

测量偏倚的方向性:基于有向无环图的结构解析

李意杰^{1,2} 曹焱敏¹ 范伟¹ 张淼¹ 刘丽丽¹ 郑英杰¹

¹复旦大学公共卫生学院流行病学教研室/国家卫生健康委员会卫生技术评估重点实验室/公共卫生安全教育部重点实验室,上海 200032;²福建医科大学公共卫生学院流行病学与卫生统计学系,福州 350108

李意杰和曹焱敏对本文有同等贡献

通信作者:郑英杰,Email:zhengshmu@gmail.com

【摘要】 测量偏倚(MB)的因果结构已有阐述但仍不完全明确;实践上,暴露与结局间效应的替代估计值(SE)的正确性是因果推断的前提,通常基于测量的暴露与测量的结局间双向无差别的错分。本文提出了基于有向无环图(DAG)的单个变量测量的结构,其MB来源于对不完美的、类似“输入/输出设备”的测量系统的选择。SE的MB受到测量系统本身和测量系统外因素的双重影响:虽然测量系统的独立性或依赖性机制仍可确保SE的MB表现为双向无差别的错分,然而测量系统外因素对SE的MB则可表现为双向无差别、单向或双向有差别的错分。此外,反向因果关系是定义在测量水平上,测量的暴露可以影响测量的结局,反之亦然。结合时序关系,DAGs有助于阐明MB的结构、机制和方向性。

【关键词】 测量偏倚; 错分偏倚; 流行病学方法; 有向无环图; 效应估计

基金项目:国家自然科学基金(82173582,81773490);国家重点研发计划(2021YFC2701800,2021YFC2701801)

The directionality of measurement bias: a directed acyclic graph-based structural perspective

Li Yijie^{1,2}, Cao Yanmin¹, Fan Wei¹, Zhang Miao¹, Liu Lili¹, Zheng Yingjie¹

¹Department of Epidemiology, Key Laboratory of Public Health Safety of Ministry of Education, Key Laboratory for Health Technology Assessment, National Commission of Health, School of Public Health, Fudan University, Shanghai 200032, China; ²Department of Epidemiology and Health Statistics, School of Public Health, Fujian Medical University, Fuzhou 350108, China

Li Yijie and Cao Yanmin contributed equally to the article

Corresponding author: Zheng Yingjie, Email: zhengshmu@gmail.com

【Abstract】 Measurement bias (MB) has been described in causal structures but is still not entirely clear. In practice, the correctness of substitution estimate (SE) of effect is a prerequisite for causal inference, usually based on a bidirectionally non-differential misclassification between the measured exposure and the measured outcome. Based on a directed acyclic graph (DAG), this paper proposes a structure for the single-variable measure, where its MB is derived from the choice of an imperfect, "input/output device-like" measurement system. The MB of the SE is influenced both by the measurement system itself and by factors outside the measurement system: while the independence or dependence mechanism of the measurement system still ensures that the MB of the SE is bidirectionally non-differential; however, the misclassification can be bidirectionally

DOI:10.3760/cma.j.cn112338-20220906-00765

收稿日期 2022-09-06 本文编辑 万玉立

引用格式:李意杰,曹焱敏,范伟,等.测量偏倚的方向性:基于有向无环图的结构解析[J].中华流行病学杂志,2023,44(4):643-649. DOI:10.3760/cma.j.cn112338-20220906-00765.

Li YJ, Cao YM, Fan W, et al. The directionality of measurement bias: a directed acyclic graph-based structural perspective [J]. Chin J Epidemiol, 2023, 44(4):643-649. DOI:10.3760/cma.j.cn112338-20220906-00765.



non-differential, unidirectionally differential, or bidirectionally differential resulted from the factors outside the measurement system. In addition, reverse causality should be defined at the level of measurement, where measured exposures can influence measured outcomes and vice versa. Combined with temporal relationships, DAGs help elucidate MB's structures, mechanisms, and directionality.

【Key words】 Measurement bias; Misclassification bias; Epidemiological methods; Directed acyclic graphs; Effect estimate

Fund programs: National Natural Science Foundation of China (82173582, 81773490); National Key Research and Development Program of China (2021YFC2701800, 2021YFC2701801)

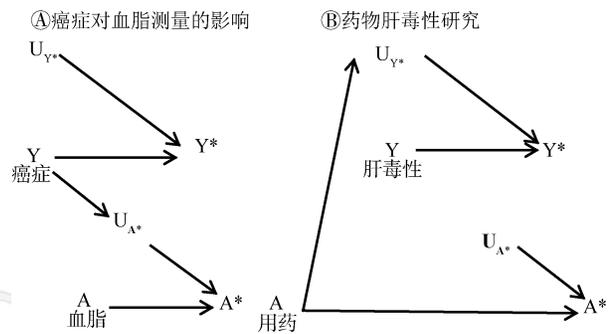
一、背景

测量是科学家认识世界的基础。测量偏倚 (measurement bias, MB) 随测量的实施而发生^[1-5]; 这是一个必须面对的困境: 测量某事物并处理 MB。

若人群中存在一个我们感兴趣的属性 (例如血糖), 其真实值为 A, 经一个测量系统 (M_{A^*}) 获得估计值 (A^*)。一般来说, M_{A^*} 是一些测量过程关键元素的组合, 包括研究人员、研究对象、工具、方法等, 贯穿于测量的全过程: 可形象地认为, M_{A^*} 就像一个“输入/输出设备 (I/O device)”, 即将 A 输入后, 输出 A^* 。A 和 A^* 之间的差异 ($U_{A^*} = A^* - A$) 即 MB。若 A 为分类变量, MB 又被称为错分偏倚^[6-9]。为阐述方便, 如未特别指出, 本文提及的变量均为分类变量; 变量的命名规则: 真实值 (以大写字母如 A 表示)、其对应的测量值 (同名变量带“*”表示, 如 A^*)、测量系统 (M_{A^*}) 及产生的 MB (MB_{A^*})。

测量及其偏倚可合理地采用有向无环图 (directed acyclic graph, DAG) 来展示^[10-23]。目前, 单个变量 MB 的结构已达成了一个共识, 即: 除了 A 影响 A^* 之外, 影响 A^* 值的所有其他因素对 A^* 的影响即 A^* 的 MB (MB_{A^*}); 但当前提出的两种 MB-DAGs 结构针对 MB 的处理策略上存在着不同: Hernán 和 Cole^[14] 以 U_A 代表了除 A 以外的所有其他影响 A^* 的因素, 同时也代表了测量值 A^* 的误差值。然而, U_A 和 A^* 是测量一刻同时产生的, 二者间不可能存在因果律, 但被 Hernán 和 Cole^[14] 用于因果图的表述中; 因此本文以 U_{A^*} 替换 Hernán 和 Cole^[14] 文中的 U_A (图 1); 而 Shahar 等^[13, 23] 则将除 A 以外的所有其他影响 A^* 的因素用一个变量数列 ($M_1 \sim M_n$) 表示, 并逐一地列出它们各自对 A^* 的同等影响。显然, 前一种方法难以提供“所有除了 A 之外的其他影响 A^* 的因素”的确切含义, 而后一种的枚举法未能捕捉到测量系统的主要特征。

此外, 这些策略尚存在其他问题: ① A 值从不会为人所知, 但在某一确定的时点, A 值理论上是



注: 以 U_{A^*} 替换 Hernán 和 Cole^[14] 研究中的 U_A ; A^*/Y^* : A/Y 的测量值; U_A/U_{Y^*} : 除 A/Y 以外的所有其他影响 A^*/Y^* 的因素

图 1 测量偏倚的结构

一个常数。因此, 一旦 A 得以测量则可得到 A^* 值, 同时也得到其对应的 U_{A^*} 值 ($U_{A^*} = A^* - A$)。因 A 未知, 故 U_{A^*} 亦为一个理论上的差值, 在现实中不会存在, 即: 某一因素只可能影响 A 或 A^* , 而不可能影响 U_{A^*} 。由此可见, 当前已发表的 MB-DAGs 研究将其他变量与 U_{A^*} 通过单箭头连结来表示二者间因果律的做法值得商榷。例如, Hernán 和 Cole^[14] 举例说明有差别错分偏倚的 DAG 结构中, 癌症 (Y) 通过影响“所有除了血脂 (A) 之外的其他影响血脂的测量值 (A^*) 的因素” (即 U_{A^*}), 进而影响 A^* , 表示为: $Y \rightarrow U_{A^*} \rightarrow A^*$ 和 $A \rightarrow A^*$ (图 1), 前者的表述显然不成立, 实际上是 Y 影响了 A 进而影响 A^* , 即 $Y \rightarrow A \rightarrow A^*$ 和 $M_{A^*} \rightarrow A^*$; 相似的问题可见于某种药物的肝毒性影响医生对用药者监测的例子 (图 1)。② 当研究 A 对 Y 的效应时, 可以明确的是: 经测量或观察而为人所知的变量 A^* (或 Y^*), 不会影响未知变量 Y (或 A) 的真实值^[24]。然而, 未知的 A (或 Y) 影响未知的 Y (或 A) 通常被称为因果关系 (或反向因果关系), 这点常令人困惑。③ 实践中, 因 A 和 Y 均未知, 以 $A^* - Y^*$ 间的关系作为效应的替代估计值 (substitution estimate, SE), 其正确性是因果推断的前提, 通常基于 A^* 与 Y^* 间存在着双向无差别的错分。然而, 迄今尚未有任何 DAG 能够阐明 MB 的方向性问题。④ 实际上, 研究中只会选择一种测量系

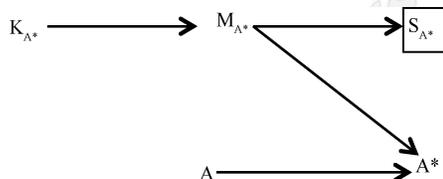
统(或称实测法),但测量系统的选择过程尚未在 MB-DAGs 中提及。

因此,本文结合实例分三部分进行阐述:①提出用于测量单个变量的结构;②测量系统对 SE 的 MB 的影响;③测量系统外因素对 SE 的 MB 的影响,以阐明 MB 的结构、机制及其方向性。

二、单个变量的测量

以葡萄糖耐量试验进行血糖检测为例,虽然被检者提前摄入的葡萄糖将导致血液葡萄糖浓度的上升,但在某一具体时点上抽血并测定其葡萄糖浓度时,该时点的血液葡萄糖的真实值(A)是明确的、固定的。由此可认为,A 是一个外源变量。与 A 不同,A*和 U_{A^*} 的值可能因 M_{A^*} 的不同而变化。

影响 A* 的原因可分为两方面:一是 A 值,遵循类似于“价值规律”的现象,通常 A* 总是决定于 A,即 $A \rightarrow A^*$;二是测量系统(M_{A^*})。因此,因果图中的 $M_{A^*} \rightarrow A^*$ 代表着 MB_{A^*} 。对 A* 的背景或先验知识(K_{A^*})的认识程度为正确理解及选择 M_{A^*} 奠定了基础。对 K_{A^*} 的认识越多,就更能帮助我们综合考虑测量性能、成本和可行性等,从而提出候选的 M_{A^*} 。在一次测量中,所有候选的 M_{A^*} 中只有一个相对科学、合理和有效的 M_{A^*} 将被最终选择(记作 S_{A^*})并付诸于实践,这可能产生相对准确且性价比较佳的 A* 值。基于此,本文提出了单个变量 A 测量的 MB-DAG(图 2)。



注:A:单个变量的真实值;A*:A的测量值; K_{A^*} :对A*的先验知识的认识程度; M_{A^*} :候选的测量系统; S_{A^*} :实际选择的测量系统; S_{A^*} 代表对 M_{A^*} 进行限制到实际选择的测量系统

图2 单个变量A测量的有向无环图

通常经由 S_{A^*} 选中并付诸实施的 M_{A^*} 只有一种(实测法),但可以想象存在着一种真实值的理论测量方法(真值法)来测量 A 而获得 A* 值(此时 $A=A^*$)。测量时,实际上进行的是通过实测法与真值法分别得到的测量值和真实值的对比研究(图 2)。实测法的性能与真值法越接近,则 $A \rightarrow A^*$ 越强,而 $M_{A^*} \rightarrow A^*$ 则越弱;当二者完全一致,此时 A 完全决定了 A*,由此路径 $M_{A^*} \rightarrow A^*$ 中的“ \rightarrow ”消失,理论上得到了与 A 值完全一样的 A* 值;若 $M_{A^*} \rightarrow A^*$ 的关系不

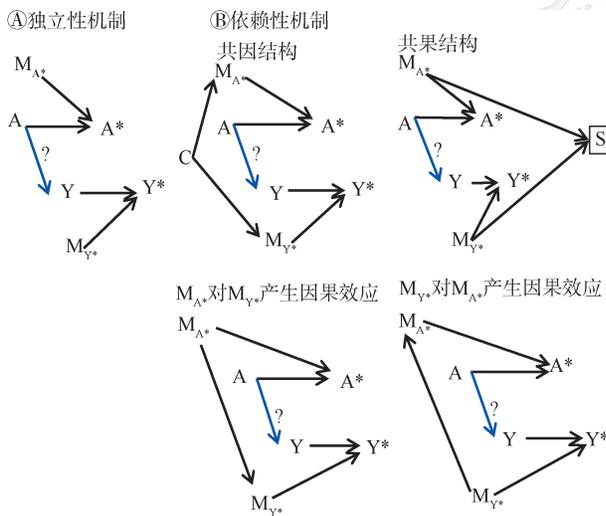
变,则意味着 $A \rightarrow A^*$ 也不变(后续针对 SE 的 MB 阐述将用到)。这进一步说明了图 2 中“ $M_{A^*} \rightarrow A^*$ ”的结构代表着 MB_{A^*} 。然而在实际测量中,即使采用一种当前性能最佳的实测法(即“金标准”),亦难以保证其与真值法完全一样。因此, $M_{A^*} \rightarrow A^*$ 中的“ \rightarrow ”不会消失,即:A* 将受到 M_{A^*} 的影响,这也意味着 MB_{A^*} 将不可避免地产生(以下的阐述均基于此种情形,涉及的 DAGs 将不再绘制 K_{A^*} 与 S_{A^*})。若候选的 M_{A^*} ($M_{A1^*}, M_{A2^*}, \dots, M_{Aj^*}$) 中有任意两种及以上被选中,即相当于进行不同测量系统的一致性研究。

上述单个变量测量的 DAG 及其思维对理解单个变量 MB 及其分类很有帮助。测量是研究人员和研究对象通过合适的 M_{A^*} 进行互动的过程;通过分析 $M_{A^*} \rightarrow A^*$ 的结构可知, M_{A^*} 中影响 A* 的一个或多个主导因素将引发 MB_{A^*} :如 M_{A^*} 的主导因素来自于研究人员,对研究对象进行不完美的观察、调查、测量、检查、检测、询问、记录信息(有意或无意的),从而相应地将产生不同的偏倚;如 M_{A^*} 的主导因素来自于研究对象,主动或被动地向研究人员进行不完美的信息报告、回忆等(有意或无意的),也将产生相应的报告或回忆偏倚等^[25];如 M_{A^*} 的主导因素来自于方法(或工具),因其不统一、检测性能低,将相应地产生测量方法(或工具)引起的偏倚、低检测性能偏倚;如 M_{A^*} 的主导因素来自于电子和纸质记录,将相应地出现记录登记不全引起的偏倚。在整个测量过程中, M_{A^*} 的上述主导因素可能部分或全部出现;因此,最终产生的 MB 是这些不同测量性能的 M_{A^*} 综合作用而产生的结果。

最后,这种思维对理解单个变量 MB 的控制方法也很有帮助。通常认为,采用统一的方法或盲法可起到控制 MB 的作用。方法统一代表着我们用同样性能的 M_{A^*} 来进行测量,而方法不统一代表着我们用不同性能的 M_{A^*} 来进行测量,后者所获得的测量结果是这些不同性能方法的一个加权值。因此,方法统一仅仅只能去除方法不统一所带来的 MB,而不代表着能减少其他原因引起的 MB。同样地,采用盲法测量只能去除由非盲法测量引起的 MB。此外,采用同一 M_{A^*} 在先后不同时间对同一人群进行两次测量,发现结果具有较高的一致性,以此说明这种 M_{A^*} 的有效性的方法显然是错误的。两次测量结果的一致性只能说明 M_{A^*} 测量性能的稳定性,以及人群中所测量指标的真实值总体不受时间的影响;但测量值的 MB 仍然存在。这也解释了当前的大数据研究不能避免 MB 的原因之一。

三、暴露-结局间效应替代估计的影响

若 A 为暴露, Y 为结局, 拟研究 A-Y 间的效应, 实践中以 A*-Y* 间的关联来估计。由上可知, A 和 Y 均难免出现 MB: ① 当 A→Y 时, 因 A*←A→Y→Y*, 故 A* 和 Y* 将出现关联, 此时 A 本身是一个影响 A*-Y* 间效应估计的混杂因素(图 3)。因此, 由这条唯一的、以 A 为混杂子的混杂路径所代替的 A*-Y* 间的关联, 即是研究感兴趣的 SE, 或因果关联^[14, 23]。② 若 A-Y 间的效应为零, 上述替代估计路径自然中断, 此时 SE 应为零关联。由此可见, 为了满足效应研究的可行性, 通常需要假设 SE 是 A-Y 间效应的合理且可行的替代估计, 这正是所有病因研究的潜在含义, 保证 SE 的准确性成为因果推断研究的现实目标。然而, SE 将受到测量系统本身及测量系统之外因素的共同影响而产生 MB(以 MB_{SE} 表示), 现分别进行阐述。



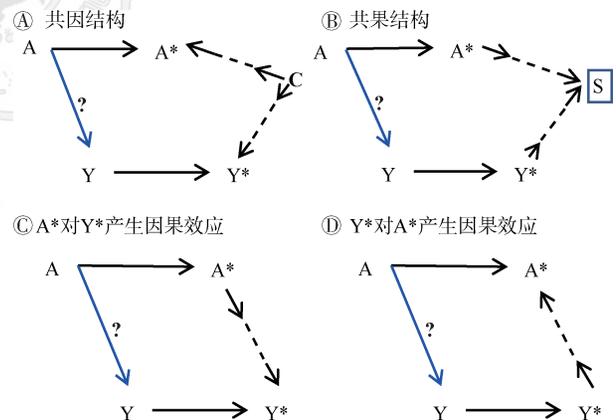
注: A/Y: A/Y 的真实值; A*/Y*: A/Y 的测量值; M_{A*}/M_{Y*}: A/Y 的测量系统; 带问号的箭头表示箭头两端变量间的关系尚在研究中; C: M_{A*} 和 M_{Y*} 的共因; S: M_{A*} 和 M_{Y*} 的共果, [S] 代表着对 S 的调整

图 3 暴露(A)和结局(Y)之间的效应研究: 测量系统的独立性机制和依赖性机制

1. 测量系统: 独立性与依赖性机制(图 3): 当 A 和 Y 的测量过程相互独立时(即测量系统的独立性机制), M_{A*} 与 M_{Y*} 间无箭头连结。因自 M_{A*} 至 A* 仅有一条路径(即 M_{A*}→A*, 即 MB_{A*}), 故 Y 不会对 MB_{A*} 产生影响; 同理, A 不会对 M_{Y*}→Y*(MB_{Y*}) 产生影响。因此, MB_{SE} 为双向无差别的错分偏倚。M_{A*} 与 M_{Y*} 间可存在 4 种因果机制(共因、共果、因果关系或反向因果关系)^[26], 从而可能导致 A* 和 Y* 相互依赖, 此即为测量系统的依赖性机制: ① 在 A→Y 的

情形下, 若 M_{A*} 和 M_{Y*} 间存在着共因结构时, 自 M_{A*} 至 A* 将新增一条路径, 即 M_{A*}←C→M_{Y*}→Y*←Y←A→A*, 因其为关闭路径, 故其不会影响 MB_{A*}(或 Y 不影响 MB_{A*}); 同样地, 自 M_{Y*} 至 Y* 将新增一条路径, 即 M_{Y*}←C→M_{A*}→A*←A→Y→Y*, 亦为关闭路径, 故其不会影响 MB_{Y*}(或 A 不影响 MB_{Y*}), 故 MB_{SE} 仍是双向无差别的错分偏倚。例如, 在慢性 HBV 的母婴传播研究中, 在配对的母婴个体中使用相同的 HBsAg 检测方法(此时 M_{A*}=M_{Y*})。同理可分析, M_{A*} 与 M_{Y*} 间存在着共果或因果关系时, MB_{SE} 亦为双向无差别的错分偏倚。② 在 A 和 Y 为零效应的情形下, 上述两条关闭路径不存在, 故 MB_{SE} 亦为双向无差别的错分偏倚。

2. 测量系统以外的因素: 对暴露和结局的测量通常只使用一种实测法, 因果图将略去测量系统(M_{A*} 与 M_{Y*}) 的展示。不论测量系统是独立性还是依赖性, 测量系统产生的 MB_{SE} 均表现为双向无差别的错分。然而, SE 仍然可能受到测量系统外因素的影响, 亦可表现为 4 种因果机制^[26], 按照 A 和 Y 间存在着非零效应首先阐述(图 4)。

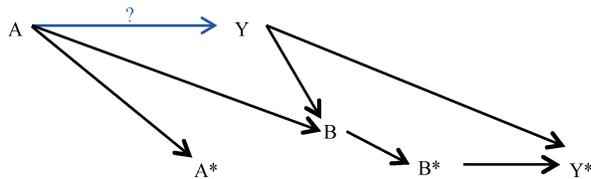


注: A/Y: A/Y 的真实值; A*/Y*: A/Y 的测量值; C: A* 和 Y* 的共因; S: A* 和 Y* 的共果, [S] 代表着对 S 的调整; - - - 代表以其前后变量为起止的路径上所有的单箭头在方向上是一致的; 带问号的箭头表示箭头两端的变量间的关系正在研究中

图 4 在 4 种不同机制下暴露(A)和结局(Y)间效应研究的有向无环图

(1) A* 和 Y* 间存在着共因结构: 共因结构存在两条新的关闭路径, 即 A→Y→Y*←...←C→...→A* 和 Y←A→A*←...←C→...→Y* (图 4)。因此, 自 A 至 A* 间仅有 A→A* 一条开放路径, 故 Y 不会对 A→A* 产生影响, 这也间接地反映了 Y 不会对 MB_{A*} 产生影响; 同理, A 不会对 MB_{Y*} 产生影响, 故 MB_{SE} 为双向无差别的错分偏倚。

以雌激素使用和子宫内膜癌发生关系研究为例^[27],由于雌激素使用(A)和子宫内膜癌(Y)均可能增加阴道出血(B)的概率,后者可以由女性自己和/或医生予以识别或诊断(B*),这将增加患者就诊和临床检查的机会,从而相应地增加患者被诊断为子宫内膜癌的机会(Y*)。这种偏倚通常被称为检出偏倚,可以用与研究兴趣无关的额外混杂路径来解释,即 $A^* \leftarrow A \rightarrow B \rightarrow B^* \rightarrow Y^*$ 和 $A^* \leftarrow A \rightarrow Y \rightarrow B \rightarrow B^* \rightarrow Y^*$ (若 $A \rightarrow Y$ 成立) (图 5)。此即共因结构机制。这里应该强调的是,是研究对象出现并知晓了已被测量的中间结果(阴道出血的测量值 B*),而非其未测量的、未知的真实值(阴道出血 B),增加了被直接诊断为子宫内膜癌(Y*)的机会。



注: A: 雌激素使用的真实情况; A*: A 的测量值; Y: 患子宫内膜癌的真实情况; Y*: Y 的测量值; B: 阴道出血的真实情况; B*: B 的测量值; 带问号的箭头表示箭头两端的变量间的关系正在研究中

图 5 关于雌激素使用和子宫内膜癌之间病因学研究的有向无环图

(2) A* 和 Y* 间存在着共果结构: 共果结构也有两条新的关闭路径, 即 $A \rightarrow Y \rightarrow Y^* \rightarrow \dots \rightarrow S \leftarrow \dots \leftarrow A^*$ 和 $Y \leftarrow A \rightarrow A^* \rightarrow \dots \rightarrow S \leftarrow \dots \leftarrow Y^*$; 但因调整 S (即图中 \boxed{S}), 这两条路径均从关闭转为开放, 此时 A* 和 Y* 的值都会受到影响。因此, 自 A 至 A* 间除了 $A \rightarrow A^*$ 一条开放路径外, 尚有因调整 S 而产生的一条新的开放路径, 故 Y 对 $A \rightarrow A^*$ 将产生影响, 这也间接地反映了 Y 将对 MB_{A^*} 产生影响。同理, A 将对 MB_{Y^*} 产生影响 (图 4)。故 MB_{S^*} 为双向有差别的错分偏倚。例如, 在基于记录系统的研究和历史队列研究中, 当 A 和 Y 同时被测量时, A* 和 Y* 不可能产生因果关系。从这个角度来看, 这类研究不太可能产生额外的 MB。然而, 由于这类研究通常局限于记录在案的研究对象 (即 $S=1$, 代表研究对象以前被记录过, 纳入研究; 若 $S=0$, 则代表研究对象未被记录过, 未纳入研究)。因此, $S=1$ 引发的选择偏倚导致了 MB 的产生, 可以用与研究兴趣无关的额外开放路径来解释, 即 $A^* \rightarrow \dots \rightarrow \boxed{S} \leftarrow \dots \leftarrow Y^*$ (图 4)。此即共果结构机制, 当被登记的人群与未被登记的人群之间在某些特征上不具有可比性时, 可能会导

致替代估计的有偏估计, 或知情存在偏倚^[28-29]。

(3) A* 和 Y* 间存在着因果关系结构: 如果 A* 和 Y* 间存在着因果关系, 有两种情形。

第一种情形, 如果 A* 对 Y* 产生因果效应, 两条新的路径将产生, 即 $A \rightarrow Y \rightarrow Y^* \leftarrow \dots \leftarrow A^*$ 和 $Y \leftarrow A \rightarrow A^* \rightarrow \dots \rightarrow Y^*$ 。前一条路径为关闭路径, 但后一条路径为开放路径。由上同样推理可知, A 对 Y* 的错分是有差别的, 而 Y 对 A* 的错分是无差别的 (图 4)。以病毒性疾病 (A) 和抗生素处方 (Y) 研究中的霍桑效应为例^[30], 由医生诊断为病毒性疾病的儿童 (A*) 参与了这项研究。由于医生在接受调查时需要被录音, 这导致了医生自身防护意识的提高, 从而降低了他们对抗生素处方开具的报告 (Y*)。此即因果关系机制; 需要强调的是, 医生对儿童的诊断 (A*) 改变了对医生开具处方行为的报告 (Y*), 或 $A^* \rightarrow Y^*$ 。

第二种情形, 当 A* 和 Y* 之间的因果关系被“逆转”时 (常见于病例对照研究中), 两条新的路径将产生, 即 $A \rightarrow Y \rightarrow Y^* \rightarrow \dots \rightarrow A^*$ 和 $Y \leftarrow A \rightarrow A^* \leftarrow \dots \leftarrow Y^*$; 这种情况正好与第一种相反, 前一条路径为开放路径, 后一条路径为关闭路径, 因此, A 对 Y* 的错分是无差别的, 而 Y 对 A* 的错分是有差别的 (图 4)。这一机制可用于说明在病例对照研究中常提到的回忆偏倚。以研究分娩先天性畸形婴儿的母亲为例, 在婴儿被确诊出现先天性畸形 (Y*) 后, 母亲可能会高估或低估既往暴露 A*。这种机制通常被称为反向因果关系, 可以由与研究兴趣无关的额外开放路径来解释, 即 $Y^* \rightarrow \dots \rightarrow A^*$ (图 4)。同样地, 这里应该强调的是, 研究对象出现并知晓了已被测量的结局 (婴儿被确诊为先天性畸形, Y*) 对暴露的测量 (A*) 产生影响, 即: $Y^* \rightarrow A^*$, 而非真实的、未知的结局 Y 对暴露 A 产生影响。因此, 反向因果关系是一个模糊且多余的概念, 是根据测量的暴露和测量的结局间的关系来定义的, 但在许多文献报道的 DAGs 中经常错误地以 $Y \rightarrow A$ 来表示。显然, 因果关系永远不会逆转。

本研究阐述了当 A 和 Y 间存在着非零效应时, 错分偏倚对应的 4 种机制。当 A 和 Y 间为零效应时, 因 A 与 Y 间无箭头存在, 故 $A \rightarrow A^*$ 或 $Y \rightarrow Y^*$ 不会受到 A* 和 Y* 间存在共因结构、共果结构和因果关系结构 (A* 对 Y* 产生效应或 Y* 对 A* 产生效应) 这 4 种路径的影响, 即其对应的 MB 也不受到影响, 均为双向无差别的错分 (表 1)。按理 A* 与 Y* 间的关联应为零; 但由于存在上述 4 种路径提及的与研

究兴趣无关的额外路径,仍将出现 MB_{SE} 。

表 1 零效应和非零效应下测量的暴露(A*)和结局(Y*)的错分偏倚

效应及关联机制	A 对 MB_{Y^*} 的影响	Y 对 MB_{A^*} 的影响
零效应		
A*和Y*间存在共因结构	无差别错分	无差别错分
A*和Y*间存在共果结构	无差别错分	无差别错分
A*对Y*产生效应(直接或间接)	无差别错分	无差别错分
Y*对A*产生效应(直接或间接)	无差别错分	无差别错分
非零效应		
A*和Y*间存在共因结构	无差别错分	无差别错分
A*和Y*间存在共果结构	有差别错分	有差别错分
A*对Y*产生效应(直接或间接)	有差别错分	无差别错分
Y*对A*产生效应(直接或间接)	无差别错分	有差别错分

注: A/Y: 暴露 A/结局 Y 的真实值; A*/Y*: A/Y 的测量值; MB_{A^*} 和 MB_{Y^*} 分别为单个变量 A 和 Y 的 MB

四、讨论与结论

测量过程可合理地由因果图来展示^[21]。本研究首先提出了单个变量测量的因果结构,接着识别 MB_{SE} 在测量系统的独立性和依赖性机制下均表现为双向无差别的错分,而测量系统外因素导致的 MB_{SE} 可能表现为双向无差别、单向或双向有差别的错分。

理论上,无论是对于单个变量或变量间的关系来说,MB 是变量或变量间关系的真实值与其测量值之间的差异。基于单个变量测量类似“输入/输出”的过程,本文结合了先验知识、因果现实和测量系统的选择过程,提出了单个变量测量的 DAG,有以下优点:首先,提出了测量系统的概念,突出了测量系统的“输入/输出”作用,识别了 U_{A^*} 不是因果图上的独立变量,是测量系统对 A 的测量值的影响,即 $MB(M_{A^*} \rightarrow A^*)$ 。其次,这种新的认识有助于理解不同类型的单个变量 MB。具有某些缺陷的测量系统会导致相应的偏差。例如,研究人员的粗心观察会导致观察者偏倚;研究对象不完善的报告可能导致报告偏倚;标准化测量过程未依从,则导致非标准化的偏倚。最后,这一思想可迅速地被应用到不同测量系统的性能评估研究中。

单个变量 MB 不可避免,由此影响到效应的估计值,即: A-Y 间的效应和 A*-Y* 间的 SE 难免不一致。实践中采用 SE 进行因果推断,基于 MB_{SE} 存在的双向无差别的错分来确保 SE 的正确性,虽然这将导致 SE 趋向于无效值。然而,这是当前因果推断唯一可行的做法。因此,判定该条件是否满足尤

为重要。

影响 SE 的正确性可有两种来源:一个来自测量系统,另一个来自测量系统之外因素产生的、与研究兴趣无关的任何额外的 A*---Y* 间的关联。本研究发现,测量系统的独立性或依赖性机制、零效应和测量系统外因素: A* 和 Y* 之间存在的共因结构可确保 MB_{SE} 表现为双向无差别的错分;然而这种特征在测量系统外因素: A* 和 Y* 之间存在共果或因果结构时则不能实现,提示着后者应加以重视,以确保 SE 的正确性。结合实例,本研究发现,测量可导致利益相关者做出任何心理或行为上的改变^[9],从而导致 MB 的产生。例如,诊断为阴道出血的女性增加了临床就诊次数,增加确诊为子宫内膜癌的机会;参与研究改变了医生开具抗生素处方的行为;分娩先天性畸形婴儿的母亲会高估或低估她们过去的暴露。

MB 往往被研究者所忽略^[31],但其本身并非无足轻重;同时尤为重要的是我们需要尽可能地获得所有单个变量的正确测量值;当有验证数据可供使用时,我们必须转向效应的重建^[18,32-33]。因此,如何正确构建测量过程的因果结构是控制 MB 的前提。本研究提出的单个变量测量结构,虽然从形式上与 Hernán 和 Cole^[14] 的一致,但其内涵已得以澄清;识别了错分偏倚的机制及其方向性,并阐明了反向因果关系是定义于测量水平上的概念。这些发现为正确认识 MB 和构建正确的 DAGs 奠定基础,并指导着 MB 控制策略的建立和形成。

综上所述,本研究提出的单个变量测量的结构为正确地构建测量 DAG 奠定基础,从而指导着研究实践;在效应估计上,应重点关注不满足“ MB_{SE} 双向无差别错分”的情景。基于测量水平,结合时序关系,DAGs 有助于阐明 MB 的结构、机制及其方向性。

利益冲突 所有作者声明无利益冲突

作者贡献声明 李意杰、曹焱敏:研究设计、论文撰写和修改;范伟、张森、刘丽丽:研究设计、论文修改;郑英杰:研究设计、研究指导、经费支持

参 考 文 献

- [1] Grimes DA, Schulz KF. Bias and causal associations in observational research[J]. Lancet, 2002, 359(9302): 248-252. DOI:10.1016/S0140-6736(02)07451-2.
- [2] Page LA, Henderson M. Appraising the evidence: what is measurement bias? [J]. Evid Based Ment Health, 2008, 11(2):36-37. DOI:10.1136/ebmh.11.2.36.
- [3] Delgado-Rodríguez M, Llorca J. Bias[J]. J Epidemiol

- Community Health, 2004, 58(8): 635-641. DOI: 10.1136/jech.2003.008466.
- [4] Vineis P. History of bias[J]. *Soz Präventivmed*, 2002, 47(3): 156-161. DOI:10.1007/bf01591887.
- [5] Höfler M. The effect of misclassification on the estimation of association: a review[J]. *Int J Methods Psychiatr Res*, 2005, 14(2):92-101. DOI:10.1002/mpr.20.
- [6] Bross I. Misclassification in 2×2 Tables[J]. *Biometrics*, 1954, 10(4):478-486. DOI:10.2307/3001619.
- [7] Diamond EL, Lilienfeld AM. Misclassification errors in 2×2 tables with one margin fixed:some further comments[J]. *Am J Public Health Nations Health*, 1962, 52(12): 2106-2110. DOI:10.2105/ajph.52.12.2106.
- [8] Greenland S, Kleinbaum DG. Correcting for misclassification in two-way tables and matched-pair studies[J]. *Int J Epidemiol*, 1983, 12(1): 93-97. DOI: 10.1093/ije/12.1.93.
- [9] Ahrens W, Pigeot I. *Handbook of epidemiology*[M]. New York: Springer, 2014: 639-658. DOI: 10.1007/978-0-387-09834-0.
- [10] Wang W, Shen G, Shahar E, et al. Forced expiratory volume in the first second and aldosterone as mediators of smoking effect on stroke in African Americans: the Jackson heart study[J]. *J Am Heart Assoc*, 2016, 5(1):1-8. DOI:10.1161/JAHA.115.002689.
- [11] Shahar E, Shahar DJ. Causal diagrams, the placebo effect, and the expectation effect[J]. *Int J Gen Med*, 2013, 6: 821-828. DOI:10.2147/IJGM.S52209.
- [12] Shahar E, Shahar DJ. Causal diagrams and change variables[J]. *J Eval Clin Pract*, 2012, 18(1): 143-148. DOI: 10.1111/j.1365-2753.2010.01540.x.
- [13] Shahar E, Shahar DJ. On the causal structure of information bias and confounding bias in randomized trials[J]. *J Eval Clin Pract*, 2009, 15(6): 1214-1216. DOI: 10.1111/j.1365-2753.2009.01347.x.
- [14] Hernán MA, Cole SR. Invited commentary: causal diagrams and measurement bias[J]. *Am J Epidemiol*, 2009, 170(8):959-962. DOI:10.1093/aje/kwp293.
- [15] Hernán MA, Hernández-Díaz S, Werler MM, et al. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology[J]. *Am J Epidemiol*, 2002, 155(2): 176-184. DOI: 10.1093/aje/155.2.176.
- [16] Robins JM. Data, design, and background knowledge in etiologic inference[J]. *Epidemiology*, 2001, 12(3): 313-320. DOI:10.1097/00001648-200105000-00011.
- [17] Blakely T, McKenzie S, Carter K. Misclassification of the mediator matters when estimating indirect effects[J]. *J Epidemiol Community Health*, 2013, 67(5):458-466. DOI: 10.1136/jech-2012-201813.
- [18] Kuroki M, Pearl J. Measurement bias and effect restoration in causal inference[J]. *Biometrika*, 2014, 101(2):423-437. DOI:10.1093/biomet/ast066.
- [19] Edwards JK, Cole SR, Westreich D. All your data are always missing: incorporating bias due to measurement error into the potential outcomes framework[J]. *Int J Epidemiol*, 2015, 44(4): 1452-1459. DOI: 10.1093/ije/dyu272.
- [20] Mansournia MA, Higgins JP, Sterne JA, et al. Biases in randomized trials: a conversation between trialists and epidemiologists[J]. *Epidemiology*, 2017, 28(1):54-59. DOI: 10.1097/EDE.0000000000000564.
- [21] Shahar E, Shahar DJ. Causal diagrams, information bias, and thought bias[J]. *Pragmat Obs Res*, 2010, 1:33-47. DOI: 10.2147/POR.S13335.
- [22] Greenland S. Bayesian perspectives for epidemiologic research: III. Bias analysis via missing-data methods[J]. *Int J Epidemiol*, 2009, 38(6): 1662-1673. DOI: 10.1093/ije/dyp278.
- [23] Shahar E. Causal diagrams for encoding and evaluation of information bias[J]. *J Eval Clin Pract*, 2009, 15(3): 436-440. DOI:10.1111/j.1365-2753.2008.01031.x.
- [24] 郑英杰, 刘海燕, 于波, 等. 观察与实验:因果视角[J]. *中华流行病学杂志*, 2021, 42(10):1863-1870. DOI:10.3760/cma.j.cn112338-20201224-01437.
- Zheng YJ, Liu HY, Yu B, et al. Observation and experiment: causal perspective[J]. *Chin J Epidemiol*, 2021, 42(10): 1863-1870. DOI: 10.3760/cma.j.cn112338-20201224-01437.
- [25] Sackett DL. Bias in analytic research[J]. *J Chronic Dis*, 1979, 32(1/2): 51-63. DOI: 10.1016/0021-9681(79)90012-2.
- [26] Rothman KJ, Greenland S, Lash TL. *Modern epidemiology* [M]. 3rd ed. Philadelphia: Lippincott Williams & Wilkins, 2008.
- [27] Horwitz RI, Feinstein AR. Alternative analytic methods for case-control studies of estrogens and endometrial cancer [J]. *N Engl J Med*, 1978, 299(20):1089-1094. DOI:10.1056/NEJM197811162992001.
- [28] Goldstein BA, Bhavsar NA, Phelan M, et al. Controlling for informed presence bias due to the number of health encounters in an electronic health record[J]. *Am J Epidemiol*, 2016, 184(11): 847-855. DOI: 10.1093/aje/kww112.
- [29] Phelan M, Bhavsar NA, Goldstein BA. Illustrating informed presence bias in electronic health records data: how patient interactions with a health system can impact inference[J]. *EGEMS (Wash DC)*, 2017, 5(1): 22. DOI: 10.5334/egems.243.
- [30] Mangione-Smith R, Elliott MN, McDonald L, et al. An observational study of antibiotic prescribing behavior and the Hawthorne effect[J]. *Health Serv Res*, 2002, 37(6): 1603-1623. DOI:10.1111/1475-6773.10482.
- [31] Brakenhoff TB, Mitroiu M, Keogh RH, et al. Measurement error is often neglected in medical literature:a systematic review[J]. *J Clin Epidemiol*, 2018, 98:89-97. DOI:10.1016/j.jclinepi.2018.02.023.
- [32] Lash TL, Fox MP, Fink AK. *Applying quantitative bias analysis to epidemiologic data*[M]. New York: Springer, 2009. DOI:10.1007/978-0-387-87959-8.
- [33] Corbin M, Haslett S, Pearce N, et al. A comparison of sensitivity-specificity imputation, direct imputation and fully Bayesian analysis to adjust for exposure misclassification when validation data are unavailable[J]. *Int J Epidemiol*, 2017, 46(3):1063-1072. DOI:10.1093/ije/dyx027.