

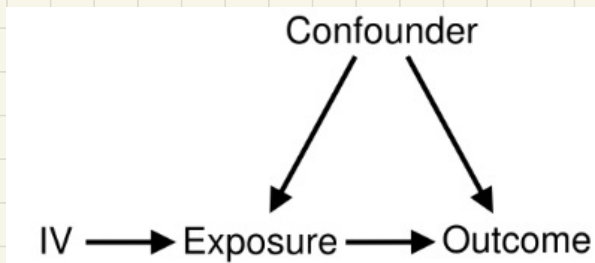
§ 1. MR第二讲

Section 1.1 复习

1.1.1 孟德尔随机化主要解决两个问题

1. 暴露和结局是否存在因果关系？
2. 如果存在因果关系，那么大小是多少？

1.1.2 孟德尔随机化的整体模型



- i. the variant is associated with the exposure,
- ii. the variant is not associated with the outcome via a confounding pathway,
- iii. the variant does not affect the outcome directly, only possibly indirectly via the exposure.

1. 工具变量与暴露密切相关；
2. 工具变量与任何影响暴露-结局关联的混杂因素均不相关；
3. 除非借助与暴露的关联，否则工具变量不会影响结局。

1.1.3 估计因果效应时还需要满足的工具变量假设

1. 单调性(Monotonicity)。这意味着对于人口中的所有个体来说，改变IV所带来的暴露的任何变化应该都是同方向的。 $X(2) \geq X(1) \geq X(0)$ 或者 $X(0) \geq X(1) \geq X(2)$

2. 同质性(Homogeneity)。这意味着暴露对结果的因果效应对人口中的所有个体来说是恒定的

1.1.4 一些别的常见假设

1. 线性性 / 对数线性性

1.1.5 孟德尔随机化抽象理解

观测结果 \rightarrow 因果结果

$$\beta(IV) = \beta(YZ) / \beta(XZ)$$

例子：Lp-PLA2与CHD之间的因果关系

Section 1.2 可能违反工具变量假设的情况

I. 生物学机制

1. 遗传多效性 / pleiotropy (垂直多效性和水平多效性)

多效性是指与多种暴露相关的遗传变异。

2. 渠化 / canalization

生物体的补偿机制，允许某基因通过不同的生物途径表达。

II. 非孟德尔式遗传

1. 连锁不平衡 / linkage disequilibrium

不同的遗传变异因为在同一染色体上物理位置接近，而被一起遗传。

2. 效应修正 / effect modification

外部因子改变了研究因素与疾病的关联程度。

III. 群体效应

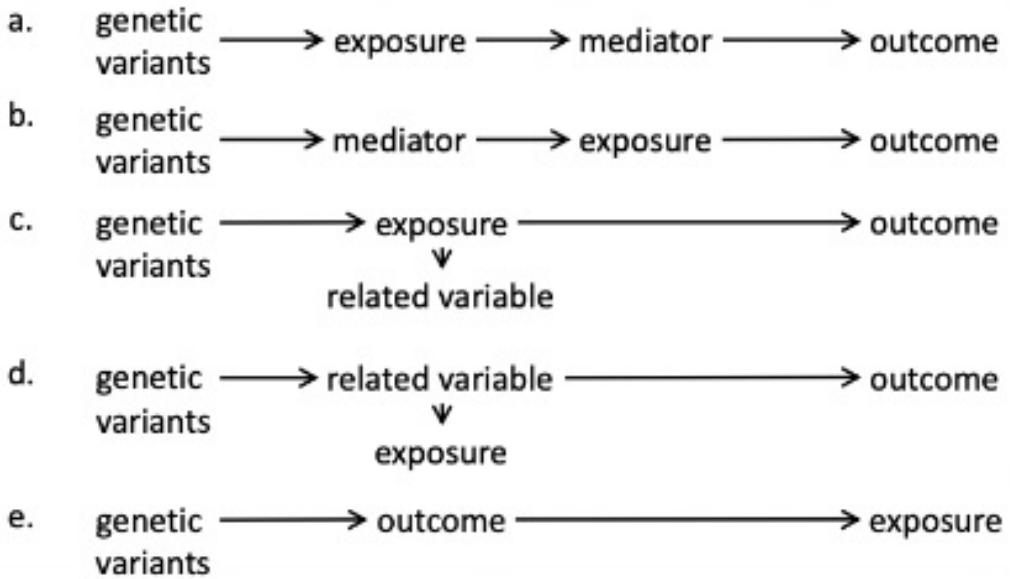
1. 人群分层 / population stratification

如果遗传变异的频率和暴露的分布在不同的人口亚群中是不同的，就会导致变异与暴露之间的错误关联。

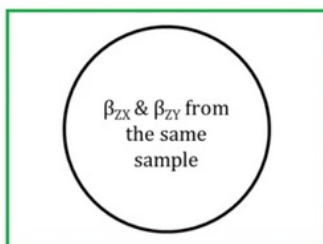
2. 确定性效应 / ascertainment effect

决策者会加重对被认为是确定性结果的选择。

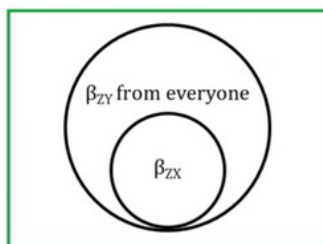
小练习



Section 1.3 One-Sample-MR and Two-Sample-MR



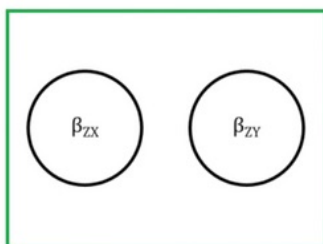
(A) One-sample MR



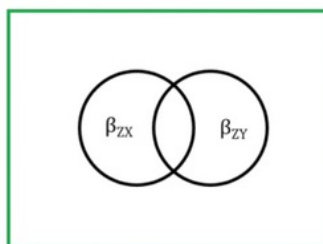
(B) One-sample MR with exposure measured in a random subgroup



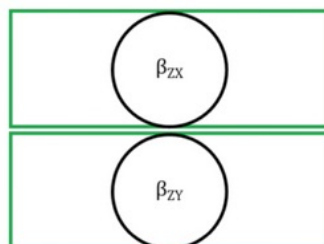
(C) One-sample MR with externally weighted PGRS



(D) Two-sample MR (ideal)



(E) Two-sample MR with overlapping samples



(F) Two-sample MR with different underlying populations

Section 1.4 One-Sample-MR: Wald-Ratio方法

注：比值法使用单一的工具变量，当有多个工具变量时，可以考虑

- (i) 每个变异体的比率法的因果估计值可以单独计算；
- (ii) 在等位基因评价法中把遗传变异合成成为一个进行估计

1.4.1 连续的结果，二分的IV

1.4.2 连续的结果，多分类或者连续的IV

1.4.3 二分的结果

1.4.4 比率估计的置信区间

Section 1.5 两阶段最小二乘法 / TSLS

Section 1.6 例子：BMI指数和吸烟的关系

SNP	Per allele change in BMI (SE)	Per allele change in smoking (SE)	Ratio estimate	95% CI (first-order)	95% CI (second-order)	95% CI (Fieller)
rs1558902	0.355 (0.011)	0.038 (0.032)	0.106	−0.073, 0.285	−0.073, 0.285	−0.073, 0.286
rs6567160	0.258 (0.013)	0.031 (0.038)	0.120	−0.166, 0.406	−0.166, 0.406	−0.167, 0.409
rs2820292	0.097 (0.011)	0.013 (0.032)	0.138	−0.514, 0.790	−0.514, 0.791	−0.525, 0.817
rs3101336	0.105 (0.011)	0.054 (0.033)	0.512	−0.093, 1.116	−0.102, 1.126	−0.093, 1.165

TABLE 4.1

Association of four SNPs with body mass index (BMI, kg/m² units) and smoking intensity (pack-years). Causal estimates are obtained from the ratio method, and 95% confidence intervals (CI) are calculated using three approaches: normal approximation with first-order standard errors, normal approximation with second-order standard errors, and Fieller’s theorem.

Section 1.7 回顾