

·基础理论与方法·

观察与实验:因果视角

郑英杰¹ 刘海燕² 于波³ 胡安群² 董遥¹ 郭阳³¹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(复旦大学)/复旦大学公共卫生学院公共卫生安全教育部重点实验室,上海 200032;²安徽省安庆市立医院 246003;³北京大学深圳医院 518036

郑英杰、刘海燕和于波对本文有同等贡献

通信作者:郑英杰,Email:yjzheng@fudan.edu.cn

【摘要】 观察和实验及其设计、相关的概念和内涵仍然模糊,影响着研究设计的正确认识及因果推断有效性的判断。本文借用物理学的相变概念,结合因果思维与因果图,首先建立了事物的属性及其状态、事件与现象之间的关系;接着识别了两种现象获取方式——人为观察和人为实验之间相反的因果结构,及其与暴露的干预、分配在因果推断研究中受各自的因果机制影响;最后,识别了干预系因果概念,是自然因果律基础上联系已知的现象和/或需经测量而获得的现象之间的核心纽带;基于此分析了研究设计二分类的策略,以干预法和非干预法的分类更为清晰。观察和实验提供了所有科学知识的基础,应是内涵统一的概念。基于因果律及其测量过程可能是实现研究设计精准分类的最佳选项之一,值得深入研究。

【关键词】 因果思维; 观察; 实验

基金项目:国家自然科学基金(81373065,81773490)

Observation and experiment: a causal perspectiveZheng Yingjie¹, Liu Haiyan², Yu Bo³, Hu Anqun², Dong Yao¹, Guo Yang³¹Department of Epidemiology/Key Laboratory for Health Technology Assessment, National Health Commission/Key Laboratory of Public Health Safety, Ministry of Education, School of Public Health, Fudan University, Shanghai 200032, China; ²Anqing Municipal Hospital of Anhui Province, Anqing 246003, China; ³Peking University Shenzhen Hospital, Shenzhen 518036, China

Zheng Yingjie, Liu Haiyan and Yu Bo contributed equally to the article

Corresponding author: Zheng Yingjie, Email: yjzheng@fudan.edu.cn

【Abstract】 Observation and experiment and their related connotations and concepts remain vague, which affects the correct understanding of research design and the judgment of the validity of causal inference. This article borrows the concept of phase transition in physics, combines causal thinking and causal diagrams, firstly establishes the relationships among the attribute, state, event, and phenomenon, and then identifies two ways with the opposite causal structures to acquire phenomena—human observations and human manipulated experiments. In causal inference, the ways mentioned above, intervention and assignment of exposure are affected by their own causal mechanisms. Finally, intervention is a causal concept, a core link among known phenomena, unknown phenomena available for measurement, and natural causality. Based on this, the two strategies in classifying research design are analyzed, and intervention method and non-intervention method are proposed, as is comprehensive and concise. Observations and experiments provide the basis for all scientific knowledge and should be viewed as concepts with a unified connotation. The accurate classification of research designs based on the law of causality and measurement process

DOI: 10.3760/cma.j.cn112338-20201224-01437

收稿日期 2020-12-24 本文编辑 万玉立

引用本文: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.



may be one of the best options worthy of in-depth study.

【Key words】 Causal thinking; Observation; Experiment

Fund programs: National Natural Science Foundation of China (81373065, 81773490)

观察(observation)和实验(experiment)提供了所有科学知识的基础数据^[1]。“观察”的概念是自明的;而“实验”通常具有多重含义,如在实验室或野外进行生物标本的检测、干预试验和现场试验等;“实验”常与“试验”一词混用,前者系通过实践活动以检验理论或假设,而后者则是通过尝试性活动以检验事物(object)的效果或性能^[2]。

观察和实验是流行病学的核心概念,通常认为有其特定的含义。然而,《流行病学词典》只定义了“观察法”,而无“观察”的定义^[3],以不干扰自然规律为特征,如研究者对自然现象的直接考察和研究^[4];《流行病学词典》将“实验”定义为实验法的核心,主要内容是对暴露和混杂子的控制^[3]。由观察和实验自然延伸为研究设计的二分类,即观察性研究(或观察法,observational studies)和实验性研究(或实验法,experimental studies),或经典利氏二分类法^[5]。其依据是暴露的分配人为控制与否:若是,即实验法;若否,即观察法,并有意地将“未随机化”一词列于观察法之下^[5];唐金陵等^[4,6]亦认为二分类法的根本区别在于是否采取随机化分组,而与干预无关。这些信息传递出“二分类法与是否随机化紧密联系”的观点。

然而,在流行病学史上,不少“实验法”经典案例均难以判断其是否满足“随机化”这一条件,如天花疫苗接种与天花发病流行的关系^[7]、12名水手坏血病的研究^[8]、维生素B₁缺乏和脚气病之间的关系等^[9]。即便是广为流传的伦敦水厂与霍乱流行之间关系的研究^[10],却被不同知名学者分别判定为“自然实验”^[11]、观察法^[12]、队列研究^[13]等。同一研究却被判定为不同的设计并不少见^[14]:涉及疾病诊断和治疗的筛检,是观察和干预的综合,却常被归为观察法^[15];实验法常常被认为等同于干预法,其中涉及的控制、实施、干预等诸多概念仍然含糊。

综上,观察(法)和实验(法)及其相关概念上不明确,特别是随机分组是否为其特征性区别以及实验法是否优于观察法,有待于进一步澄清。本文从因果思维的角度出发,结合因果图,基于观察和实验的通俗含义所反映的具体过程^[12,16],探讨这些概念的确切含义及其可能的价值。为阐述方便,本文假定变量的真实值以大写字母表示,其对应的测量

值以同名字母加上标“*”表示,变量的下标代表不同时间点同一变量值,并假定无测量误差和选择性偏倚。因混杂系自然固有的现象,参照 Pearl 的因果图法^[17],以下均以“双箭头虚线”来表示。

1. 事物的属性和现象:

(1)事物的属性:状态的基础。自然界中存在的一切为物、描述的一切为事。任一具体的物体总是具有许多不同的性质,即物体的属性(attribute);若物体的某一属性持续存在,则意味着该物体处于该属性的状态(state)。物体属性的产生和发展有其固有的规律,其所处于的状态是人们认识物体的基础;对物体属性的认识是人类有意的行为,通过对物体属性的描述而实现。因此,事物是客观存在的一切事和物,对事的描述加深了人们对物的认识。掌握事物产生和发展的规律,既是人类认识自然和改造自然的需要,又是推动人类文明发展的动力。

物质(substance)是物体属性的基础。物理学中,以“相”来指代物质的状态,譬如水分子(H₂O)有固、液、气相等。相变(phase transition)是物质两种不同状态之间的转变,如固态冰融化成液态水、液态水经加热而成为水蒸汽(图1)。与此类似,事物属性从一种状态向另一种状态的相变,即意味着事物发生了某一事件(event),如张三因出现第一个细胞癌变而发生从非病理状态向病理状态的相变,李四因感染某一病原体而发生了从健康状态向疾病状态的相变,王五的血压值因被测量而发生了从未知状态向已知状态的相变等。这种相变(或事件的发生)通常是瞬间的,或可以说,相变具有时点(point time)——没有时间长度的特征。事件发生后,该事物即拥有已发生事件所具有的相应属性;若该属性持续一段时间,即意味着该事物处于相应的状态。状态持续时间因事物的不同属性而有不同的名称,如疾病状态持续时间即病程(duration of disease),张三患病后存活的时间即生存期,易感的李四因接触某一病原体开始直至出现疾病启动而使得其易感状态结束的这段时间即潜伏期。因此,状态具有时期(period time)——有时间长度的特征,这为我们认识这种状态提供了必需的时间。随着时间的推移,一种状态可能再次出现相变,如疾

病经治疗后痊愈表现为从有病状态向无病状态的相变,经接种乙型肝炎(乙肝)疫苗而获得的对 HBV 的免疫力随着时间推移而消失。因此,事物的属性,既是前一事件发生后的终点状态,又是后一事件发生前的起点状态;事件的发生是两个不同状态间发生相变的瞬间结果,是相变后状态的起点。事件的发生是两个不同状态相变的结果,因此探测事件需要明确相变前后的两个状态。这意味着:唯有时期方可探测并识别事件;而时点和时期均能探测并识别状态。如图 1 所示,时点 t_1 可探测水的液相, t_2 可探测水的固相,而时期(t_1, t_2)可探测水的液相、固相及其相变。综上,事物属性的状态是我们认识事物属性的基础。

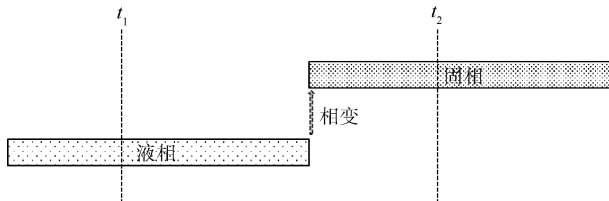


图 1 事件是两个不同状态的相变:以水分子为例

(2)事物的现象:状态的测量。当事物处于某一属性的状态并能被人们所“感知”或测量时,即为该属性的现象(phenomenon)。科学研究中,通过感官可以感知到事物的状态,如看到个体脸部发黄,听到咳嗽声,闻到口腔异味,摸到隆起的肿块等;此外,现象也可通过一定的仪器、设备和方法等各种手段测量而获知,如高血压、肿瘤、糖尿病等不易“感知”的状态可通过其相应的测量手段得以诊断,利用 HBsAg 标准品可标定乙肝疫苗中 HBsAg 的含量。通过对事物属性的测量,获得其现象,从而实现认识事物的需要。因此,通常认为,科学研究始于对事物属性的测量——现象的认识。

2. 现象的获取方式:包括人为观察(基本结构、非干预型结构)和人为实验(基本结构、非干预型结构、干预型结构)。见图 2。

(1)人为观察:因事物的现象是事物属性经测量而获知的状态,故属性首先必须存在,方能通过测量获得其测量值(或现象),如 2020 年春季武汉市出现许多咳嗽(已被看到或听到)、发热(经体温计测量)等患者,经检测后明确诊断为新型冠状病毒感染,从而确立了新型冠状病毒肺炎暴发(现象)。通过测量,一方面可以判断事物的属性是否存在,进而提示了这个属性已经发生,并表现为确切的先后时间顺序(或时序, time order),即:先出现

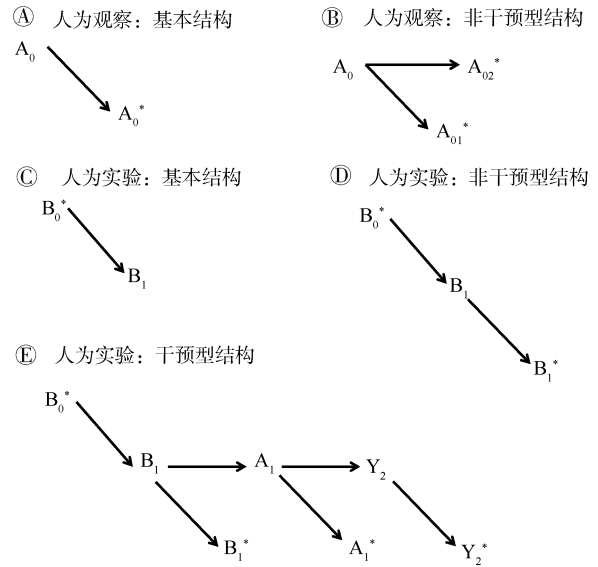


图 2 人为观察和人为实验结构示意图

事件的发生,接着处于事件所具有的状态,最后经测量而获知现象。

以上阐述了我们通过现象认识大自然的第一种方式(图 2A)——人为观察(基本结构)^[18];表示为 $A_0 \rightarrow A_0^*$,即在事物属性的真实值(A_0)虽然未知但已存在的自然情形下,人们对该事物属性 A_0 进行测量并获得其测量值(A_0^*)的过程,即:通过人为观察的方式实现了对事物属性从未知向已知的相变,获得了属性 A_0 的替代值 A_0^* 。

若两个不同研究者对同一属性(A_0)进行观察(图 2B),即 $A_0 \rightarrow A_{01}^*$ 和 $A_0 \rightarrow A_{02}^*$,则可以进行 $A_{01}^* - A_{02}^*$ 间的关系研究。例如, A_{02}^* 采用金标准法获得,而 A_{01}^* 采用评估法获得,即相当于进行评估法的有效性评价;若 A_{01}^* 和 A_{02}^* 系同一属性经不同方法测量的结果,即相当于进行不同方法间的重复性(或一致性)研究。这种思维同样地适用于不同时间内、不同人群间等对同一属性的比较研究。因 A_{01}^* 和 A_{02}^* 均为同一属性(A_0)的替代者,研究涉及同一属性的不同现象间关系,而不涉及一个属性改变另一属性的情形,因此将这类获取现象的过程定义为非干预型人为观察。

(2)人为实验:经人为观察获得已知现象,为人类改造自然奠定基础。改造意味着改变,人类将一个已知现象施加于事物或事物接受这一已知现象的过程,即人为实验(图 2C)。人为实验有两个目的:①观察这个已知现象在事物上产生与该现象对应的真实值的测量值。例如,慢性乙肝患者服用拉米夫定(已知现象:已知剂量、用药途径和疗程等, B_0^*)后,我们并不知道在患者体内发挥作用的、真

实的拉米夫定剂量(B_1)有多少,这可因个体不同特征(如生理状态、食用食物或服用其他药物等)而异,但我们知道真实剂量(B_1)受到患者服用剂量(B_0^*)的影响,即 $B_0^* \rightarrow B_1$ (图 2C)。对 B_1 进行人为观察,获得其测量值 B_1^* ,即 $B_1 \rightarrow B_1^*$,从而实现人为实验的第一个目的(图 2D),这在毒物或药物的代谢动力学研究、外暴露与内暴露间的关系研究中常见。②观察这个已知现象施加于事物后,改变事物另一属性的情况(图 2E)。上例中,慢性乙肝患者服用拉米夫定后,虽然可以不测量或忽略患者体内真实发挥作用的拉米夫定剂量(B_1),但我们知道该变量(B_1)将降低慢性乙肝患者体内的 HBV 含量(A_1),并进而改变患者发生肝癌的风险(Y_2),后者经人为观察而获得(Y_2^*),即 $B_1 \rightarrow A_1$ 、 $A_1 \rightarrow Y_2$ 和 $Y_2 \rightarrow Y_2^*$ 。可见,人为实验的这两个目的均在于获得效应估计^[19]。

以上,我们阐述了认识或改造自然的第二种方式——人为实验。上述已知的现象(B_0^*)即人为实验措施,这种措施总是人为预先设置的^[6]。若研究目的是获得该措施施加于事物后其对应的同一属性将改变后续事物的另一种属性,如图 2E 中的 $B_1 \rightarrow C_1$,即干预(intervention),人为实验中所采取的具体方法即人为干预措施(intervention measure)。在上述慢性乙肝患者服用拉米夫定后改变肝癌发生概率的例子中,这种人为实验为干预型人为实验,其目的在于检验病因的第二个含义——因变致果变,以事物的一种属性改变其另一种属性、进而改变后者的后续效应,这可通过 B_0^* (或 B_1^* 或 A_1^*)与 Y_2^* 之间的关联来估计(图 2E)。若研究目的是获得该措施施加于事物后对应真实值 B_1 的替代者(B_1^*),而不涉及一个属性改变另一属性的情形,我们可将其命名为非干预型人为实验。

综上,科学研究总是一门测量的学问^[20],是一个以获取的现象替代属性的测量过程。我们界定了人类获得现象的两种方式——人为观察和人为实验:人为观察是在事物的属性(A_0)已存在的前提下,回顾性观察历史的属性,通过测量该属性而获得其现象(A_0^*)的过程,即 $A_0 \rightarrow A_0^*$,是人类认识自然的基础;而人为实验是在现象已知(B_0^*)的前提下,通过给事物实施该现象,即 $B_0^* \rightarrow B_0$,实验的影响具有前瞻性,是人类认识及改造自然的基础。

3. 自然实验及其测量:观察总是人为的,而实验则可分为人为实验和自然实验。自然界中,事物一个属性(如暴露, A)影响着另一个属性(如结局,

Y)或因果关系,即通过以未知的暴露 A 的真实值施加于个体,进而改变未知的结局 Y 的真实值,即自然实验(natural experiment)。显然,自然实验总是干预型,其中的干预是大自然赋予的。

围绕特定的暴露与结局间关系,大自然影响暴露属性的分配有两种方式,即混杂变量(双箭头虚线,图 3A)和工具变量(IV,图 3B)^[21-23]:前者是大自然为人类设置的障碍,促使现有设计的首要目标是控制混杂;而后者则是大自然馈赠人类的礼物,是控制前者的办法之一^[24],也是观察法所期望的^[25-26]。由此可见,大自然是事物发生发展的真正启动者,在暴露影响结局的自然发生发展过程中,自然实验及其中暴露的自然干预和自然分配是基于因果律的统一体。

在自然因果律规律的基础上,人类通过各种手段和方法,以认识并达到改变自然因果律为最终目的。因果律为研究所感兴趣;若要为我们所知,则需要通过对代表因果律的变量进行测量,以获得其相应的替代值。因此,认识因果律是人为观察和/或人为实验的结果。从因果推断的角度来看,得到结局 Y 的替代值只有一种途径——人为观察而获得,也就是 $Y \rightarrow Y^*$ ^[27-28],即结局需先发生、接着存在,经过测量,方为我们所知。然而,暴露 A 替代值的获取有两种方法,即上述的人为观察和人为实验。

(1)人为观察:因自然实验与暴露的自然分配和自然干预是统一的,所有变量均可经人为观察而获得。对于暴露的分配,若是仅在混杂子的影响下(图 3A),则这种情形为自然干预和自然非随机化分配,通过观察获得 A_0 和 Y_2 的测量值,研究 A_0^* - Y_2^* 间的关联以估计 A_0 - Y_2 间的效应;然而,这种关联受到混杂结构的影响(图 3 中虚线表示),需通过有效控制混杂而实现正确的效应估计,这是因果推断无法避免的常态现象^[27]。若在混杂结构上叠加了工具变量对暴露分配的影响(图 3B),经观察可获得 IV 的测量版(IV^*),然而 A_0^* - Y_2^* 间的关联估计仍然无法逃离混杂子的影响。虽然这样,在此情形下,可通过估计 IV^* - A_0^* 间和 IV^* - Y_2^* 间的关联,从而绕过可能未知或未测量的混杂子的影响,以间接获得 A_0 - Y_2 间的正确效应估计值。与图 3A 不同,暴露的分配因 IV 的影响而表现为自然随机分配,因此这种结构可被称为自然干预和自然随机化的人为观察(图 3B),如以基因变异为工具变量的孟德尔随机化设计^[22,29]。令人惊讶的是,在人类认识自然的过程中,大自然既设置障碍,又提供解法,寻找合适

的工具变量成为科学研究的梦想^[26]。

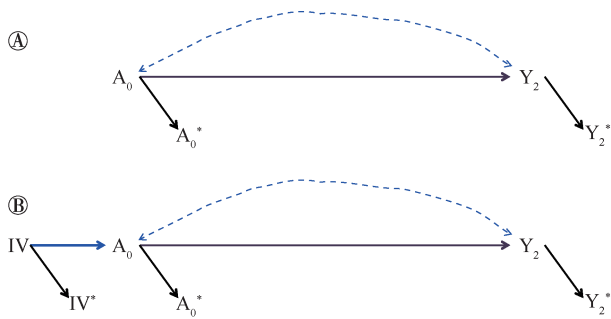


图3 人为观察法:自然分配的两种情形及其测量

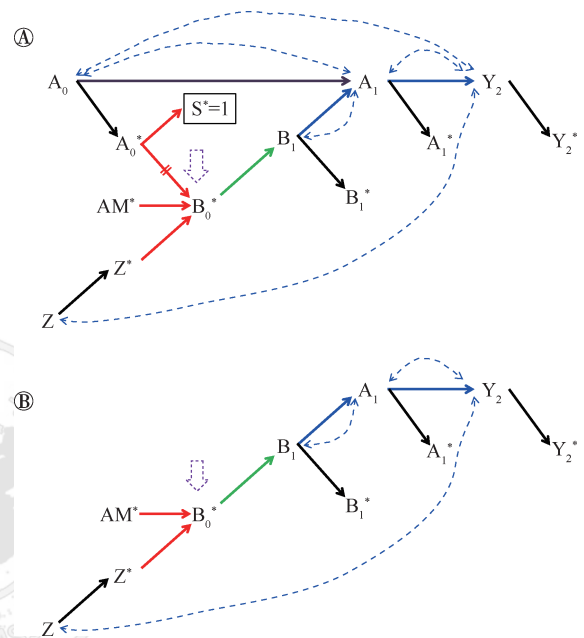
(2)人为实验:在自然实验的基础上进行干预型人为实验,方能达到人类改造自然的目的。通常,实施这类实验的前提是,暴露影响结局(如图2E中的 $B_1 \rightarrow A_1$)必须成立或得到足够的证据支持,这样才能通过人为实验施加某一措施(B_0^* ,图4),经 B_1 达到改变暴露(A_1),进而影响暴露致结局的自然进程;否则,实施人为实验将是无效的。

值得指出的是,人为实验施加的措施通常是已知的,所施加的人群通常是同质的,如已知剂量和疗程的拉米夫定(B_0^*)只用于慢性乙肝患者而不对非患者使用(否则违背伦理),已知活性HBsAg抗原量的乙肝疫苗的接种通常只针对乙肝易感者而不对乙肝免疫者实施,已知剂量和疗程的二甲双胍只用于2型糖尿病患者治疗。

因干预型人为实验拟实施的人群是同质的,故在实施前需进行研究人群的选择或同质化。仍以拉米夫定治疗慢性乙肝以预防肝癌的效果评价研究为例(图4A),首先需进行人群的调查或医院招募(筛查和诊断)以选择合格的、已明确诊断的慢性乙肝患者,即 $A_0 \rightarrow A_0^*$;接下来,诊断为慢性乙肝者($A_0^*=1$)将被纳入研究($S^*=1$),而非慢性乙肝者($A_0^*=0$)不需要进行拉米夫定治疗,排除出研究($S^*=0$)。最后,慢性乙肝者($A_0^*=1$)将被分组,例如:一组接受拉米夫定治疗($B_0^*=1$),另外一组接受对照治疗方法($B_0^*=0$)。

接受拉米夫定治疗与否的分配机制由人为分配方法($AM^*=1$,随机分配; $AM^*=0$,非随机分配)和分配方法外的其他现象($Z^*=1$,是; $Z^*=0$,否;为阐述方便,以二分类来表示)来决定:若随机分配($AM^*=1$),则 B_0^* 完全由其决定,这将导致 $Z^* \rightarrow B_0^*$ 的路径消失,则构成了人为分配的随机实验;否则为人为分配的非随机实验,即 $AM^*=0$ 时, Z^* 对 B_0^* 的影响不能消除。

实践中,因研究始于干预措施的施加时刻,故研究人群的招募和同质化(或筛选和选择)的过程被忽略,图4A可进一步简化为图4B;若 B_0^* 的分配经随机分组而确定,则其分配过程可进一步被忽略,从而图4B中 B_0^* 左侧的变量不显示;有时只突出与关联计算有关的 B_0^* 和 Y_2^* 而忽略干预过程,从而也不显示 $B_1 \rightarrow A_1$ 的结构。



注:蓝色单箭头代表大自然实验;当涉及暴露时,也是暴露的自然分配和干预;蓝色虚线双箭头代表箭头指向变量间因果律受到混杂的影响;黑色单箭头代表人为观察;绿色单箭头代表人为实验;红色单箭头代表推理;紫色虚线框单箭头代表研究起点;红色双斜杆代表着当研究人群同质化后,因 $A_0^*=1$ 、只有一个数值, $A_0^* \rightarrow B_0^*$ 之间的单箭头消失

图4 人为实验:人为分配及其测量

4. 研究设计分类:两种策略的对比。以上,我们首先建立了现象获取方式——人为观察和人为实验的过程;当涉及因果推断时,基于自然实验/因果律,应用上述两个过程分析其在暴露的分配和干预中的作用,从而得出:①非干预研究不涉及分配机制,以单一属性的现象或现象间关系为研究目的。②干预研究必然涉及暴露的分配,是否随机化系自然或人为的选择,而与现象获取的方式无关,以探讨属性间因果律为目的。人为观察系在自然实验/干预/分配的基础上进行,而人为实验必然涉及人为的暴露分配,并需结合人为观察,以达到因果推断的目的。不论是何种因果推断研究,干预有其自然因果律为基础,如图3中的 $A_0 \rightarrow Y_2$ 和图4中的 $B_1 \rightarrow A_1$ 的结构。因此,干预系因果概念^[17],在此

基础上,获取与干预过程相关变量的替代者,从而实现我们认识的目的,这为研究设计的分类奠定基础。

基于现象获取方式(观察/实验)和干预与否的先后顺序,我们探讨了两种策略进行研究设计的二分类,即人为观察法和人为实验法(延续传统思维),或非干预法和干预法(是否因果推断),并与利氏二分类法进行了比较(图 5)。两种分类法均全面地包括了非干预和干预研究,其中增补了非干预研究和自然随机实验(上述 IV 的情形,图 3B),并明确地指出了非干预人为实验中涉及的是现象间因果关系(以属性变量为中介子,图 4A 和图 4B 中的变量 B_1),这些均是利氏二分类法所不具备的。同时不难发现,利氏二分类法主要对以因果推断为目

标的研究进行分类,而忽略非因果律的研究。因干预的因果性基础,以及干预必须结合暴露分配方式的特点,以干预与否为起点,首先直接涉及不同研究的目的,接着涉及现象获取方式,因此第二种分类法更为清晰(图 5B)。

5. 讨论:借用物理学的相变概念,本文首先建立了事物的属性及其状态、事件与现象之间的关系,接着识别了人为观察和人为实验不同的因果结构,以及其与暴露的干预、分配在因果推断研究中分别受各自的因果机制影响,并指出以干预法和非干预法的二分类法较佳。

观察总是人为的,系历史的属性从未知到已知的相变,需经测量而实现获得并认识现象;而人为实验系已知现象到未知的未来属性的相变,所施加

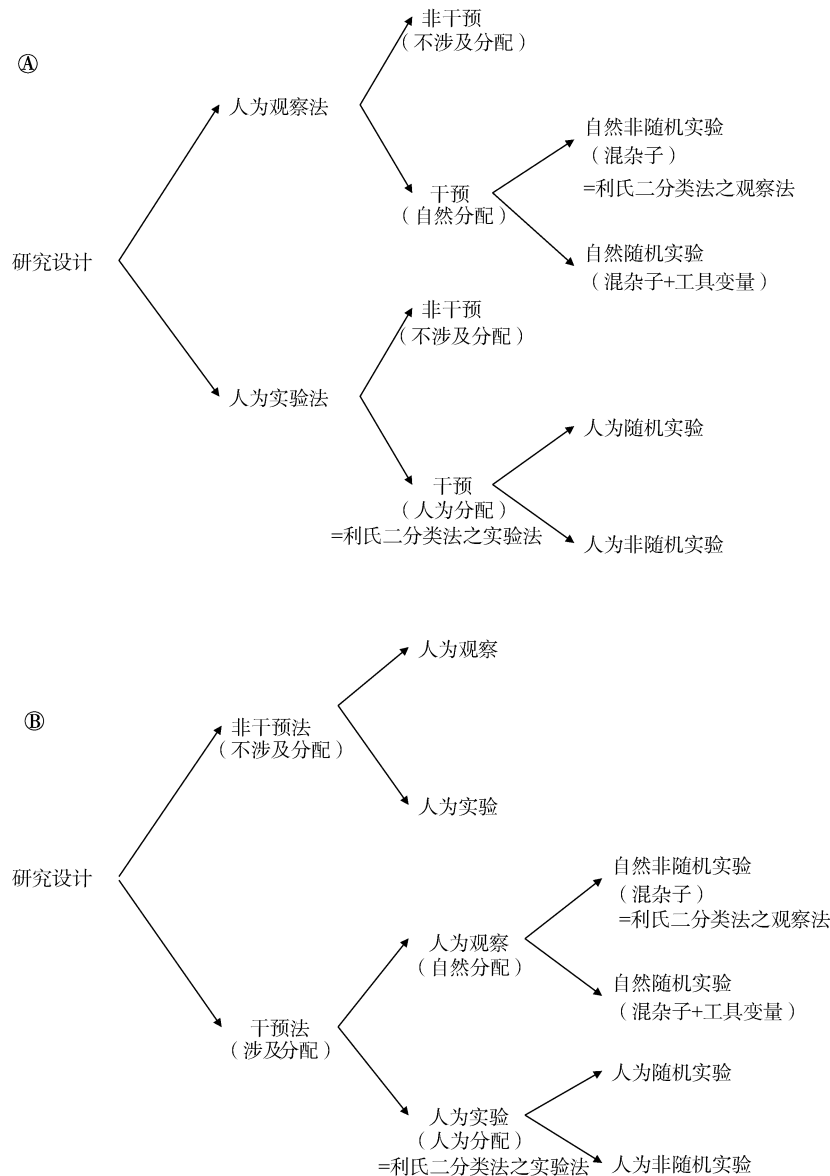


图 5 研究设计的新二分类:两种策略的对比及其与利氏二分类法的比较

的现象是已知的、无需测量。以属性间是否相变界定了干预,确立了其因果性基础^[17],明晰地区分出两类研究目标:非干预和干预。非干预研究不涉及分配,其中:关注单一属性的单一现象的现状,或单一属性的不同现象间关系,后者实质上是以该属性为混杂子影响现象间的关系,系我们研究所需要的,这可通过人为观察来实现;而人为实验涉及单一属性的不同现象间关系,系以该属性为中介子的因果结构,同样地需结合人为观察来实现。

干预研究以因果推断为目标,干预的因果性奠定了其作为自然实验和人为实验的基础,从而起到将干预结构、相关变量及其对应测量值变量统一于因果框架下的核心纽带作用。不论是自然实验还是人为实验,这个干预的结构需假设是成立的或得到足够证据的支持。实践中,以干预结构两侧变量的替代者来估计干预的效应:对于自然实验,因果推断相关的自然变量需要通过人为观察后方可获知;而对于人为实验,除了暴露属性的替代者有两种情形:属性之因(已知的现象)和属性之果(需经观察而获得的),这为研究的灵活性提供条件,其余与自然实验类似。因此,干预与暴露、结局及干扰暴露与结局间关系的因素如何获得或测量无关,也不涉及暴露的分配。

因果推断以获得因果效应的正确估计值为目标,因此暴露的分配与混杂的控制密切相关。自然实验中暴露的自然分配若只由混杂子所致或人为实验中暴露以非随机方式分配,则因果推断的可交换性假设无法满足;实践中,研究设计阶段以混杂的控制为首要目标,条件随机化得以满足成为必须的假设(虽然永远不知道是否成立),从而得以实现因果效应的估计^[30],这是当前因果推断研究最为基本的假设之一。若在此基础上,自然实验中暴露的自然分配叠加了自然工具变量的影响(自然随机实验),或人为实验中经由人为随机化(相当于设置人为工具变量)确定暴露的分配(人为随机实验),则二者均可确保因果推断中正性假设(assumption of positivity,即不同混杂子层的暴露比例在0~1的开区间)和可交换性假设(assumption of exchangeability,即暴露组发生结局的风险和非暴露组在暴露的情况下发生结局的风险一致)的成立^[30],达到有效控制混杂的目的。因此,不论是人为实验还是自然实验,随机化仅是暴露分配方式之一,取决于研究者或自然的选择。研究者拟通过设计随机化人为实验以期得到一个正确的效应估计

值;然而,这仅仅反映了研究实施之前研究者实现了组间可比性的期望,这需要研究实施的完美过程方能得以持续,若出现诸如未实施盲法、霍桑效应、检出偏倚、测量偏倚、失访等,都将使得这个期望化为泡影。

传统利氏二分类法及其扩展版^[5,15,31-32],从测量的角度出发,主要围绕以因果推断为目的、人为控制的研究进行分类,有时实验法又区分出类实验。本文从自然因果律出发,厘清了状态与事件、现象、人为观察和人为实验、干预、分配等概念,为认识研究设计的二分类铺平道路。两种分类策略均包括了全部的非干预和干预研究(图5),其中:增补了非干预研究和自然随机实验,并明确地指出了非干预人为实验中涉及的是现象间因果关系。因干预的因果性,以及干预必须结合暴露分配方式的特点,以干预为起点直接涉及研究目的、接着涉及现象获取方式的第二种分类策略,更为清晰。由此容易判断,观察法并非不涉及随机化,实验法不全是随机化,干预不等于实验,实验法并不优于观察法等^[33-35]。

综上所述,观察和实验提供了所有科学知识的基础,应有统一的内涵,研究设计离不开人为观察和人为实验。基于自然因果律,本文厘清了人为观察和人为实验及其概念,并提出以干预法和非干预法的二分类法较佳。基于大自然的因果律及其测量过程可能是实现研究设计精准分类的最佳选项之一,值得深入研究。

利益冲突 所有作者均声明不存在利益冲突

志谢 复旦大学公共卫生学院李意杰、阙慧、张森、牟育彤、曹焱敏、王凯琳对本文成稿提供帮助

参 考 文 献

- [1] Woodward RS. Observation and experiment[J]. Ann New York Acad Sci, 1901, 14(1): 69-84. DOI: 10.1111/j.1749-6632.1901.tb55051.x.
- [2] 唐慈富. "临床实验"? "临床试验"! [EB/OL]. (2020-04-24) [2020-12-24]. http://blog.sina.com.cn/s/blog_4c0596f60102ynob.html.
Tang CF. "Clinical experiment"? "clinical trial"! [EB/OL]. (2020-04-24) [2020-12-24]. http://blog.sina.com.cn/s/blog_4c0596f60102ynob.html.
- [3] Porta M. A dictionary of epidemiology[M]. Oxford: Oxford University Press, 2014.
- [4] 唐金陵, 杨祖耀. 观察与实验:效力与效果[J]. 中华流行病学杂志, 2014, 35(3): 221-227. DOI: 10.3760/cma.j.issn.0254-6450.2014.03.001.
Tang JL, Yang ZY. Observation versus experiment: efficacy versus effectiveness[J]. Chin J Epidemiol, 2014, 35(3):

- 221-227. DOI:10.3760/cma.j.issn.0254-6450.2014.03.001.
- [5] Lilienfeld AM, Lilienfeld DE. Foundations of epidemiology [M]. Oxford: Oxford University Press, 1980.
- [6] 唐金陵. 再论观察与实验: 大数据现实世界研究不能取代随机对照试验[J]. 中华流行病学杂志, 2018, 39(8):1121-1124. DOI:10.3760/cma.j.issn.0254-6450.2018.08.021.
- Tang JL. Revisit to observation and experiment: real world study cannot replace randomized controlled trial[J]. Chin J Epidemiol, 2018, 39(8):1121-1124. DOI:10.3760/cma.j.issn.0254-6450.2018.08.021.
- [7] Best M, Neuhauser D, Slavin L. "Cotton Mather, you dog, dam you! I'll inoculate you with this; with a pox to you": smallpox inoculation, Boston, 1721[J]. Qual Saf Health Care, 2004, 13(1):82-83. DOI:10.1136/qshc.2003.008797.
- [8] Blair JS. James Lind[J]. J Roy Army Med Corps, 2007, 153(1):69. DOI:10.1136/jramc-153-01-15.
- [9] Pietrzak K. Christiaan Eijkman (1856-1930) [J]. J Neurol, 2019, 266(11):2893-2895. DOI:10.1007/s00415-018-9162-7.
- [10] Snow J. On the mode of communication of cholera[J]. Edinb Med J, 1856, 1(7):668-670.
- [11] MacMahon B, Pugh TF. Epidemiology: principles and methods[M]. Boston: Little, Brown and Company, 1970.
- [12] Hill AB. Observation and experiment[J]. N Engl J Med, 1953, 248(24):995-1001. DOI: 10.1056/nejm195306112482401.
- [13] Ahrens W, Pigeot I. Handbook of epidemiology[M]. New York: Springer, 2014.
- [14] 朱以敏, 李亚欣, 牟育彤, 等. 暴发调查设计: 以急性戊型肝炎暴发为例[J]. 中华疾病控制杂志, 2020, 24(11):1241-1245, 1347. DOI:10.16462/j.cnki.zhjbkz.2020.11.001.
- Zhu YM, Li YX, Mu YT, et al. The designs for outbreak investigation: an example of acute hepatitis E outbreak[J]. Chin J Dis Control Prev, 2020, 24(11): 1241-1245, 1347. DOI:10.16462/j.cnki.zhjbkz.2020.11.001.
- [15] Bhopal R. Concepts of epidemiology: an integrated introduction to the ideas, theories, principles and methods of epidemiology[M]. Oxford: Oxford University Press, 2002.
- [16] Sansone EB, Slein MW. Application of the microbiological safety experience to work with chemical carcinogens[J]. Am Ind Hyg Assoc J, 1995, 37(12):711-720. DOI:10.1080/0002889768507553.
- [17] Pearl J. Causality, models, reasoning, and inference[M]. 2nd ed. Cambridge: Cambridge University Press, 2009.
- [18] 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.
- [19] Merriam-Webster. Learner's dictionary[EB/OL]. [2020-12-17] [2020-12-24]. <https://www.learnersdictionary.com/>.
- [20] Codd JA. Experiment and observation in medicine[J]. Br Med J, 1910, 2(2598): 1111-1115. DOI: 10.1136/bmj.2.2599.1287-a.
- [21] Miettinen OS. Theoretical epidemiology[M]. New York: Wiley, 1985.
- [22] 郑英杰, 赵耐青. 有向无环图: 语言、规则及应用[J]. 中华流行病学杂志, 2017, 38(8):1140-1144. DOI:10.3760/cma.j.issn.0254-6450.2017.08.029.
- Zheng YJ, Zhao NQ. Directed acyclic graphs: languages, rules and applications[J]. Chin J Epidemiol, 2017, 38(8): 1140-1144. DOI: 10.3760/cma.j.issn.0254-6450.2017.08.029.
- [23] Greenland S. An introduction to instrumental variables for epidemiologists[J]. Int J Epidemiol, 2000, 29(6):1102. DOI:10.1093/oxfordjournals.ije.a019909.
- [24] 何一宁, 刘丽丽, 蔡倩莹, 等. 研究设计时混杂控制策略的结构分类[J]. 中华流行病学杂志, 2018, 39(7):999-1002. DOI:10.3760/cma.j.issn.0254-6450.2018.07.025.
- He YN, Liu LL, Cai QY, et al. A structural classification of strategies for confounding control in research design[J]. Chin J Epidemiol, 2018, 39(7): 999-1002. DOI: 10.3760/cma.j.issn.0254-6450.2018.07.025.
- [25] Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease[J]. Epidemiology, 2008, 19(6): 766-779. DOI:10.1097/EDE.0b013e3181875e61.
- [26] Hernán 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.
- [27] 郑英杰, 赵耐青, 何一宁. 客观世界的因果关系: 基于有向无环图的结构解析[J]. 中华流行病学杂志, 2018, 39(1): 90-93. DOI:10.3760/cma.j.issn.0254-6450.2018.01.019.
- Zheng YJ, Zhao NQ, He YN. Causality in objective world: Directed Acyclic Graphs-based structural parsing[J]. Chin J Epidemiol, 2018, 39(1): 90-93. DOI: 10.3760/cma.j.issn.0254-6450.2018.01.019.
- [28] 王天雷, 牟育彤, 阚慧, 等. 基于大自然时间轴的测量时序分类法[J]. 中华流行病学杂志, 2020, 41(5):782-787. DOI: 10.3760/cma.j.cn112338-20190929-00711.
- Wang TL, Mu YT, Kan H, et al. A new classification of measured temporalities: based on the time axis in nature [J]. Chin J Epidemiol, 2020, 41(5):782-787. DOI:10.3760/cma.j.cn112338-20190929-00711.
- [29] Smith GD, Holmes MV, Davies NM, et al. Mendel's laws, Mendelian randomization and causal inference in observational data: substantive and nomenclatural issues [J]. Eur J Epidemiol, 2020, 35(2): 99-111. DOI: 10.1007/s10654-020-00622-7.
- [30] Hernán MA, Robins JM. Causal inference: what if[M]. Boca Raton: Chapman & Hall/CRC, 2020.
- [31] Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principle and quantitative methods[M]. Belmont, California: Lifetime Learning, 1982.
- [32] Rothman JK. Modern epidemiology[M]. Boston: Little, Brown and Company, 1986.
- [33] Silcocks P. Experiment and observation[J]. Lancet, 1994, 344(8939-8940): 1775-1776. DOI: 10.1016/S0140-6736(94)92916-5.
- [34] Herman J. Experiment and observation[J]. Lancet, 1994, 344(8931): 1209-1211. DOI: 10.1016/s0140-6736(94)90516-9.
- [35] Hughes, Edmund. Experiment and observation[J]. Lancet, 1994, 204(5276):782-783. DOI:10.1016/S0140-6736(01)39629-0.