significant genetic correlations were found with 14 of 25 tested substance use and mental health-related traits, including smoking, alcohol use, schizophrenia and risk-taking. Mendelian randomization analysis showed evidence for a causal positive influence of schizophrenia risk on cannabis use. Overall, our study provides new insights into the etiology of cannabis use and its relation with mental health.

Interpretation of MR studies in which disease A is the apparent exposure



International Journal of Epidemiology, 2016, 1866–1886
doi: 10.1093/ije/dyw314
Advance Access Publication Date: 20 January 2017
Original article

Approaches to causal inference

Triangulation in aetiological epidemiology

Debbie A Lawlor,[1,2,*] Kate Tilling[1,2] and George Davey Smith[1,2]

[1]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK and [2]School of Social and Community Medicine, University of Bristol, Bristol, UK

*Corresponding author. MRC IEU, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK. E-mail: d.a.lawlor@bristol.ac.uk

Accepted 3 October 2016

4th International Mendelian randomization conference

17th to 19th July, 2019

School of Chemistry, Bristol, UK

Welcome

The meeting will focus on the development, application and translation of Mendelian randomization methods to a range of fields. We encourage all with an interest in causality, including from social science, clinical sciences, public health, biomedical research, epidemiology, statistics and the pharmaceutical industry to join us and contribute to discussion on new methodologies and wide scale implementation.

Key Dates

20th Feb 2019

First call abstract submission deadline

10th July 2019

Registration closes

www.mendelianrandomization.org.uk

Further Reading

- Davies NM et al. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ 2018;362:k601
- Holmes MV et al. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nature Reviews Cardiology 2017;14:577-590
- Spiller W et al. Detecting and correcting for bias in Mendelian randomization analyses using gene-by-environment interactions. Int J Epidemiol 2018: dyy204.
- Sanderson E et al. An examination of multivariable Mendelian randomization in the single sample and two-sample summary data settings. Int J Epidemiol 2018, in press (on bioRxiv)
- Hemani G et al. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Human Molecular Genetics 2018;27:R195-R208.
- Bowden J et al. Improving the visualisation, interpretation and analysis of two-sample summary data Mendelian randomization via the radial plot and radial regression. Int J Epidemiol 2018;47:1264-1278
- Bowden J et al. Detecting individual and global horizontal pleiotropy in Mendelian randomization: a job for the humble heterogeneity statistic. Am J Epidemiol 2018, kwy185.