

# 有向无环图:语言、规则及应用

郑英杰 赵耐青

200032 上海,复旦大学公共卫生学院卫生微生物学教研室(郑英杰),生物统计学教研室(赵耐青); 200032 上海,复旦大学公共卫生学院教育部公共卫生安全重点实验室(郑英杰); 200032 上海,复旦大学卫生部卫生技术评估重点实验室(郑英杰)

通信作者:郑英杰, Email: yjzheng@shmu.edu.cn; 赵耐青, Email: nqzhao@shmu.edu.cn

DOI: 10.3760/cma.j.issn.0254-6450.2017.08.029

**【摘要】** 几乎所有的科学研究都在探索因果关系,有向无环图(DAGs)是因果关系研究的图形工具。本文系统地介绍了DAGs的图形语言、基本规则和干扰规则,及其在识别研究问题、理解和实施研究设计、指导数据分析、偏倚分类等方面的应用。DAGs对因果关系的研究具有重要的指导价值。

**【关键词】** 有向无环图; 因果关系; 研究设计; 偏倚

**基金项目:** 国家自然科学基金(81373065)

**Directed acyclic graphs: languages, rules and applications** Zheng Yingjie, Zhao Naiqing

Department of Public Health Microbiology (Zheng YJ), Department of Biostatistics (Zhao NQ), School of Public Health, Fudan University, Shanghai 200032, China; Key Laboratory of Public Health Safety, Ministry of Education, Fudan University, Shanghai 200032, China (Zheng YJ); Key Laboratory of Health Technology Assessment, Ministry of Health, Fudan University, Shanghai 200032, China (Zheng YJ)

Corresponding authors: Zheng Yingjie, Email: yjzheng@shmu.edu.cn; Zhao Naiqing, Email: nqzhao@shmu.edu.cn

**【Abstract】** Nearly all scientific studies explore causality, which will be met by directed acyclic graphs (DAGs). This paper systematically introduces graphic language, basic and interference rules of DAGs, and their applications into identifying research questions, understanding and undertaking research designs, guiding data analysis, classifying biases, etc. DAGs play key roles in causality studies.

**【Key words】** Directed acyclic graphs; Causality; Research designs; Bias

**Fund program:** National Natural Science Foundation of China (81373065)

几乎所有的科学研究都在探索因果关系,如药物/疫苗的效果、政策干预的有效性、引起疾病/促进健康等事件的病因等等。客观世界事物的发生发展有其固有的规律,因果关系研究是科学工作者对沉默世界的主动探索。因此,正确的因果关系推断必须建立在对客观世界病因框架的整体认识上。

客观世界纷繁复杂决定着病因研究的困难。基于有效的背景信息,如何科学地组织并呈现出现实世界中各种因果效应之间的复杂关系,成为病因研究需要解决的优先问题。有向无环图(directed acyclic graphs, DAGs)成为解决这一问题的重要工具,即基于数据的产生过程, DAGs以合乎逻辑的图形(变量、符号以及两者组合构成的路径)及其规则来揭示其所反映的概率关系<sup>[1-3]</sup>,从而刻画客观世界定性的因果关系。

## 一、DAGs 图形语言

1. 符号: DAGs 的符号主要为单箭头“→”,少数有问号“?”、方框“□”、括号“( )”、虚线“---”和双箭头虚线“---”等。

(1)单箭头“→”和问号“?”:因果效应如何描述?简单

地,对病因A及其效应B,应用单箭头“→”表示为“A→B”(图1):箭头方向反映时序关系,箭头两端变量反映了它们之间存在的因果关系。“A→B”代表变量A发生时间在前,变量B发生时间在后;变量A是B的病因,变量B是A的效应。

如果A→B的效应是我们拟估计或研究的,那么在DAGs中,可在单箭头上添加问号“?”,即“A<sup>?</sup>→B”,代表效应的不确定性。

(2)方框“□”和括号“( )”:DAGs中的变量被方框“□”所包围的结构,称之为带框变量,如□C,代表着以该变量C限制

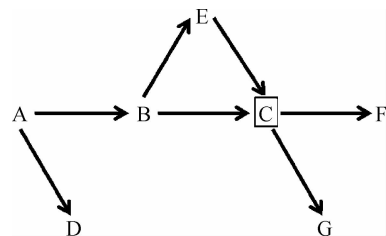


图1 DAGs示意图

于某一固定值(图1),如数据分析只对纳入研究的人群进行,等同于对C进行调整。这种调整对变量C所在的路径及其反映的变量之间的关系将可能产生影响。有时,用括号“()”来代替方框“□”。

(3)虚线“---”:用于代表着在对变量进行调整时,导致在DAGs中的变量之间将产生新的关联,可能影响变量之间的效应估计。虚线主要用于调整碰撞变量时产生的新关联<sup>[4]</sup>:如图1中,因变量C为变量B和变量E的共同效应,调整C将使得B和E之间在原有B→E的基础上,产生新的关联B---E。因这种关联并非客观世界所固有的,而是因调整而产生,也称B---E为触发路径;通常在提供设计和分析思路时用于简化DAGs。

(4)双箭头虚线“⇌”:代表着等同于图2-6的混杂结构:当有一未测量的因素(U)共同影响着拟研究的两个因素时,如吸烟(S)与饮酒(D),此时可省略U,而用双箭头虚线直接连接S和D来表示,即S⇌D。实际使用时,多采用图2-6的表示法,较为清晰。

2. 变量(节点)与路径:DAGs中的变量,也称为节点(Vertex或Node),如图1中的A-G,其中□为带框变量。路径是一条由2个或以上节点及其相应的连接符号组成的、不分叉的连续线路。如图1所示的路径可有:A→D, A→B→□←E, A→B→E→□→G等。DAGs由一条或多条路径组成,不同路径之间可共享某些变量。

(1)变量(或节点)间的关系(相邻性和代际关系):在一特定路径上,由一个连接符接在一起的两个变量(或节点),称为相邻变量,表现为相邻关系,如图1中的A→B, A→D, E→□等。相邻变量互为母子关系,以图1中的B→□为例,称B为带框变量□的母变量,或□为B的子变量。如路径A→B→□→F上,□有母变量或更早的病因变量,则将这些变量(如A、B)称为□的祖先变量;如路径A→B→□→F上, B有子变量或更晚的效应变量,则将这些效应变量,称为后代变量,如F、□均为B的后代变量。当C的祖先变量没有更早的母变量影响时,则称此时的祖先变量为外源变量(或根节点,如A);该路径上所有的其他变量均称为内源变量。代际关系是基于特定的路径,因此同一变量在不同的路径可表现为不同的代际关系。

(2)碰撞变量:在一特定路径上,如果该路径的一个变量存在着两个或以上的母变量或祖先变量,则称该变量为该路径上的碰撞变量。如图1中的路径A→B→□←E, □同时是B和E的效应变量(子变量),因此□被称为该路径的碰撞变量。碰撞变量同样是针对特定的路径而言的。

(3)路径及其分类:①有向路径和无向路径:只由多个同向单箭头组成的路径,称为有向路径。如图1中A→B→□→F, A→E→□→F, A→B→E→□→G等均为有向路径。否则,则称为无向路径。所有的有向路径都是病因路径,也是因果关系研究的目标所在。②开放路径和关闭路径:一条路径上如存在碰撞变量,则称该路径为关闭路径(或失活路径、阻断路径),如图1中的路径A→B→□←E;如不

存在碰撞变量,则称该路径为开放路径(或活动路径、未阻断路径)。所有的有向路径都是开放路径,只存在单向的信息流动;但开放路径可包括有向路径和无向路径。所有的关闭路径都是无向路径。开放路径上的变量之间存在单向或双向的信息流动,即提示它们之间可能存在关联(非独立性)。路径的开放与关闭也是针对特定的路径而言的。③自然路径和触发路径:DAGs上所显示的所有路径均称为自然路径(natural path)。自然路径可包括:病因路径(有向路径)、混杂路径(如图2中E←A6→D)、碰撞路径等。碰撞路径可进一步分为单碰撞路径和双碰撞路径(图2-5和7)。如出现带框变量,如图1中的□,而导致其他变量之间出现新的关联,我们称之为触发路径(induced path),有时可在DAGs中增加虚线“---”来表示。

## 二、DAGs基本规则

DAGs反映变量之间的因果效应,基本规则:

1. 一条路径上所包含的变量只能出现1次,并且不能以其自身作为祖先变量或母变量,即一条路径上不能出现回环,如图2-9中的路径E→D→A<sub>6</sub>→E中的E;
2. 多个同向单箭头(有向路径)显示变量间存在着连续的因果效应,仅有向路径刻画这种关系;
3. 相邻变量(互为母子关系)总是有向连接,因此这两个变量总是存在依赖性;
4. 对于开放路径,通常表明,该路径上的某些变量之间存在关联,这种特征称为有向联接(directed connection)。所有的有偏路径,如混杂路径,都是开放路径;但开放路径不全是有偏路径,如有向路径是提供因果效应的病因路径;
5. 对于关闭路径,碰撞变量两侧的变量之间无条件独立的,这种特征成为有向分离(directed separation);这种独立性可表现在紧邻碰撞变量两侧的无符号连接的两个变量之间,也可表现在远离碰撞变量两侧的无符号连接的两个变量之间。

## 三、DAGs干扰变量

1. 干扰变量DAGs的9种基本结构:由于客观世界的复杂性,E→D所表述的因果效应(如E为病因,D是E的效应,为不可再发的非传染性事件),将不可避免受到第三个或更多因素(如干扰变量/向量A)的影响,这是因果效应研究无法回避的常态现象。

根据干扰因素A与E和D的关系,以DAGs表示,将出现以下9种基本结构(图2):

(1)A<sub>1</sub>是独立变量:A<sub>1</sub>与E或D均无箭头联结,说明A<sub>1</sub>不影响E或D,E-D的边际关联即可反映E-D的因果效应。因此,在分析时,无需调整A<sub>1</sub>;如调整A<sub>1</sub>,则调整的E-D关联仍然等于边际关联,但调整的E-D关联比边际关联的方差大。研究中诸如姓名、流水号及其他无关变量等。

(2)A<sub>2</sub>是工具变量:A<sub>2</sub>是E的病因,并只通过E影响D,因此,A<sub>2</sub>具有工具性(instrumentality),为工具变量(instrumental variable)<sup>[5-6]</sup>。E-D边际关联可反映E-D效应。因此,在分析时,没有必要调整A<sub>2</sub>;如调整A<sub>2</sub>,则调整的E-D

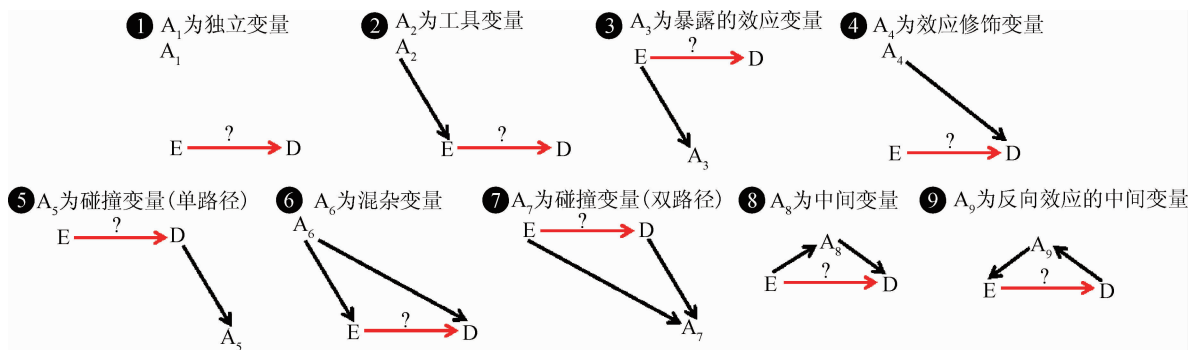


图2 干扰变量对因果效应的影响

关联仍然等于边际关联,但调整的E-D关联比边际关联的方差大。例如,乙醛脱氢酶 2 (aldehyde dehydrogenase 2, ALDH<sub>2</sub>)催化酒精的初次代谢产物—乙醛,而起到分解酒精的作用;编码 ALDH<sub>2</sub>的基因多态影响酒精摄入,从而建立 ALDH<sub>2</sub>基因多态作为研究饮酒与健康关系的工具变量<sup>[7]</sup>。

(3) A<sub>3</sub>是暴露E的效应变量:A<sub>3</sub>与D同是E的效应。E-D 边际关联可反映E-D效应。没有必要调整A<sub>3</sub>;如调整A<sub>3</sub>,则调整的E-D关联仍等于边际关联,但调整的E-D关联比边际关联的方差大。

(4) A<sub>4</sub>是效应修饰变量:A<sub>4</sub>与E同是D的病因,因此,A<sub>4</sub>为E-D效应修饰变量(effect modifier)<sup>[4, 8]</sup>。单独的E-D 边际关联不能正确反映E-D效应,E-D效应估计需要同时考虑到A<sub>4</sub>对其的效应修饰作用。调整A<sub>4</sub>将导致效应修饰偏倚(effect modification bias)<sup>[4]</sup>。

(5) A<sub>5</sub>是碰撞变量(单路径):如果存在E→D,即E-D为非零效应,则E通过D对A<sub>5</sub>产生效应,则称A<sub>5</sub>为单路径碰撞变量(Collider)<sup>[4, 8-9]</sup>。E-D 边际关联可反映E-D效应。没有必要调整A<sub>5</sub>;如调整A<sub>5</sub>,将产生单路径碰撞偏倚(uni-path collider bias),影响E-D效应估计。这也助于理解为何调整碰撞变量的子代(或后代)变量将引入偏倚。

(6) A<sub>6</sub>是混杂变量:A<sub>6</sub>既是E又是D的病因(共病因),为混杂的病因结构,A<sub>6</sub>为混杂变量(confounder)。E-D效应受到混杂路径E←A<sub>6</sub>→D的影响。因此,E-D 边际关联不能正确反映E-D效应。必需调整A<sub>6</sub>,才能获得正确的E-D效应。

(7) A<sub>7</sub>是碰撞变量(双路径):A<sub>7</sub>既是E又是D的效应(共效应),即E和D共同碰撞于A<sub>7</sub>(E→A<sub>7</sub>←D为关闭路径),因此,E-D 边际关联正确反映E-D效应。如调整A<sub>7</sub>,将产生双路径碰撞偏倚(bi-path collider bias)或选择性偏倚<sup>[10-12]</sup>,干扰E-D效应的估计。在病例对照研究中,因病例被纳入研究的概率通常高于对照,因此D→A<sub>7</sub>一般都成立,如E→A<sub>7</sub>能成立,则该DAGs结构形成。这也解释了病例对照研究中选择性偏倚为何容易发生。

(8) A<sub>8</sub>是中间变量:E可通过A<sub>8</sub>对D产生影响,因此A<sub>8</sub>是中间变量(mediator)。E-D 边际关联可反映E-D效应—总效应,包括直接效应(E→D)和间接效应(E→A<sub>8</sub>→D)。估计总效应时,没有必要调整A<sub>8</sub>;如估计直接效应,则需要调整A<sub>8</sub>。因此,A<sub>8</sub>是否调整,取决于研究目标是获得总效应、直接效应

还是间接效应。如估计的效应与其病因路径不一致,则将产生病因通路偏倚(causal pathway bias)<sup>[4]</sup>。

(9) A<sub>9</sub>是反向效应的中间变量:对E-D效应进行研究,但D通过A<sub>9</sub>对E产生因果效应,从而产生循环路径。因此,A<sub>9</sub>作为中间变量而产生反向的因果效应。从时序关系上看,E显然处于2个不同时间点,可分别用E<sub>0</sub>和E<sub>1</sub>予以区别(类似于时间协变量),以图3来表示,从而可不违背DAGs基本规则(见“DAGs基本规则”)。如在回忆偏倚中,E<sub>0</sub>为真实暴露,E<sub>1</sub>为测量的暴露,结局D通过中间变量A<sub>9</sub>影响暴露的测量E<sub>1</sub>。

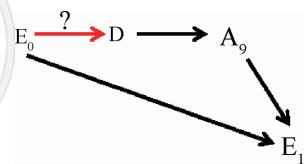


图3 反向效应的中间变量

2. DAGs 干扰变量调整规则:显然,对任一E-D因果关系的研究,其整体的病因图框架将全部由图2所提供的8种基本结构组成;列出自E至D及其干扰因素的所有路径,从而可形成一个完整的E-D因果关系研究的DAGs。因此,建立这个完整的DAGs是E-D因果关系研究的关键。

在此基础上,需要区分:①研究问题已明确(图2-8),只需确定识别何种效应—总效应、直接效应还是间接效应;②识别不影响E-D效应估计的关闭路径,予以排除;③识别非病因路径的开放路径,区分其中影响和不影响E-D效应估计的路径,确定最终用于调整的干扰因素,以获得正确的E-D效应估计。

如何处理干扰变量对E-D因果效应的估计? DAGs的规则如下:在E至D的路径上,调整干扰变量将切换其所在路径的开放或关闭特性;即当干扰变量为碰撞变量或碰撞变量的后代变量时,对其进行调整,将使得该路径从关闭转换为开放。如图2-7中,E→A<sub>7</sub>←D本身为关闭路径,并不影响E-D效应估计,但在病例对照研究的实际中因A<sub>7</sub>仅为纳入研究的人群而成为带框变量,事实上形成调整,从而开放了这条通路而导致偏倚;如干扰变量不是碰撞变量时,对其进行调整,将使得该路径从开放变换为关闭。如图2-6的混杂路径E←A<sub>6</sub>→D本身为开放路径,将影响E-D效应估计;调整A<sub>6</sub>将关闭这条路径,从而解释了为何混杂变量总是要进行调整。



同样道理,对从E-D的路径来说,当对一组变量A进行调整时,如果出现以下2种情形中的任何一种,则将使得从E-D的路径关闭:①该路径存在非碰撞变量Z,并且Z被包含在A里面,那么该路径将在Z处关闭;②该路径存在碰撞变量W,但该碰撞变量和其后代变量不被包含在A里面,那么该路径将在碰撞变量处关闭。

如果同时满足以下2种情形,则将使得从E-D的路径开放:①A不包含在该路径上的任何非碰撞变量,并且,②A包含在该路径上的任一碰撞变量或其后代变量。

四、DAGs的应用

1. 识别研究问题:提供建立的E-D效应估计的DAGs,识别干扰路径及其中的干扰变量,估计其是否可测量;否则,应在E-D研究前,优先需要解决是否存在偏倚?否则,应在整体的研究设计中整合适当的偏倚设计,以获得相应的偏倚参数,用于后续效应的定量偏倚分析<sup>[13]</sup>。

2. 理解和指导研究设计及实施:DAGs帮助并加深了对多种设计的认识和理解,如传统观察性设计(横断面研究<sup>[9]</sup>、病例对照研究<sup>[4]</sup>和队列研究<sup>[9]</sup>)、匹配设计<sup>[4, 14]</sup>、Meta分析<sup>[15]</sup>、准实验设计<sup>[16]</sup>等。

在研究设计阶段,可避免测量不必要的变量,从而达到简化研究的目的。提供研究人群选择上的思路,如病例对照研究设计本身已存在单变量碰撞结构(图2-5);选择对照的时候,应采取适当策略使得对照选择独立于暴露;否则,暴露影响到研究对象是否纳入,即E→A<sub>0</sub>成立,则形成类似图2-7的结构,而引起碰撞偏倚。

在混杂(图2-6)控制上,依据DAGs可容易理解混杂控制策略的依据:①打断混杂路径:常用策略如限制、匹配等,使得A<sub>0</sub>固定于某个值,而打断E←A<sub>0</sub>→D的路径;②外部干预:通过外部变量Z对E的干预,当Z可完全决定E,则形成了随机对照试验设计;当Z不完全决定E,则形成了工具变量设计及孟德尔随机化设计(mendelian randomization design)<sup>[17]</sup>(图4)。

3. 指导数据分析:在DAGs的指导下,围绕E-D效应及其所有路径,确定拟进行调整的干扰变量充分子集,以正确进行效应估计。向韧等<sup>[18]</sup>对如何应用DAGs以实现混杂因素最小子集的获取、具体步骤及优缺点等进行了详尽的介绍。

在资料分析阶段,避免加入不必要变量的调整,如图2-1,2,3<sup>[19-20]</sup>;或调整不应调整的变量,如图2-7中的A<sub>0</sub>,导致在分析阶段中人为地引入选择性偏倚,影响E-D效应估计<sup>[19-20]</sup>;需要考虑效应修饰变量的影响(图2-4);评估变量之间的共线性等<sup>[21]</sup>。

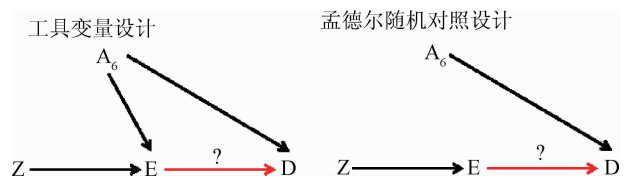


图4 外部干预对E-D效应估计

选择合适的统计分析策略,如图5,可通过分别估计无偏的E→M和调整E的M→D,以实现绕过A<sub>0</sub>所致的混杂效应,实现E-D效应的正确估计,形成了中间变量的分析方法<sup>[22-24]</sup>。

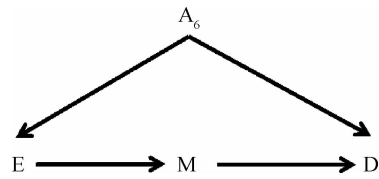


图5 中间变量分析法正确估计E-D效应

4. 偏倚的理解及结构分类:目前已识别的各类偏倚达上百种<sup>[25]</sup>,常常令人困惑并且难以掌握<sup>[26]</sup>。依据DAGs,Shahar和Shahar<sup>[4]</sup>将众多的偏倚归结为三对(6种)对立的偏倚,即混杂偏倚和选择性偏倚、效应修饰偏倚和病因通路偏倚、信息偏倚和思维偏倚(thought bias)。这种分类对偏倚的正确认识和理解非常高效。

5. 其他应用:此外,DAGs可用于评估模型的忠实性(faithfulness)或稳定性(Stability)<sup>[27]</sup>、效应的可折叠性(Collapsibility)<sup>[28]</sup>、研究人群的可交换性<sup>[29]</sup>等。

DAGs刻画了客观世界已明确的和待研究的因果效应的定性关系,是对整个研究设计及其实施过程的预演,有助于理清病因假设的基础、合理性和可行性,对可能存在的问题进行预判,有助于提出针对性的解决方案,同时也有助于提出新的研究假设。因此,DAGs对因果关系研究具有重要的指导价值。

利益冲突 无

参 考 文 献

[1] Judea P. Causality: Models, Reasoning, and Inference [M]. Cambridge University Press, 2009: 1-102.  
[2] Greenland S, Brumback B. An overview of relations among causal modelling methods [J]. Int J Epidemiol, 2002, 31 (5) : 1030-1037.  
[3] Pearl J. An introduction to causal inference [J]. Int J Biostat, 2010, 6(2): Article 7. DOI: 10.2202/1557-4679.1203.  
[4] Shahar E, Shahar DJ. Causal diagrams and the logic of matched case-control studies [J]. Clin Epidemiol, 2012, 4: 137-144. DOI: 10.2147/CLEPS.31271.  
[5] Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? [J]. Epidemiology, 2006, 17 (4) : 360-372. DOI: 10.1097/01.ede.0000222409.00878.37.  
[6] Greenland S. An introduction To instrumental variables for epidemiologists [J]. Int J Epidemiol, 2000, 29(6) : 1102.  
[7] Au Yeung SL, Jiang C, Cheng KK, et al. Is aldehyde dehydrogenase 2 a credible genetic instrument for alcohol use in Mendelian randomization analysis in Southern Chinese men? [J]. Int J Epidemiol, 2013, 42 (1) : 318-328. DOI: 10.1093/ije/dys221.  
[8] Shahar E, Shahar DJ. On the definition of effect modification [J]. Epidemiology, 2010, 21 (4) : 587; author reply 587-588. DOI:

- 10.1097/EDE.0b013e3181e0995c.
- [9] Shahar E, Shahar DJ. Causal diagrams and the cross-sectional study [J]. *Clin Epidemiol*, 2013, 5: 57-65. DOI: 10.2147/CLEP.S42843.
- [10] Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias [J]. *Epidemiology*, 2004, 15 (5) : 615-625.
- [11] Snoep JD, Morabia A, Hernandez-Diaz S, et al. Commentary: A structural approach to Berkson's fallacy and a guide to a history of opinions about it [J]. *Int J Epidemiol*, 2014, 43 (2) : 515-521. DOI: 10.1093/ije/dyu026.
- [12] Pearce N, Richiardi L. Commentary: Three worlds collide: Berkson's bias, selection bias and collider bias [J]. *Int J Epidemiol*, 2014, 43(2) : 521-524. DOI: 10.1093/ije/dyu025.
- [13] Lash TL, Fox MP, Fink AK: Applying quantitative bias analysis to epidemiologic data [M]. New York: Springer Science + Business Media, 2009: 13-32.
- [14] Mansournia MA, Hernan MA, Greenland S. Matched designs and causal diagrams [J]. *Int J Epidemiol*, 2013, 42 (3) : 860-869. DOI: 10.1093/ije/dyt083.
- [15] Shrier I. Structural approach to bias in Meta-analyses [J]. *Res Synth Methods*, 2011, 2(4) : 223-237. DOI: 10.1002/jrsm.52.
- [16] Steiner PM, Kim Y, Hall CE, et al. Graphical models for quasi-experimental designs [J]. *Sociol Methods Res*, 2015, 46 (2) : 0049124115582272. DOI: 10.1177/0049124115582272.
- [17] Bowden J, Davey SG, Haycock PC, et al. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator [J]. *Genet Epidemiol*, 2016, 40 (4) : 304-314. DOI: 10.1002/gepi.21965.
- [18] 向韧,戴文杰,熊元,等. 有向无环图在因果推断控制混杂因素中的应用 [J]. *中华流行病学杂志*, 2016, 37 (7) : 1035-1038. DOI: 10.3760/cma.j.issn.0254-6450.2016.07.025.
- Xiang R, Dai WJ, Xiong Y, et al. Application of directed acyclic graphs in control of confounding [J]. *Chin J Epidemiol*, 2016, 37 (7):1035-1038. DOI:10.3760/cma.j.issn.0254-6450.2016.07.025.
- [19] VanderWeele TJ. On the relative nature of overadjustment and unnecessary adjustment [J]. *Epidemiology*, 2009, 20 (4) : 496-499. DOI: 10.1097/EDE.0b013e3181a82f12.
- [20] Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies [J]. *Epidemiology*, 2009, 20(4) : 488-495. DOI: 10.1097/EDE.0b013e3181a819a1.
- [21] Schisterman EF, Perkins NJ, Mumford SL, et al. Collinearity and causal diagrams: a lesson on the importance of model specification [J]. *Epidemiology*, 2017, 28 (1) : 47-53. DOI: 10.1097/EDE.0000000000000554.
- [22] VanderWeele TJ. Mediation and mechanism [J]. *Eur J Epidemiol*, 2009, 24(5) : 217-224. DOI: 10.1007/s10654-009-9331-1.
- [23] Hafeman DM. A sufficient cause based approach to the assessment of mediation [J]. *Eur J Epidemiol*, 2008, 23 (11) : 711-721. DOI: 10.1007/s10654-008-9286-7.
- [24] de Stavola BL, Daniel RM, Ploubidis GB, et al. Mediation analysis with intermediate confounding: structural equation modeling viewed through the causal inference lens [J]. *Am J Epidemiol*, 2015, 181(1) : 64-80. DOI: 10.1093/aje/kwu239.
- [25] Delgado-Rodríguez M, Llorca J. Bias [J]. *J Epidemiol Community Health*, 2004, 58(8) : 635-641. DOI: 10.1136/jech.2003.008466.
- [26] Schwartz S, Campbell UB, Gatto NM, et al. Toward a clarification of the taxonomy of "bias" in epidemiology textbooks [J]. *Epidemiology*, 2015, 26 (2) : 216-222. DOI: 10.1097/EDE.0000000000000224.
- [27] Greenland S, Mansournia MA. Limitations of individual causal models, causal graphs, and ignorability assumptions, as illustrated by random confounding and design unfaithfulness [J]. *Eur J Epidemiol*, 2015, 30(10) : 1101-1110. DOI: 10.1007/s10654-015-9995-7.
- [28] Mansournia MA, Greenland S. The relation of collapsibility and confounding to faithfulness and stability [J]. *Epidemiology*, 2015, 26(4) : 466-472. DOI: 10.1097/EDE.0000000000000291.
- [29] Flanders WD, Eldridge RC. Summary of relationships between exchangeability, biasing paths and bias [J]. *Eur J Epidemiol*, 2015, 30(10) : 1089-1099. DOI: 10.1007/s10654-014-9915-2.

(收稿日期: 2017-01-10)

(本文编辑: 万玉立)