Causal mediation
因果中介效应分析

北京师范大学
金承刚

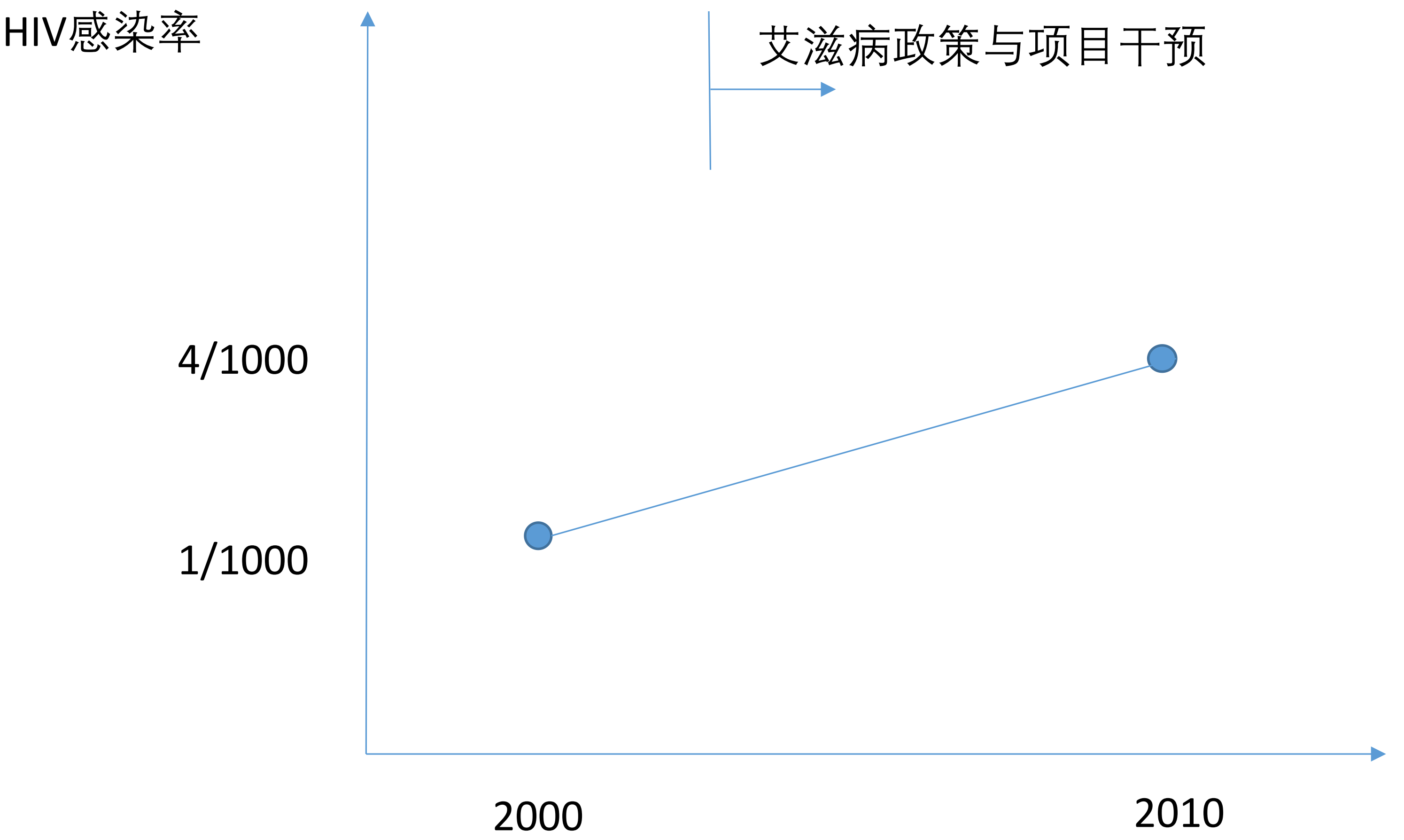
内容

- 中介效应
  - 因果推断
  - 作用机制或作用路径
- 传统中介效应的分析方法
  - 估计方法
  - 所受到的限制
  - 因果解释
- 因果中介分析的STATA语句
  - 前提假设
  - 效应的分解
    - .paramed
    - .med4way

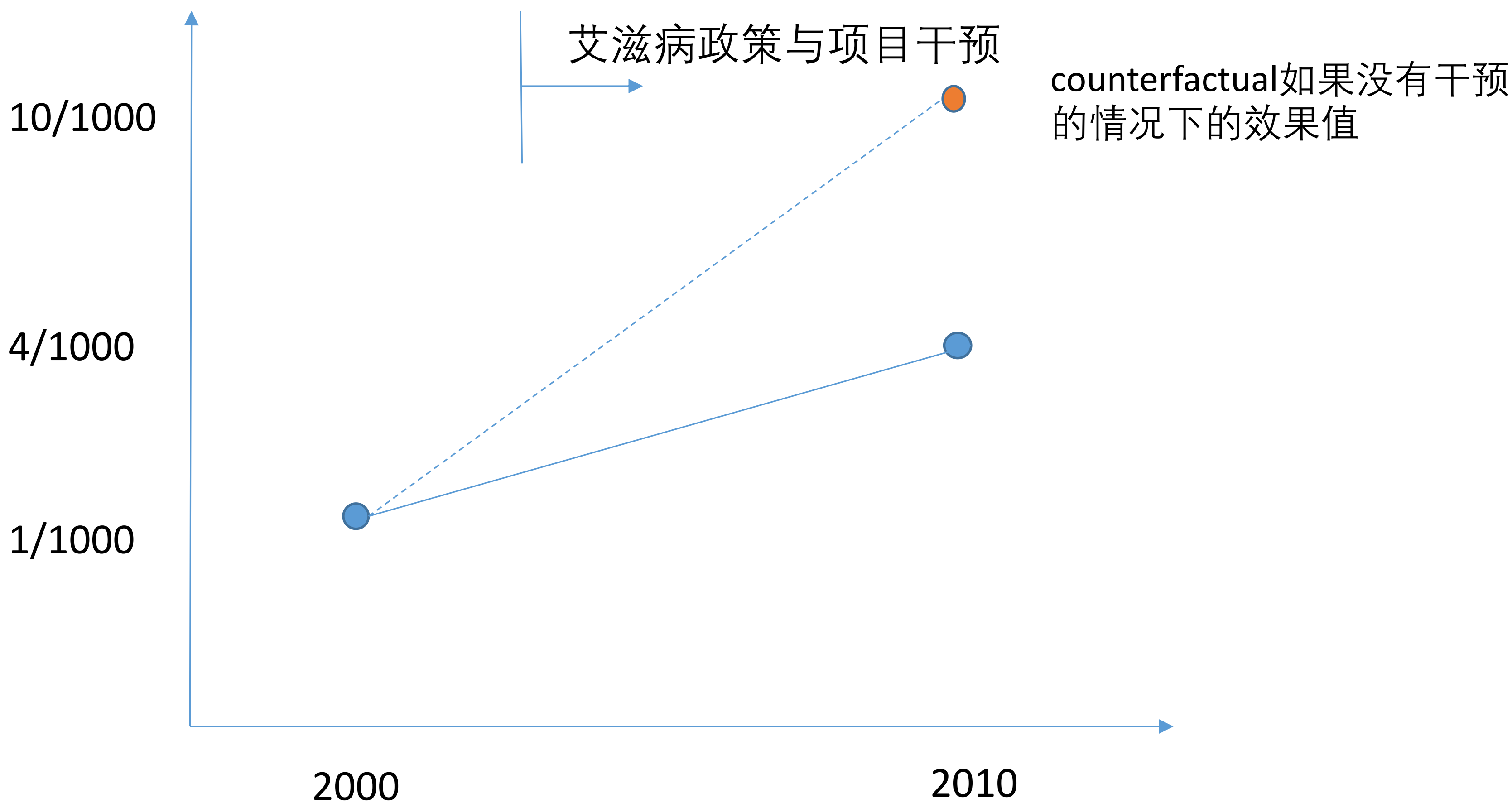
内容

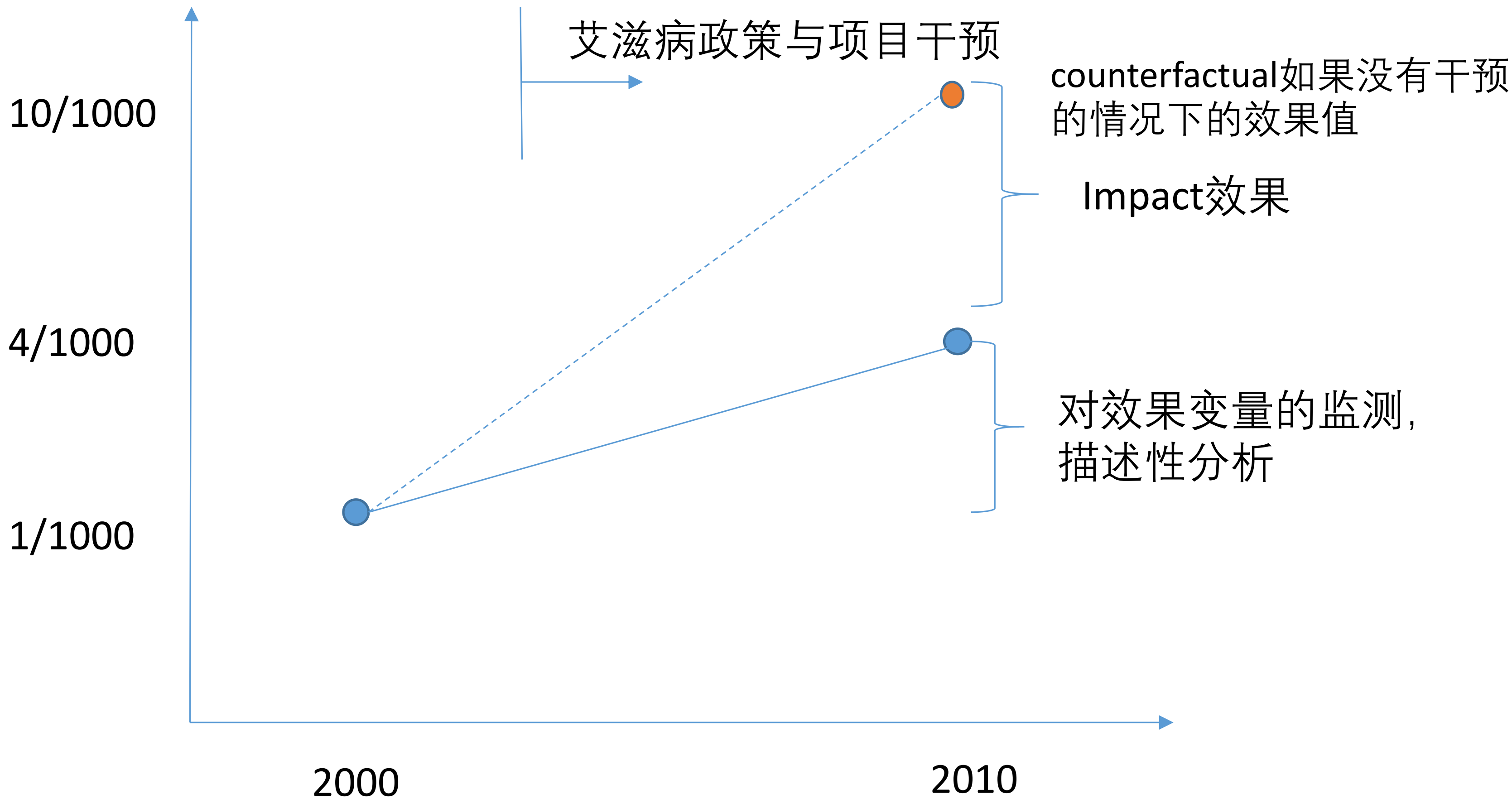
- 中介效应
  - 因果推断
  - 作用机制或作用路径
- 传统中介效应的分析方法
  - 估计方法
  - 所受到的限制
  - 因果解释
- 因果中介分析的STATA语句
  - 前提假设
  - 效应的分解
    - .paramed
    - .med4way

Counter-factual



前后比较：HIV病毒感染率由1/1000提高到4/1000,提高了3/1000
干预导致艾滋病病毒感染率提高，促进了流行


艾滋病政策与项目干预
counterfactual如果没有干预
的情况下的效果值
10/1000
4/1000
1/1000
2000
2010

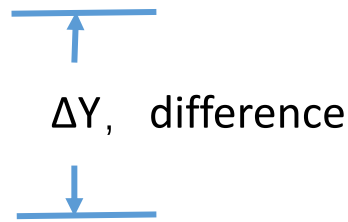

艾滋病政策与项目干预
counterfactual如果没有干预
的情况下的效果值
Impact效果
对效果变量的监测,
描述性分析
10/1000
4/1000
1/1000
2000
2010

效果的估计

- 对于具体的个体(individual)，无法得到干预的效果，因为总有一半的数据缺失
- 但是对于group来说，可以进行估计：
- E(Y1 – Y0)

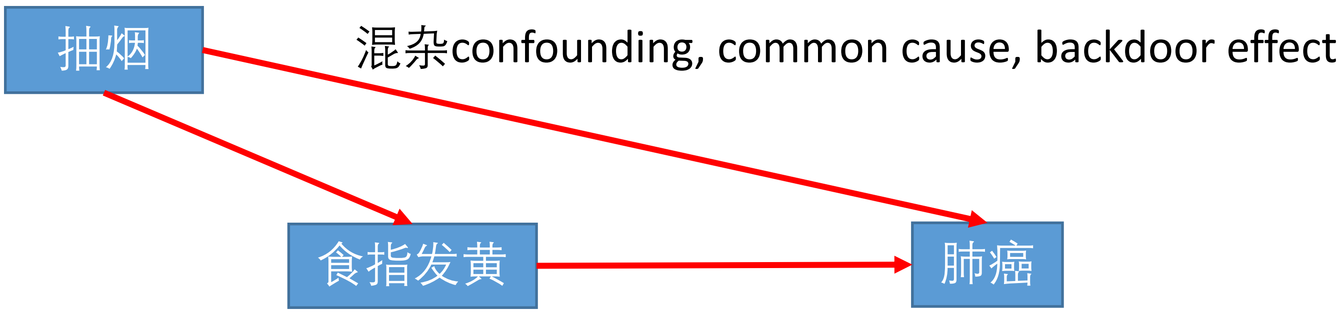
暴露效应的因果推断

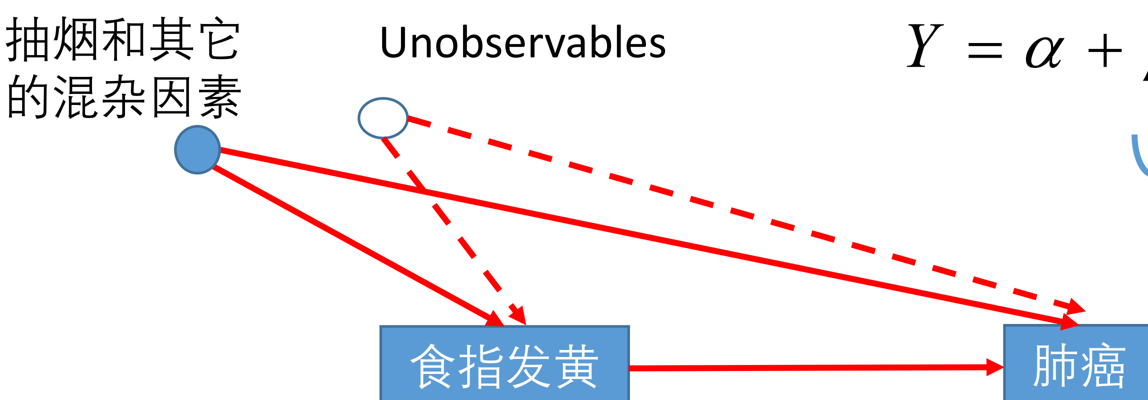
- 当暴露组和非暴露组的效果变量有一个差值(difference)



- 其它可能的解释(alternative explanations)?
  - 1、由混杂导致的（confounding）
  - 2、由于选择偏倚导致的(selection bias)
  - 3、由于抽样误差导致的(sampling error, sample variation, or chance)
  - 4、由于measurement bias导致的
- 把4个可能的解释排除后，才能形成有效的统计联系（因果推断）

暴露作用的效果:混杂的作用





unobservables

$$Y = \alpha + \underbrace{\beta smoke + \sum \gamma_j X_j}_{\text{observables}} + u$$

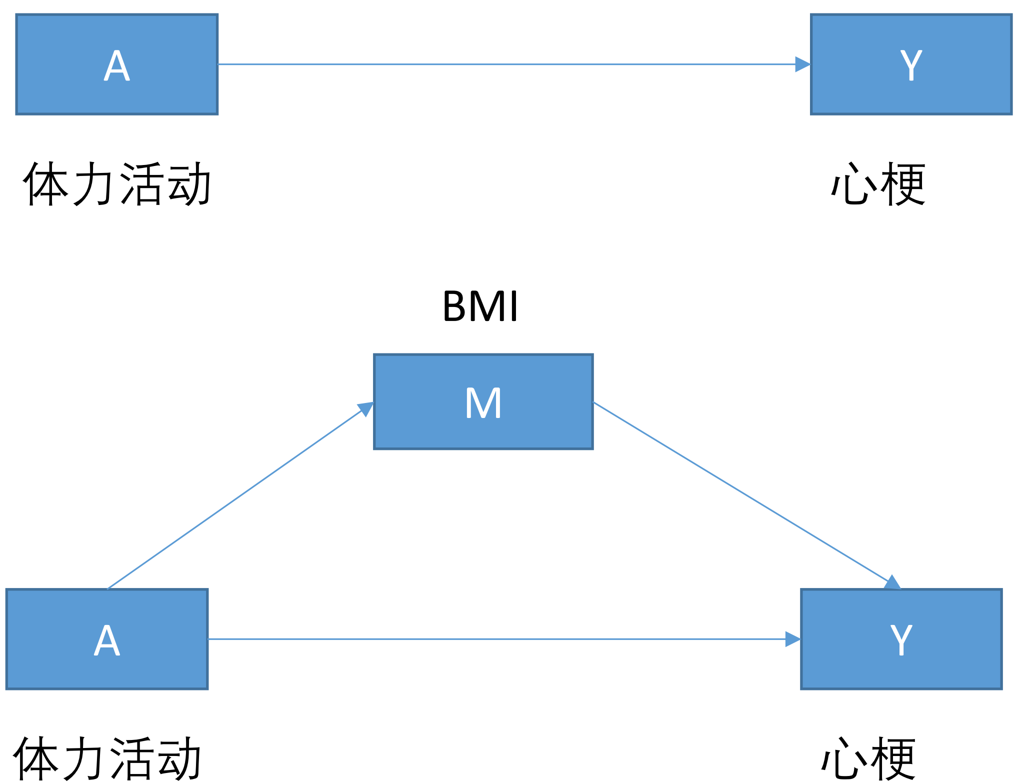
内容

- 中介效应
  - 因果推断
  - 作用机制或作用路径
- 传统中介效应的分析方法
  - 估计方法
  - 所受到的限制
  - 因果解释
- 因果中介分析的STATA语句
  - 前提假设
  - 效应的分解
    - .paramed
    - .med4way

暴露作用的效果

暴露的效果

- 我们想研究体力活动(physical activity)对心梗(MI)的作用
- 体力活动是如何所用到心梗发生的？作用的机理、路径？
- 体力活动导致肥胖下降，然后导致心梗的发生的减少？



社区干预控制高血压

干预组: T O

非随机对照组: O

$$Y = \alpha + 0.7 * \text{项目干预} + \sum \gamma_j X_j + u$$

$$Y = \alpha + 0.98 * \text{项目干预} + 0.6 * \text{服药依从性} + \sum \gamma_j X_j + u$$

Y: 血压是否控制,

1: 血压没控制, 0: 血压控制

项目干预(treat):

1:干预组, 0对照组

- 不放入服药依从性变量，则项目干预有效（OR=0.7， p<0.05）
- 放入服药依从性变量，则项目干预没效（OR=0.98， p=0.23）
- 进一步了解项目干预的活动和内容，发现项目的干预内容之一是提高患者的服药依从性

柳叶刀期刊发表的文章
中国武汉成年新冠肺患者临床过程和死亡相关风险因素—回顾性队列研究

北京友万信息科技有限公司
www.uone-tech.cn

Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study

Fei Zhou, Ting Yu*, Ronghui Du*, Guohui Fan*, Ying Liu*, Zhibo Liu*, Jie Xiang*, Yeming Wang, Bin Song, Xiaoying Gu, Lulu Guan, Yuan Wei, Hui Li, Xudong Wu, Jiuyang Xu, Shengjin Tu, Yi Zhang, Hua Chen, Bin Cao*

Summary

Background Since December, 2019, Wuhan, China, has experienced an outbreak of coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Epidemiological and clinical characteristics of patients with COVID-19 have been reported but risk factors for mortality and a detailed clinical course of illness, including viral shedding, have not been well described.

Methods In this retrospective, multicentre cohort study, we included all adult inpatients (≥18 years old) with laboratory-confirmed COVID-19 from Jinyintan Hospital and Wuhan Pulmonary Hospital (Wuhan, China) who had been discharged or had died by Jan 31, 2020. Demographic, clinical, treatment, and laboratory data, including serial samples for viral RNA detection, were extracted from electronic medical records and compared between survivors and non-survivors. We used univariable and multivariable logistic regression methods to explore the risk factors associated with in-hospital death.

当时的关注的问题之一是年龄对新冠肺死亡的作用

	Univariable OR (95% CI)	p value	Multivariable OR (95% CI)	p value
Demographics and clinical characteristics				
Age, years*	1·14 (1·09–1·18)	<0·0001	1·10 (1·03–1·17)	0·0043
Female sex (vs male)	0·61 (0·31–1·20)	0·15	··	··
Current smoker (vs non-smoker)	2·23 (0·65–7·63)	0·20	··	··
Comorbidity present (vs not present)				
Chronic obstructive lung disease	5·40 (0·96–30·40)	0·056	··	··
Coronary heart disease	21·40 (4·64–98·76)	<0·0001	2·14 (0·26–17·79)	0·48
Diabetes	2·85 (1·35–6·05)	0·0062	··	··
Hypertension	3·05 (1·57–5·92)	0·0010	··	··
Respiratory rate, breaths per min				
≤24	1 (ref)	··	··	··
>24	8·89 (4·34–18·19)	<0·0001	··	··
SOFA score 器官衰竭分值	6·14 (3·48–10·85)	<0·0001	5·65 (2·61–12·23)	<0·0001
qSOFA score	12·00 (5·06–28·43)	<0·0001	··	··
Laboratory findings				
White blood cell count, × 10⁹ per L				
<4	0·73 (0·26–2·10)	0·56	··	··
4–10	1 (ref)	··	··	··
>10	6·60 (3·02–14·41)	<0·0001	··	··
Lymphocyte count, × 10⁹ per L*	0·02 (0·01–0·08)	<0·0001	0·19 (0·02–1·62)	0·13
ALT, U/L				
≤40	1 (ref)	··	··	··
>40	2·87 (1·48–5·57)	0·0018	··	··
(Continued from previous column)				
Creatinine, μmol/L				
≤133	1 (ref)	··	··	··
>133	4·39 (1·01–19·06)	0·048	··	··
Lactate dehydrogenase, U/L				
≤245	1 (ref)	··	··	··
>245	45·43 (6·10–338·44)	0·0002	··	··
Creatine kinase, U/L				
≤185	1 (ref)	··	··	··
>185	2·56 (1·03–6·36)	0·043	··	··
High-sensitivity cardiac troponin I, pg/mL				
≤28	1 (ref)	··	··	··
>28	80·07 (10·34–620·36)	<0·0001	··	··
D-dimer, μg/mL				
≤0·5	1 (ref)	··	1 (ref)	··
> 0·5	1·96 (0·52–7·43)	0·32	2·14 (0·21–21·39)	0·52
> 1	20·04 (6·52–61·56)	<0·0001	18·42 (2·64–128·55)	0·0033
Prothrombin time, s				
<16	1 (ref)	··	··	··
≥16	4·62 (1·29–16·50)	0·019	··	··
Serum ferritin, μg/L				
≤300	1 (ref)	··	··	··
>300	9·10 (2·04–40·58)	0·0038	··	··
IL-6, pg/mL*	1·12 (1·03–1·23)	0·0080	··	··
Procalcitonin, ng/mL*	13·75 (1·81–104·40)	0·011	··	··

OR=odds ratio. SOFA=Sequential Organ Failure Assessment. qSOFA=Quick SOFA. ALT=alanine aminotransferase. IL-6=interleukin-6. *Per 1 unit increase.

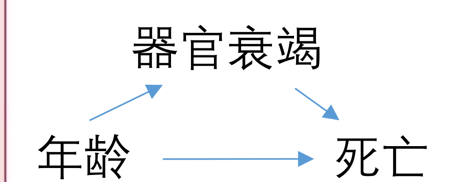
Table 3: **Risk factors associated with in-hospital death**

选择了认为5个重要的自变量放入模型:

1. 年龄
2. 冠心病
3. 器官衰竭分值 sofa
4. 淋巴细胞计数
5. D-dimer:高凝或纤融状态

Logistic回归
年龄是一个重要的研究的自变量

但是，患新冠肺，会导致器官衰竭，是年龄与死亡的causal path,应按mediator处理



内容

北京友万信息科技有限公司
www.uone-tech.cn

- 中介效应
  - 因果推断
  - 作用机制或作用路径
- 传统中介效应的分析方法
  - 估计方法
  - 所受到的限制
  - 因果解释
- 因果中介分析的STATA语句
  - 前提假设
  - 效应的分解
    - .paramed
    - .med4way

传统的中介效应分析

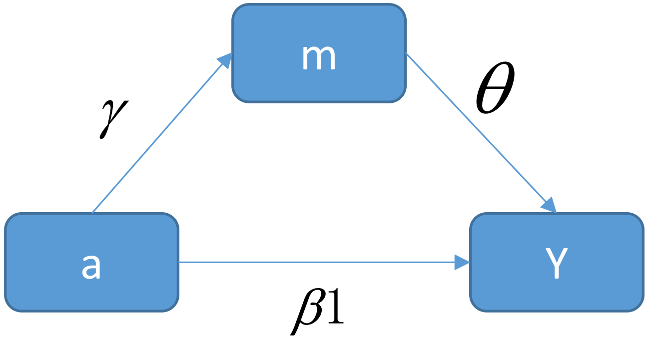
- Baron and Kenny

对效果变了的估计模型

$$E[Y \mid a,m] = \alpha_1 + \beta_1 a + \theta m$$

对中介变量的回归模型

$$E(M \mid a) = \alpha_3 + \gamma a$$



直接效应(direct effect)：β_1

间接效应(indirect effect, product method)：$\theta\gamma$

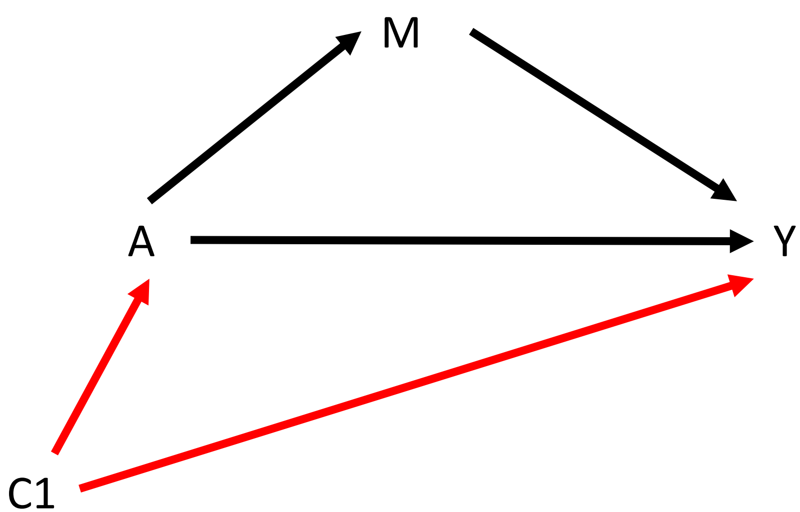
传统的中介效应

- 一般方法的适用条件
  - 线性模型
  - 没有exposure-mediator interaction
  - Causal interpretation?

传统的方法

北京友万信息科技有限公司
www.uone-tech.cn

- 传统的mediation的分析，仅仅控制exposure-outcome的混杂(C1)
- 即使暴露是随机分组的，或者exposure-outcome confounders都放入模型，但是，如果mediator-outcome confounders没有控制，则会导致偏倚



Arlinghaus A, Lombardi DA, Willetts JL, et al. A structural equation modeling approach to fatigue-related risk factors for occupational injury. Am J Epidemiol. 2012;176(7):597–607.

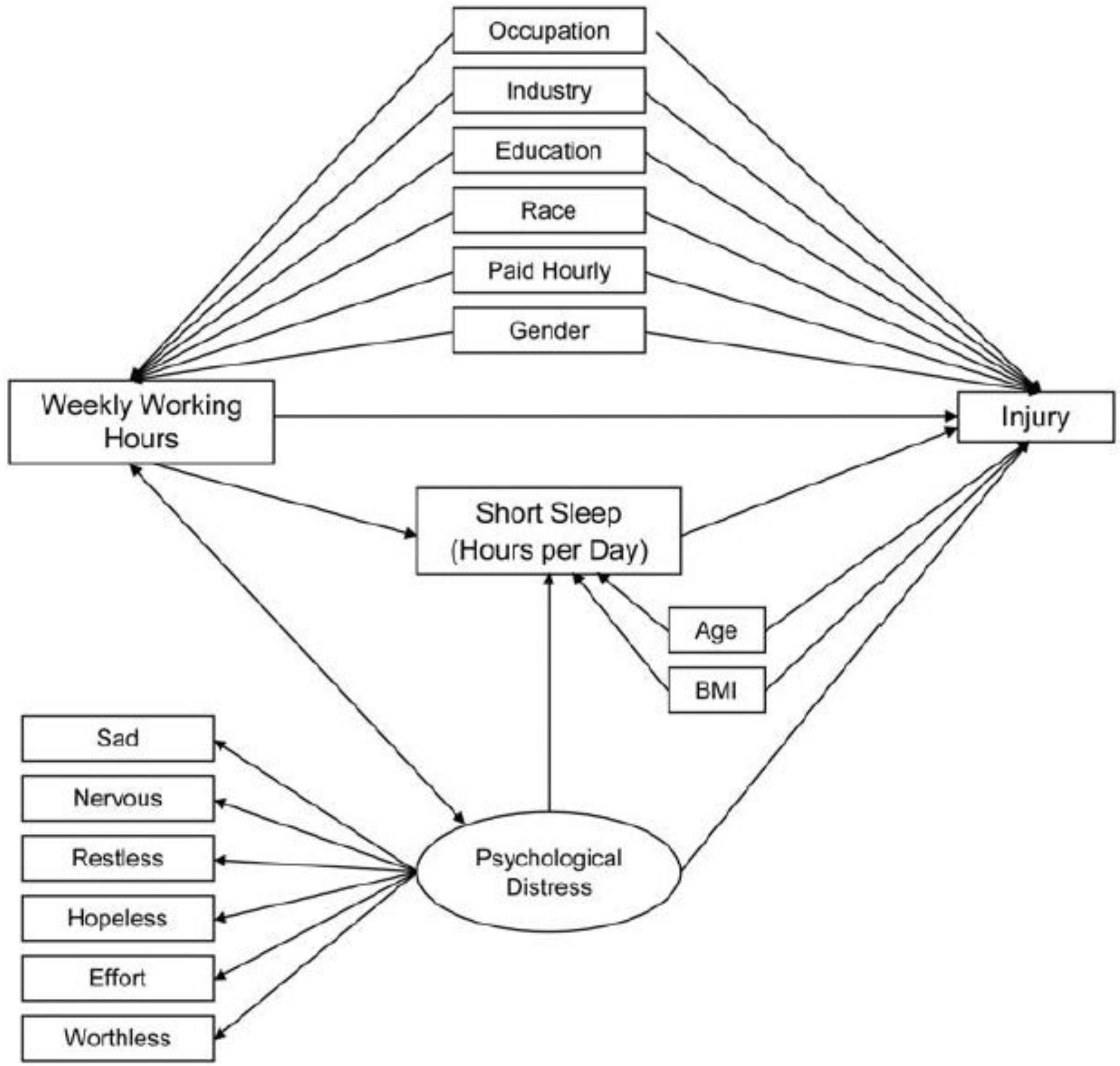


Figure 1. Structural equation model analyzed by Arlinghaus et al. (1). (BMI, body mass index).

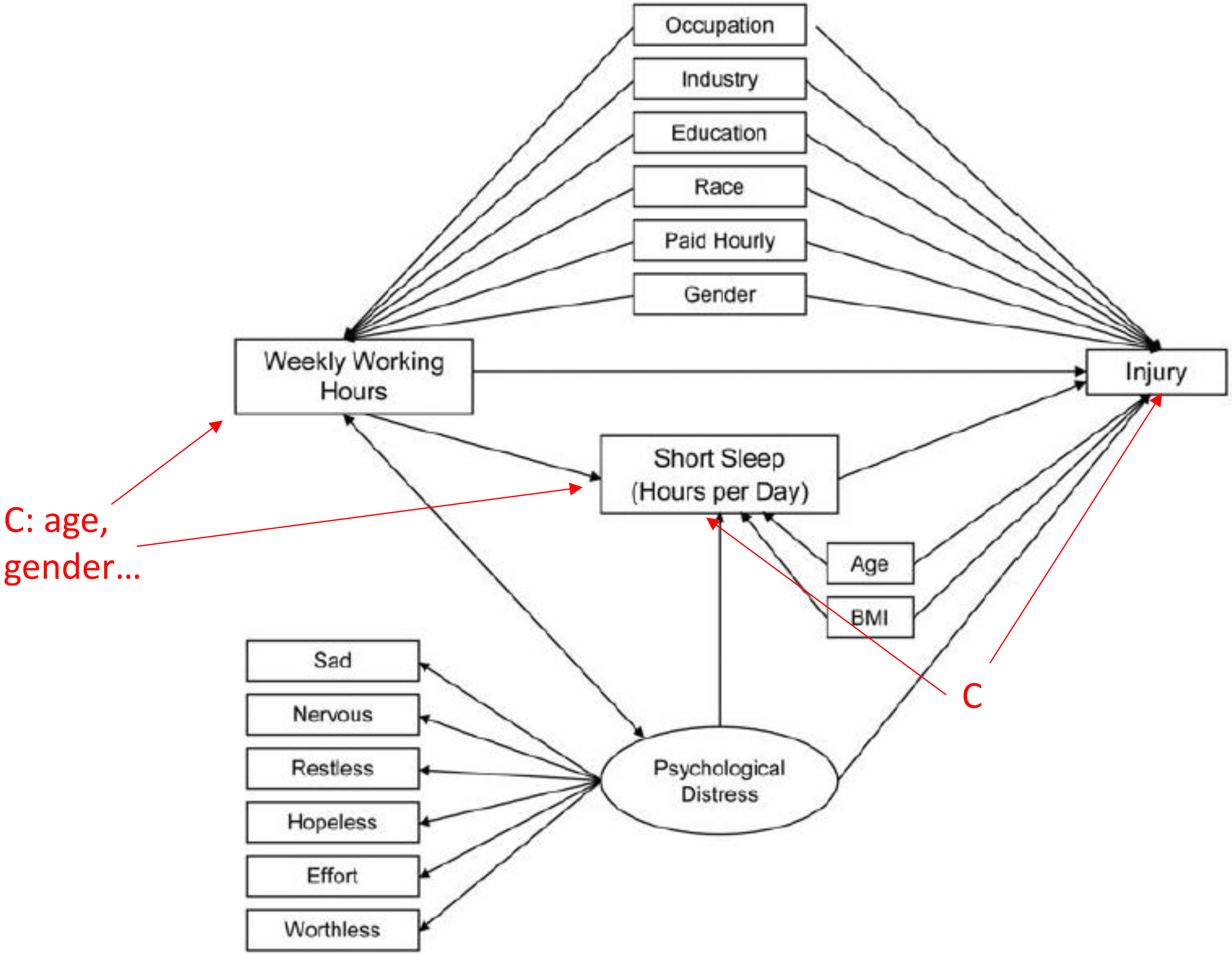


Figure 1. Structural equation model analyzed by Arlinghaus et al. (1). (BMI, body mass index).

内容

- 中介效应
  - 因果推断
  - 作用机制或作用路径
- 传统中介效应的分析方法
  - 估计方法
  - 所受到的限制
  - 因果解释?
- 因果中介分析的STATA语句
  - Counterfactual framework
  - 前提假设
  - 暴露和中介变量的交互
  - 效应的分解
    - .paramed
    - .med4way

传统方法

$$E(M \mid A=a, C=c) = \beta_0 + \beta_1 a + \beta_2 ' c$$

$$E(Y \mid A=a, M=m, C=c) = \theta_0 + \theta_1 a + \theta_2 m + \theta_4 ' c$$

Counterfactual Approach

当暴露因素和中介变量有交互时：

$$E(Y \mid A=a, M=m, C=c) = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4 ' c$$

当暴露水平从a*变化到a

Controled Derect Effect :

$CDE(m) = (\theta_1 + \theta_3 m)(a - a^*)$, 当没有交互时$\theta$3=0，$CDE = \theta 1(a - a^*)$

Natural Direct Effect :

$$NDC = (\theta_1 + \theta_3\beta_0 + \theta_3\beta_1 a^* + \theta_3\beta_2 ' c)(a - a^*)$$

Natural Indirect Effect

$$NIE = (\theta_2\beta_1 + \theta_3\beta_1 a)(a - a^*)$$

新的方法:causal mediation

- Robins和Greenland(1992), pearl(2001), Vanderweele(2009)
- 采取counterfactual方法
- 适用于更多的情况
- 敏感性分析（sensitivity analysis, hidden bias）

效果评价的两难：counter-factual

- 效果：
  - 针对第i个诊断为肺癌的患者，经过治疗活了5年(y_{1i})
  - 同样这个患者i，如果不治疗，他活了多长时间(y_{0i})
  - 针对这个人的项目干预效果= $Y_{1i} - Y_{0i}$
- 但是，但是，针对同一个患者，要么治疗、要么不治疗。因此，无法知道如果不干预的发病率。
- 因此，Y0缺失(missing)
- 评价设计和分析策略主要是如何解决counterfactual（解决missing data）

北京友万信息科技有限公司
www.uone-tech.cn

Counterfactual approach

- A:暴露、或处理变量exposure or treatment
- Y：效果变量、结局变量
- M：中介变量
- 假设A为基因型，Y为肺癌，M中介变量为抽烟程度
- Ya:如果和事实(factual)相反，暴露因素设为a，因此，Ya则为counterfactual or potential outcome
- 假设A是二分类的，那么每一个个体就有两个potential outcome
  - Y0：和事实相反，如果没有得到暴露(A=0)的效果
  - Y1：和事实相反，如果得到暴露(A=1)的效果
- 例如，如果每个个体具有某个基因，是否患肺癌Y1。如果每个个体没有该基因型，是否患肺癌Y0
- 那么，针对个体的暴露效果=Y1-Y0
- 但是，对于个体，总有一半的数据缺失，无法得到效果

- 而是利用组别数据进行估计平均暴露的效果
- 如果两组是随机分组，则Causal effect=E(Y1-Y0)
- 如果两组不是随机分组，当exposure-outcomede混杂变量控制的较好时：
  - 则暴露的效果=E(Y1 – Y0|c), 为conditional causal effect

Mediation effect

- 对于中介变量mediator，同样有counter-factual or potential outcome
- 当被暴露时，即A=a, M的取值为Ma
- 假设暴露为二分类：暴露和非暴露，则 M:
  - 和事实相反，如果暴露时，M1
  - 和事实想法，如果没有暴露，则为M0
- 同样，每个个体都有两个值M1, M0
- 例如具有该基因的，平均抽烟的支数/天；如果没有该基因的，平均抽烟的支数/天

同时暴露和中介顺序一起作用时

- Yam为当A=a,M=m时的counterfactual(potential outcome)值
- 例如：当具有该基因，抽烟=10支，则效果的counterfactual为：Ya=1,m=10
- 而Ya=1, m=20则为如果暴露，抽烟为20支的效果值
- 因此，会有多个多个取值，但是，我们只能观察到一个值

效果的术语

- 采用counterfactual方法， Robin, Greenland(1992) and Pearl(2001)定义了
  - Controlled Direct Effect(CDE)
  - Natural Direct effect (NDE)
  - Natural Indirect Effect (NIE)

CDE (Controlled Direct Effect)

- 如果M=m,比较A=1和A=0的暴露效果：
- Y1m-Y0m
- 把M固定地设定的m值,即没有通过中介路径的作用
- 例如，比较基因暴露的效果，则将抽烟支数设定为具体的m
- 但是，CDE的值会随着m的取值不同而不同
- 当暴露效果=0,
  - 把抽烟支数固定为0，M=0，CDE(0)=Y1,0 – Y0,0 = 0
  - 当把抽烟支数固定到20时，CDE(20)=Y1,20 – Y0,20 =1
- 可以估计组间数据，即E(Y1m – Y0m)
- Conditional average CDE(m)=E(Y1m – Y0m|c)

Natural Direct Effect (NDE)

- NDE和CDE不同，它
- 如果暴露A=0时，M为其自然取值，此时比较暴露和非暴露二值之间Y的差异
- NDE=$Y_{1M0} - Y_{0M0}$
- 作为群体数据，则可以得到平均期望值
  - $E(Y_{1M0} - Y_{0M0})$
  - 当混杂控制时，则$E(Y_{1M0} - Y_{0M0} | c)$

NIE(Natural Indirect Effect)

- 如果将中介变量M分别设定为M1、M0，并且将暴露设定为A=1，然后对效果变量进行比较
- 对于个体(individual)：$Y_{1M1} - Y_{1M0}$
- 对于每组（group）数据来说:

$$E=(Y_{1M1} - Y_{1M0})$$

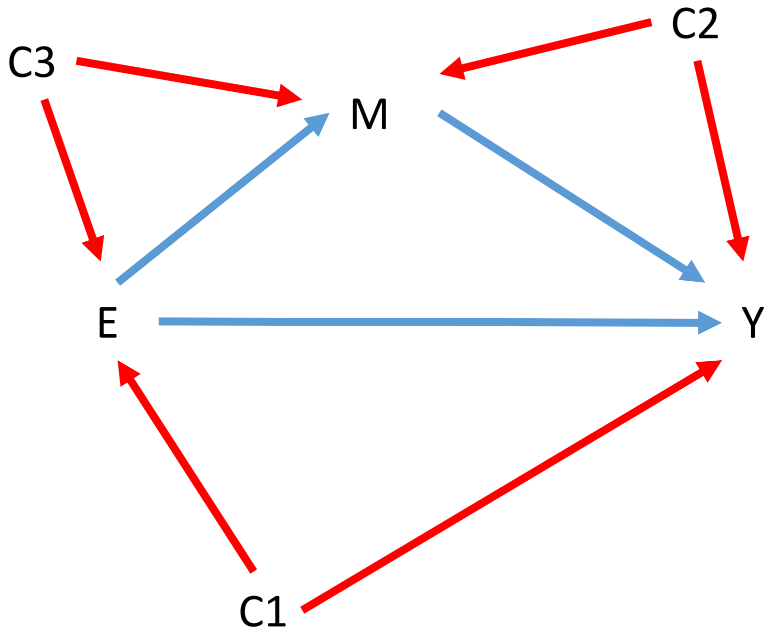
- 非随机分组时，则控制足够的混杂，得到：

$$E=(Y_{1M1} - Y_{1M0}|c)$$

Mediation效应估计的混杂假设 Confounding assumptions

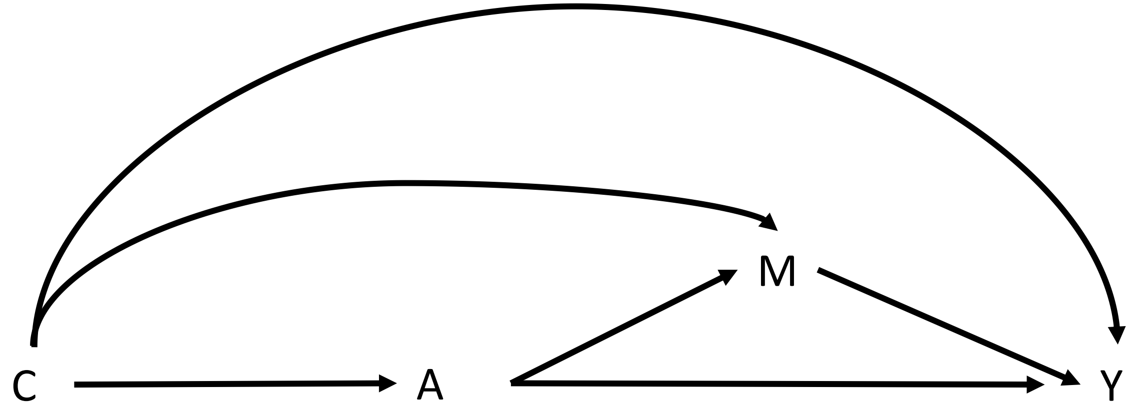
- Assumption A1
  - exposure-outcome 没有混杂(confounder C1)
- Assumption A2
  - mediator-outcome没有混杂(confounder C2)
- Assumption A3
  - Exposure-mediator没有混杂(confounder C3)
- Assumption A4
  - Mediator-outcome的confounder不受暴露因素和exposure-outcome的混杂的作用

混杂的控制

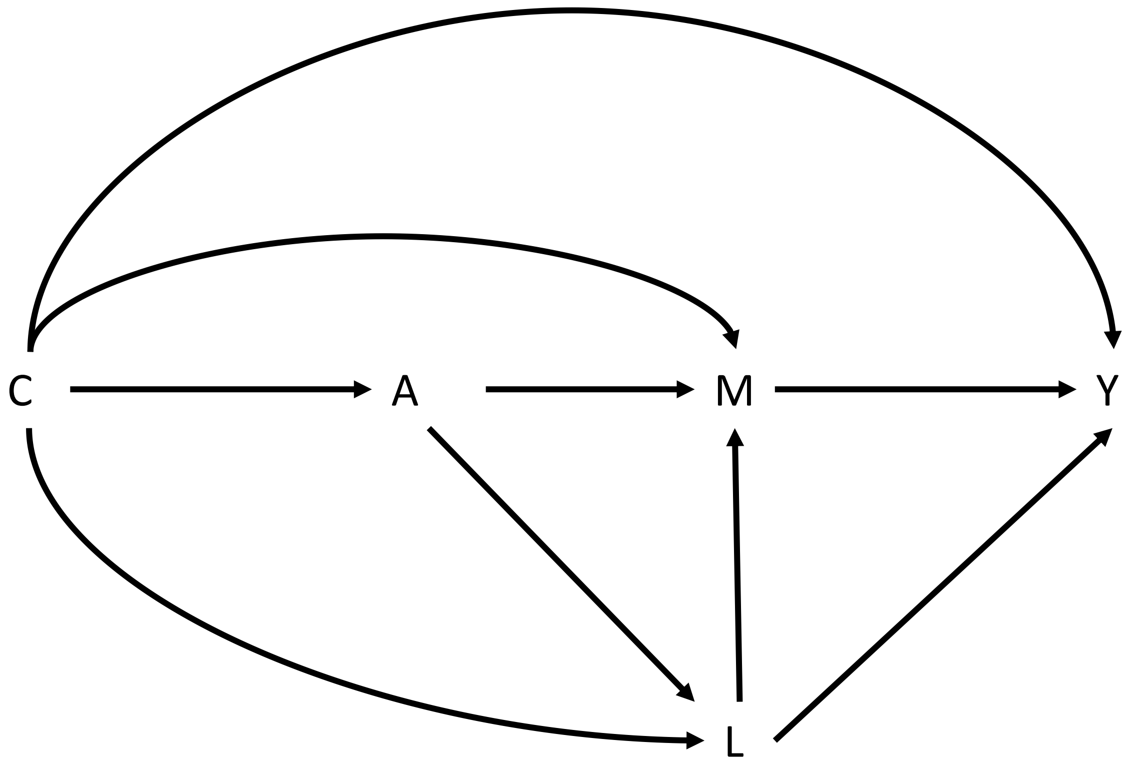


简化的DAG图

C： c1: exposure-outcome confounders
c2: mediator-outcome confounders
c3:exposure-mediator confounders



没有受暴露因素导致的Mediator-outcome confounder (no mediator-outcome confounder caused by exposure)



暴露因素和中介变量的交互

- 传统的方法没有考虑到暴露因素和中介变量的交互。
  - 有可能是中介作用：体力活动对心梗发生的作用，体力活动减少肥胖(BMI)而作用于心梗的发生
  - 也可能体力活动和肥胖交互而产生作用

Stata分析命令： .paramed

- .ssc install paramed

- paramed allows continuous, binary or count outcomes, and continuous or binary mediators, and
- requires the user to specify an appropriate form for the regression models.

	Outcome	Exposure	Mediator
变量类型	连续变量 二分类 计数变量(count)	连续变量 二分类	连续变量 二分类

.help paramed

Title

paramed -- causal mediation analysis using parametric regression models

Syntax

paramed varname, avar(varname) mvar(varname) a0(real) a1(real) m(real) yreg(string) mreg(string) [cvars(varlist) nointeraction casecontrol fulloutput c(numlist) bootstrap reps(integer 1000) level(cilevel) seed(passthru)]

--varname – 效果变量the outcome variable.

--avar(varname) – 暴露变量 the treatment (exposure) variable.

--mvar(varname) – 中介变量mediator variable.

--a0(real) - this specifies the natural level of the treatment (exposure).

--a1(real) - this specifies the alternative treatment (exposure) level.

--m(real) – 具体的中介变量取值水平this specifies the level of mediator at which the controlled direct effect is to be estimated

--yreg(string) – 效果模型regression model to be fitted for the outcome variable. This can be either linear, logistic, loglinear, Poisson or Negative binomial.

-- mreg(string) – 中介模型regression model to be fitted for the mediator variable. This can be either linear or logistic.

. *1 y:连续变量 m:连续变量, 暴露：两分类，无交互

.paramed y_cont, avar(treat) mvar(m_cont) cvars(var1 var2) ///
a0(0) a1(1) m(1) yreg(linear) mreg(linear) nointer

y_cont	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
treat	-23.97229	.4352565	-55.08	0.000	-24.82641	-23.11816
m_cont	-2.643668	.1122969	-23.54	0.000	-2.864034	-2.423302
var1	.9816636	.0731527	13.42	0.000	.8381124	1.125215
var2	.0722772	.1537972	0.47	0.638	-.2295268	.3740812
_cons	13.27515	.9656767	13.75	0.000	11.38015	15.17014

m_cont	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
treat	.0485887	.1228043	0.40	0.692	-.1923961	.2895736
var1	-.0349786	.0206113	-1.70	0.090	-.0754252	.0054679
var2	.9860252	.0301176	32.74	0.000	.9269239	1.045126
_cons	.453117	.2721012	1.67	0.096	-.0808404	.9870743

	Estimate	Std Err	P>|z|	[95% Conf	Interval]
cde	-23.972289	.43525651	0.000	-24.825391	-23.119186
nie	-.12845254	.32469969	0.692	-.76486392	.50795885
te	-24.100741	.54297198	0.000	-25.164966	-23.036516

cde:controlled direct effect, nie:natural indirect effect, te:total effect

* 2 Y:连续变量， m:两分类变量，暴露：两分类，有交互, bootstrap

北京友万信息科技有限公司
www.uone-tech.cn

```
.paramed y_cont, avar(treat) mvar(m_bin) cvars(var1 var2) ///
              a0(0) a1(1) m(1) yreg(linear) mreg(logistic) boot
```

y_cont	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
treat	-12.90945	.5110227	-25.26	0.000	-13.91226	-11.90664
m_bin	3.292577	.5693594	5.78	0.000	2.175292	4.409861
_treat_X_m_bin	-22.23864	.7217306	-30.81	0.000	-23.65493	-20.82235
var1	.9897736	.0606257	16.33	0.000	.8708045	1.108743
var2	-1.412419	.1084454	-13.02	0.000	-1.625228	-1.199611
_cons	5.556967	.8206041	6.77	0.000	3.946651	7.167282

Logistic regression

m_bin	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
treat	.0681243	.1576183	0.43	0.666	-.2408019	.3770504
var1	-.0100659	.0263047	-0.38	0.702	-.0616221	.0414903
var2	.8696374	.0571742	15.21	0.000	.7575781	.9816968
_cons	-4.242807	.4044849	-10.49	0.000	-5.035583	-3.450031

	Estimate	Std Err	P>|z|	[95% Conf	Interval]
cde	-35.148087	.50975379	0.000	-36.147205	-34.14897
nde	-23.860326	.71321343	0.000	-25.258225	-22.462428
nie	-.32263925	.74640509	0.666	-1.7855932	1.1403147
mte	-24.182965	4.7226287	0.000	-33.439318	-14.926613

4 ways 的整合 unification of mediation and interaction

- 没有中介作用，仅有暴露作用
- 当中介的暴露为0，中介仅有交互作用
- 有与中介的交互作用
- 完全的总结作用

Stata分析命令： med4way

- net install med4way, from("https://raw.githubusercontent.com/anddis/med4way/master/") replace

*1. Binary outcome, binary mediator

```
. med4way y_bin treat m_bin cvar1 cvar2 cvar3, yreg(logistic) mreg(logistic) ///
>                    a0(0) a1(1) m(0) c(0 0 .)
```

-> Model for the outcome

y_bin	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interva
treat	-.2944844	.289416	-1.02	0.309	-.8617292	.2727
m_bin	1.046198	.2400432	4.36	0.000	.5757223	1.516
_treatXm_bin_000	.9821457	.3321434	2.96	0.003	.3311566	1.633
cvar1	.1219094	.1451617	0.84	0.401	-.1626023	.4064
cvar2	.1908434	.158203	1.21	0.228	-.1192287	.5009
cvar3	.0451507	.0065279	6.92	0.000	.0323562	.0579
_cons	-3.823604	.4667596	-8.19	0.000	-4.738436	-2.908

-> Model for the mediator

m_bin	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
treat	.2400897	.1477509	1.62	0.104	-.0494968	.5296762
cvar1	.3107653	.1491474	2.08	0.037	.0184418	.6030888
cvar2	-.5623266	.154374	-3.64	0.000	-.8648942	-.259759
cvar3	.0335621	.0065424	5.13	0.000	.0207393	.0463849
_cons	-1.188081	.4246957	-2.80	0.005	-2.02047	-.3556931

-> 4-way decomposition: delta method

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
tereri	.933078	.2971722	3.14	0.002	.3506311	1.515525
ereri_cde	-.1102558	.1088553	-1.01	0.311	-.3236083	.1030967
ereri_intref	.944044	.2872328	3.29	0.001	.3810781	1.50701
ereri_intmed	.062001	.0437227	1.42	0.156	-.0236938	.1476959
ereri_pie	.0372887	.0244705	1.52	0.128	-.0106725	.0852499

tereri=total excess relative risk; ereri_cde=excess relative risk due to controlled direct effect; ereri_intref=excess relative risk due to reference interaction; ereri_intmed=excess relative risk due to mediated interaction; ereri_pie=excess relative risk due to pure indirect effect.

3 Continuous outcome, continuous mediator

```
. sum

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
   SubjectID |        300       150.5    86.74676          1        300
     FamSize |        300    3.803333    1.620724          0          8
   SocStatus |        300    25.07333    3.625618         14         33
   Encourage |        300    34.41667    1.906613         29         39
  Motivation |        300       38.47    1.923912         33         43
-------------+---------------------------------------------------------
  CogPerform |        300    90.77333    15.36997         49        129
```

北京友万信息科技有限公司
www.uone-tech.cn

.med4way CogPerform Encourage Motivation SocStatus FamSize, ///
yreg(linear) mreg(linear) a0(33.91667) a1(34.91667) m(38.47)

-> Model for the outcome

CogPerform	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interva
mu						
Encourage	.9679754	.102095	9.48	0.000	.7678729	1.1680
Motivation	.9973695	.0893475	11.16	0.000	.8222517	1.1724
_EncourageXMotivation_000	.0834855	.0022753	36.69	0.000	.079026	.0879
SocStatus	.0121183	.017629	0.69	0.492	-.0224339	.04667
FamSize	-.0247094	.0126186	-1.96	0.050	-.0494414	.00002
_cons	-91.95793	3.082423	-29.83	0.000	-97.99937	-85.916
sigma2						
_cons	.0380791	.0031091	12.25	0.000	.0319853	.04417

-> Model for the mediator

Motivation	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
mu						
Encourage	.6878576	.0360386	19.09	0.000	.6172232	.7584921
SocStatus	.1562103	.0180483	8.66	0.000	.1208363	.1915843
FamSize	-.1103726	.0129439	-8.53	0.000	-.1357422	-.085003
_cons	11.2993	.8411747	13.43	0.000	9.650632	12.94798
sigma2						
_cons	.0498804	.0040727	12.25	0.000	.041898	.0578628

-> 4-way decomposition: delta method

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
te	6.842126	.1430073	47.84	0.000	6.561837	7.122415
cde	4.179662	.0469639	89.00	0.000	4.087614	4.271709
intref	-.0287129	.0020085	-14.30	0.000	-.0326496	-.0247762
intmed	.0574261	.0033914	16.93	0.000	.0507791	.0640732
pie	2.633751	.1423044	18.51	0.000	2.35484	2.912663

Title

med4way — 4-way decomposition using parametric regression models

Syntax

med4way [*yvar*] *avar* *mvar* [*cvars*] [*if*] [*in*], **a0(**#**)** **a1(**#**)** **m(**#**)** **yreg(***string***)** **mreg(***string***)** [*options*]

yvar is the variable name for the outcome. Note that *yvar* must not be specified when the model for the outcome is an Accelerated Failure Time or a Cox proportional hazards model. In these cases, you must stset your data before r ed4way.

avar is the variable name for the exposure. If binary, it must be coded as 0/1.

mvar is the variable name for the mediator. If binary, it must be coded as 0/1.

cvars are the variable names for the covariates to be included in the model for the outcome and for the mediator.

options	Description
* **a0(**#**)**	referent exposure level
* **a1(**#**)**	actual exposure level
* **m(**#**)**	level of the mediator at which to compute the 4-way decomposition
* **yreg(***string***)**	form of the regression model for the outcome
* **mreg(***string***)**	form of the regression model for the mediator
c(*string***)**	values of the covariates *cvars* at which to compute the 4-way decomposition
yregoptions(*string***)**	pass options to the regression model for the outcome
mregoptions(*string***)**	pass options to the regression model for the mediator
casecontrol	treat the data as coming from a case-control study
fulloutput	compute also the proportion of the Total Effect attributable to its 4 components and the proportions attributable to mediation, interaction, and either mediation or interaction or both
nodeltamethod	suppress the calculation of the standard errors using the delta method

```
.      tabstat Encourage Motivation
   stats |   Encour~e   Motiva~n
---------+--------------------
    mean |   34.41667      38.47
------------------------------

. * Linear regression model for the outcome; Linear regression model for
. * the mediator; Delta method standard errors.
. * Note: a0=mean(Encourage)-0.5; m=mean(Motivation)

.med4way CogPerform Encourage Motivation SocStatus FamSize, ///
            yreg(linear) mreg(linear) a0(33.91667) a1(34.91667) m(38.47)
-> Summary
   Outcome    (yvar):  CogPerform
   Exposure   (avar):  Encourage
   Mediator   (mvar):  Motivation
   Covariates (cvars): SocStatus FamSize

   Model for the outcome  (yreg): linear
   Model for the mediator (mreg): linear

   Referent exposure level (a0):              33.91667
   Actual exposure level   (a1):              34.91667
   Mediator level for the decomposition (m): 38.47
   Fixed values of the covariates (c):        25.07 3.803
```

-> Model for the outcome

CogPerform	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
mu						
Encourage	.9679754	.102095	9.48	0.000	.7678729	1.168078
Motivation	.9973695	.0893475	11.16	0.000	.8222517	1.172487
_EncourageXMotivation_000	.0834855	.0022753	36.69	0.000	.079026	.087945
SocStatus	.0121183	.017629	0.69	0.492	-.0224339	.0466705
FamSize	-.0247094	.0126186	-1.96	0.050	-.0494414	.0000225
_cons	-91.95793	3.082423	-29.83	0.000	-97.99937	-85.91649
sigma2						
_cons	.0380791	.0031091	12.25	0.000	.0319853	.0441729

-> Model for the mediator

Motivation	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
mu						
Encourage	.6878576	.0360386	19.09	0.000	.6172232	.7584921
SocStatus	.1562103	.0180483	8.66	0.000	.1208363	.1915843
FamSize	-.1103726	.0129439	-8.53	0.000	-.1357422	-.085003
_cons	11.2993	.8411747	13.43	0.000	9.650632	12.94798
sigma2						
_cons	.0498804	.0040727	12.25	0.000	.041898	.0578628

-> 4-way decomposition: delta method

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
te	6.842126	.1430073	47.84	0.000	6.561837	7.122415
cde	4.179662	.0469639	89.00	0.000	4.087614	4.271709
intref	-.0287129	.0020085	-14.30	0.000	-.0326496	-.0247762
intmed	.0574261	.0033914	16.93	0.000	.0507791	.0640732
pie	2.633751	.1423044	18.51	0.000	2.35484	2.912663

te=total effect; cde=controlled direct effect; intref=reference interaction; intmed=mediated interaction; pie=pure indirect effect.