

DOI:10.13288/j.11-2166/r.2026.14.005

思路与方法

自然实验在中医药疗效评价中的优势、挑战及实施路径

刘日荷¹, 刘庭秀¹, 刘凯¹, 孙源¹, 张晓寒¹, 刘建平^{2,3}, 刘兆兰¹✉

1. 北京中医药大学循证医学中心, 北京市朝阳区北三环东路11号, 100029; 2. 成都中医药大学附属医院循证医学中心; 3. 四川省中医药循证医学中心

[摘要] 自然实验是一种基于真实世界研究的因果推断研究设计, 其核心思路在于利用由外部因素所引发的、近似随机化的暴露变异, 评估干预措施或暴露因素对结局的影响, 尤其适用于无法实施随机对照试验(RCT)或存在伦理限制的研究情景, 并兼具因果推断能力与较好的外部真实性。总结自然实验的定义、基本特征与研究现状; 提出自然实验在中医药疗效评价中的适用场景, 如政策实施或变更评价、医院内部医疗实践差异评价、地区或机构间资源禀赋差异评价; 认为设计路径包括从中医药临床或政策问题出发、系统收集研究数据、对数据进行统计建模与效应分析、结合临床实际对效应值进行解读等; 并进一步讨论自然实验策略应用于中医药疗效评价的优势与挑战。

[关键词] 自然实验; 中医药疗法; 疗效评价; 真实世界研究

真实世界研究(real-world study, RWS)能够反映干预措施在日常医疗中的实际效果^[1-2], 具有较高的外部真实性, 可有效弥补随机对照试验(randomized controlled trial, RCT)的适用场景局限。自然实验(natural experiment)作为RWS中一类具有“准实验”特征的研究设计^[3-4], 可以对大规模医疗政策或干预措施的真实实施效果进行评价, 为真实医疗环境下评估中医药干预的疗效和安全性提供方向与视角。自1910年自然实验被提出^[5]以来, 已经被应用于西药治疗肝病^[6]、心血管与呼吸系统疾病^[7]、糖尿病^[8]等疾病的疗效与安全性评价。自然实验并非专为中医诊疗评价而设计, 但其利用真实世界暴露差异、关注干预实施情境和人群异质性的特点, 为评价具有辨证论治、个体化调整和复杂干预过程^[9]特征的中医诊疗措施提供了可借鉴的方法学思路^[10]。本文总结自然实验的定义、应用现状、实施路径, 以及在中医药应

用中的优势与挑战, 以期对中医药干预措施的疗效和安全性评价提供思路。

1 自然实验的特征、应用现状与方法学定位

1.1 自然实验的核心特征

自然实验是指在真实世界环境中, 利用由外部因素(如政策实施、环境变化、社会事件或自然灾害)所引发的近似随机化的暴露变异, 来评估某一干预或暴露因素对结局影响的因果推断研究设计^[3-4]。该方法起源于1854年John Snow对伦敦宽街霍乱疫情的经典调查^[11], 由流行病学家、经济学家和社会学家共同发展并广泛应用^[12-13]。

自然实验的核心特征在于其干预分配的近似随机性。自然实验中的“处理”并非由研究者主动施加, 而是利用自然发生的事件或政策变化所形成的分组, 这种分组方式被称为“准随机”。“准随机”指由于外部因素的作用, 在满足暴露组与对照组具有可比性的前提下, 个体进入暴露组或对照组的概率在理论上近似于随机。研究者可以基于这种“准随机”分组比较暴露组与非暴露组之间的结局差异, 在一定程度上模拟了RCT的分组机制^[14]。

基金项目: 国家自然科学基金(82374298); 北京中医药大学学科后备带头人支持项目(90010960920033); 北京市科技新星计划交叉合作课题(20240484546); 国家中医药管理局高水平中医药重点学科建设项目(zyyzdxk-2023249)

✉ 通信作者: lzl1019@163.com

1.2 自然实验的应用现状

在临床与卫生服务领域，自然实验多依托临床指南更新^[15]、医生诊疗倾向^[16]，或医疗政策变化^[17]等外部变异开展。尽管自然实验具有补充RCT适用场景的潜力，但尚未成为干预措施临床疗效评价的常用研究设计。近年来，大数据技术的应用突破了传统研究数据获取的局限，也推动着自然实验扩展至城市建设与环境健康^[18]等交叉领域，成为连接理论研究与实践应用的桥梁。一项比较北京市乡镇推行煤改清洁取暖政策与否对急性心肌梗死发病率影响的自然实验^[17]表明，该政策的推行可使乡镇内急性心肌梗死发病率下降6.6%，且在女性和老年人中下降更为显著。

1.3 自然实验在医学研究设计中的方法学定位

自然实验在证据等级体系中通常被视为中等等级别观察性证据^[19]。在常用的医学研究设计中，自然实验与队列研究、非随机对照试验的证据等级相近，但在设计逻辑与因果推断能力上存在本质差异，三项设计的比较见表1。自然实验的论证强度通常高于一般队列研究和非随机对照研究，但低于严格实施的RCT。自然实验的优势在于利用外部因素形成近似随机的暴露差异，使暴露组与对照组在个体选择因素上的差异相对减少，从而降低选择偏倚和内生性问题，增强因果推断能力。然而，自然实验并不具备人为随机分配和严格干预控制，仍可能受到残余混杂、同期事件和分组机制不完全外生等因素影响，因此其内部真实性通常低于RCT。RCT更适合回答理想控制条件下干预是否有效的问题，自然实验则更适合回答干预在真实世界中是否产生预期效果的问题^[20]。

2 自然实验应用于中医药疗效评价的实施路径

2.1 应用场景

在医院环境内，应用自然实验评价中医药干预

措施的疗效，需先识别与利用自然变异来源。根据变异来源的不同，自然实验的应用场景主要可分为以下三类。

一是政策实施或变更。政策实施或变更往往具有外生性和突发性，能直接影响医院的管理模式和临床行为，是自然实验的理想变异来源。已有研究^[21]将中国医疗卫生体制改革看作准自然实验的暴露因素，通过不同综合医改政策试点的时间和地域差异形成自然分组，探讨政策改革对居民健康的提升效应，发现政策改革显著提升了居民主观健康和客观健康。该研究从城乡差异、性别差异、年龄差异方面进行异质性分析，论证过程具有借鉴价值。以此类推，在施行中医药疗法的医疗机构内，同样可以探究中医药疗效在政策执行前后的变化，如医保支付方式改革、中医药干预措施服务价格调整等自然变异对中医药临床应用的影响。

二是医院内部医疗实践差异。医院内部固有的、非研究者人为干预所致的医疗实践差异，也能提供“准随机”变异来源。例如，同一家医院内，不同医师或科室可能因师承渊源、学术流派或个人经验的不同，对同一病证的辨证思路和处方用药存在显著且相对稳定的差异^[22]。若将同一疾病患者分配至科室内不同医师就诊的过程视为“准随机”，同时控制其他变量如医师年资、就诊频次等，则可比较不同医师用药习惯下的疗效差异，从而评估真实世界中不同中医药治疗方案的有效性。此类自然变异尤其适用于探索中医个体化治疗路径的实际应用效果。

三是地区或机构间资源禀赋差异。由于历史、文化、经济等原因，不同地区或不同级别的医疗机构在中医药服务能力和资源禀赋上存在天然差异，且各地区中医类医院的资源配置（如每千人中医床位数、执业中医师数、中药师数）存在明显不均

表1 自然实验与队列研究、非随机对照试验设计的对比

Table 1 Comparison of the Designs of Natural Experiments, Cohort Studies, and Non-randomized Controlled Trials

项目	自然实验	队列研究	非随机对照试验
研究类型	观察性研究	观察性研究	干预性研究
分组依据	外部事件导致的准随机	自然暴露/自选择	研究者主观非随机分配
是否主动分配干预	否	否	是
偏倚主要来源	混杂偏倚	选择偏倚、混杂偏倚	选择偏倚、测量偏倚
因果推断强度	中等	中等	中等偏低
优势	外部真实性高	可同时研究多种结局，计算发病率	操作简便
适用场景	疗效评价、政策评估、资源配置差异	长期暴露因素与疾病关系	探索性研究、可行性研究

衡。这种地域性的资源配置差异，为研究地域、医院级别等对中医药疗效的评价提供了自然实验条件。例如，可比较不同资源配置水平下，临床实践中实际发生的中医药干预措施对特定疾病的疗效差异，从而揭示资源配置差异对临床结局的潜在影响。

2.2 研究设计

在医院环境中开展中医药自然实验研究，可遵循以下设计路径，确保研究过程的规范性与结果的可解释性（详见附图1，请扫描本文二维码获取）。

首先，应从中医药临床或政策问题出发，明确研究对象、干预措施和目标结局，构建“准随机”实验。确定一个外生的、近似随机的自然变异来源（如中医药政策试点、中医特色科室间固有的诊疗差异），据此定义分组依据，确认暴露组、对照组及时间窗口，并核实患者分组过程是否受研究者干预，变异是否独立于患者病情与治疗意愿。

其次，系统收集研究数据。结合研究问题，设定主要疗效指标（如中医证候积分、事件发生率）及次要疗效指标。数据收集需全面涵盖干预前的基线特征、干预中的过程与安全指标，干预后的结局指标及关键混杂因素。基于电子病历、处方系统、医保结算、检验检查、随访记录和患者报告结局等数据建立研究数据库，并完整记录中医药干预的组成、剂量、疗程、加减变化和实际执行情况。

再次，对数据进行统计建模与效应分析。采用倾向得分匹配进行数据预处理，平衡组间基线，根据自然变异的类型选择相应的因果推断方法，如政策分批实施或前后变化可采用双重差分法，基于明确阈值的分配可采用断点回归法，医师诊疗倾向导致的治疗差异可采用工具变量法。

最后，结合临床实际对效应值进行解读，核实结局稳健性，并明确研究的局限性及证据的外推性。

2.3 数据分析要点

2.3.1 样本量计算 自然实验研究的样本量计算应基于该研究所采用的组间比较统计方法^[23]。步骤如下：根据结局指标的数据类型确定组间比较统计检验方法；设定显著性水准(α)与把握度($1-\beta$)；确定具有临床意义的效应值；计算得出样本量。

2.3.2 数据预处理 在自然实验研究中，为减少暴露组与对照组之间由非随机分配导致的基线不平衡和选择偏倚，研究者常采用倾向得分匹配（propensity score matching, PSM）作为数据预处理方法。通过构建一个综合性的倾向得分，用以概括多个已知混杂变量（如人口学特征、临床指标及病史

等）对处理分配的影响，从而将多维协变量信息压缩为单一维度得分^[24-25]。在收集足够的样本数据后，研究者根据得分对暴露组与对照组个体进行匹配，确保两组在匹配后的协变量分布上达到平衡，有效减少观察性数据中由于混杂因素带来的偏倚。在实际应用中，PSM通常可借助SPSS、R等统计软件实现其分析流程。

2.3.3 因果推断 1) 双重差分法：常用于处理观察性数据的未测量混杂因素^[26]。该法比较暴露组与对照组在自然变异前后结局指标的差值，以识别自然变异的效应。此法控制选择偏倚的有效性依赖于平行趋势假设^[27]，即在没有自然变异暴露的情况下，暴露组与对照组的结果随时间变化的趋势平行。此时对照组观察前后的变化，可视对暴露组在未受变异影响时潜在结果的模拟。将暴露组的实际变化减去对照组的变化，所得差值即可被解释为自然变异带来的因果效应。通过引入组别、时间虚拟变量及二者的交互项可估计这一效应^[28]。此法能同时控制不随时间变化的组间固有差异以及共同的时间趋势，识别干预与结果之间的因果关系，设定如下^[29]：

$$Y_{it} = \alpha + \beta(treat_i \times post_t) + \gamma X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

其中， Y_{it} 表示自然实验受试者个体 i 在时期 t 的结局变量， $treat_i$ 为暴露组虚拟变量， $post_t$ 为时间虚拟变量，交互项系数 β 反映自然变异的平均处理效应， X_{it} 为控制变量， μ_i 和 λ_t 分别为调查时期和地区的固定效应， ε_{it} 为扰动项。

2) 断点回归法：利用连续变量存在明确临界值（断点）以模拟随机化实验^[30]。当个体是否接受某种变异，取决于某连续变量是否超过特定临界值时，该值附近的个体具有高度可比性，可视为“准随机”地被分配至暴露组或对照组，其核心在于估计自然变异在断点附近的局部平均处理效应。在临界值附近，除变异因素外，个体的其他所有特征被视为连续变化，人为操纵分组变量的混杂因素得以控制。研究者通常采用局部线性回归或多项式拟合，对临界点两侧的数据分别拟合回归函数，并计算两者在临界点处的差值作为处理效应。

3) 工具变量法：可处理不可观测混杂因素导致的内生性问题^[31]。当研究中存在无法测量或未观测到的混杂因素时，工具变量法通过引入变量来剥离干预中的外生变异。此变量需满足以下条件：相关性，工具变量与内生解释变量高度相关；排他

性,工具变量只能通过影响内生解释变量来间接影响结局变量,不存在其他作用路径;随机性,工具变量与误差项不相关,即不受不可观测混杂因素的影响。

2.3.4 稳健性检验 稳健性检验可降低偶然效应及未观测混杂的干扰。安慰剂检验通过采用虚构处理时间或处理组别的方法,观察处理效应是否仍然显著,若上述虚构情境下均未发现显著效应,则证明分析结果可靠;对倾向得分匹配的过程进行敏感性分析,更换匹配策略可检验估计结果是否保持稳定,若不同匹配策略下所得结果与基线估计基本一致,则表明研究结论对匹配方案的选择不敏感,进一步确认了暴露组与对照组之间具有可比性;在双重差分法模型基础上也可引入更多协变量,例如患者合并用药情况、医生从业年限等可能影响结局的变量,若关键解释变量的系数估计值未发生实质性变化,且保持统计显著性,则可在较大程度上排除遗漏变量偏倚对研究结论的影响。

3 自然实验应用于中医药疗效评价的优势与挑战

自然实验能够有效保留中医药个体化诊疗特色,证据外部有效性高。与需要在高度受控环境下进行的随机对照试验不同,自然实验不主动干预患者的诊疗决策,从而完整保留中医药辨证论治的个体化诊疗模式精髓,其结果更能反映常规医疗条件下的实际疗效,具有较高的临床推广价值^[32]。

自然实验在伦理层面具有包容性,为部分研究提供可行路径。该方法中对照组来源于自然形成的分组,研究者无需为了研究目的而主动剥夺对照组患者接受可能有效治疗的机会,这为许多因伦理限制难以实施的研究课题开辟了可行的路径^[33]。

自然实验适用于难以通过 RCT 验证的中医药干预措施。针灸、针刀、推拿、太极等物理疗法以及民间特色疗法,由于难以实现盲法、操作标准化困难、治疗个性化程度高等原因,开展 RCT 存在方法学障碍。自然实验则可在不改变其原有实践方式的前提下,系统评估这类复杂干预在真实世界中的疗效与安全性,有效填补传统试验的证据空白。

自然实验存在未知的混杂因素,易影响疗效估计的准确性。由于分组非随机,接受不同干预的患者群体在基线特征,如病情、体质、社会经济状况乃至健康信念,可能存在系统性差异,这些因素会干扰对疗效的准确估计。为此,应在设计与分析阶段积极采用倾向得分匹配等统计方法,模拟随机化

以平衡组间差异,并结合敏感性分析来评估未测量混杂因素对结论的影响。

自然变异难以标准化。在真实场景下,中医药干预(如中药汤剂、针灸手法)的实施细节可能因医师、医院而异,存在剂量、配方、操作手法等方面的异质性^[34]。这为清晰界定“治疗是什么”带来了困难。针对这一问题,应在研究中引入标准化报告规范与过程评价,详尽记录治疗的实际内容、剂量与流程,并结合定性访谈等混合研究方法,深入理解医师决策与患者依从的深层逻辑,从而将不同的干预措施转化为可测量、可分析的科学数据。

自然实验因果论证易受测量偏倚影响。观察性研究的本质决定了其因果论证力度弱于 RCT,而中医药侧重证候、生存质量等主观指标的评估体系又易引入测量偏倚。为此,应优先选择分组机制近乎随机、因果链条清晰的自然实验情境,辅以相对客观的疗效指标,并加用经过严格验证的患者报告结局量表来提升结局测量的可信度。

4 小结

自然实验在 RCT 难以实施、伦理上不宜开展或外部真实性不足时,为中医药疗效评价提供了补充性的研究路径。中医药诊疗具有辨证论治、个体化调整和复杂干预等特点,其疗效往往受政策环境、医疗资源配置、医师诊疗习惯和患者真实依从性影响。自然实验能够利用这些真实世界中的外部变异,评价中医药干预在常规医疗环境下的实际效果。其证据强度取决于自然变异来源外生性、暴露组与对照组间可比性、结局测量可靠性,以及处理混杂偏倚的充分性。未来在中医药领域应用自然实验,应从研究问题选择、变异来源识别、数据质量控制、干预过程记录、因果推断方法选择和稳健性检验等环节建立更规范的实施流程,以形成更贴近临床实践且具有较强解释力的真实世界证据。

参考文献

- [1] Zipursky J, Juurlink D N. Studying drug safety in the real world [J]. JAMA Intern Med, 2018, 178 (11): 1533-1534.
- [2] Delpassand E S, Yazdi S M, Ghantaji S, et al. Effectiveness and safety of retreatment with ¹⁷⁷Lu-DOTATATE in patients with progressive neuroendocrine tumors: a retrospective real-world study in the United States [J]. J Nucl Med, 2024, 65(5): 746-752.
- [3] Craig P, Katikireddi S V, Leyland A, et al. Natural experi-

- ments: an overview of methods, approaches, and contributions to public health intervention research[J]. *Annu Rev Public Health*, 2017, 38:39-56. DOI: 10.1146/annurev-publhealth-031816-044327.
- [4] Craig P, Campbell M, Deidda M, et al. Using natural experiments to evaluate population health and health system interventions: new framework for producers and users of evidence[J]. *BMJ*, 2025, 388:e080505. DOI: 10.1136/bmj-2024-080505.
- [5] Lazursky A F. Natural experiment [J]. *Exp Psychol*, 2010, 3(2):87-98.
- [6] Buchanan R M, Smith A, Rowe I. The role of natural experiments in hepatology research: filling the gap between clinical trials and service evaluations[J]. *Hepatol Commun*, 2023, 7(5):e0121.
- [7] Vestbo J. Natural experiments and large databases in respiratory and cardiovascular disease[J]. *Eur Respir Rev*, 2016, 25(140):130-134.
- [8] Gregg E W, Duru O K, Shi L, et al. Filling the public health science gaps for diabetes with natural experiments [J]. *Med Care*, 2020, 58 Suppl6 Suppl 1:S1-S3. DOI: 10.1097/MLR.0000000000001330.
- [9] Skivington K, Matthews L, Simpson S A, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance [J]. *Int J Nurs Stud*, 2024, 154:104705. DOI: 10.1016/j.ijnurstu.2024.104705.
- [10] Sun X, Li L, Liu Y, et al. Assessing clinical effects of traditional Chinese medicine interventions: moving beyond randomized controlled trials [J]. *Front Pharmacol*, 2021, 12:713071. DOI: 10.3389/fphar.2021.713071.
- [11] Mcleod K S. Our sense of Snow: the myth of John Snow in medical geography [J]. *Soc Sci Med*, 2000, 50(7/8): 923-935.
- [12] Zhang H, Chen Z. Financial reform and haze pollution: a quasi-natural experiment of the financial reform pilot zones in China [J]. *J Environ Manage*, 2023, 330: 117196. DOI: 10.1016/j.jenvman.2022.117196.
- [13] Deidda M, Geue C, Kreif N. et al. A framework for conducting economic evaluations alongside natural experiments [J]. *Soc Sci Med*, 2019, 220: 353-361. DOI: 10.1016/j.socscimed.2018.11.032.
- [14] Khullar D, Jena A B. "Natural Experiments" in health care research [J]. *JAMA Health Forum*, 2021, 2(6): e210290.
- [15] Bor J, Fox M P, Rosen S, et al. Treatment eligibility and retention in clinical HIV care: a regression discontinuity study in South Africa [J]. *PLoS Med*, 2017, 14(11): e1002463.
- [16] Barnett M L, Olenski A R, Jena A B. Opioid-prescribing patterns of emergency physicians and risk of longterm use [J]. *N Engl J Med*, 2017, 376(7): 663-673.
- [17] Lee M, Chang J, Deng Q, et al. Effects of a coal to clean heating policy on acute myocardial infarction in Beijing: a difference-in-differences analysis [J]. *Lancet Planet Health*, 2024, 8(11): e924-e932.
- [18] Ogilvie D, Foley L, Nimegeer A, et al. Health impacts of the M74 urban motorway extension: a mixed-method natural experimental study [M]. Southampton (UK): NIHR Journals Library, 2017:1.
- [19] Brozek J L, Akl E A, Alonso-coello P, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines. Part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions [J]. *Allergy*, 2009, 64(5): 669-677.
- [20] Stadhouders A, Van Rossenberg L X, Kik C, et al. Natural experiments as a study method in spinal trauma surgery: a systematic review [J]. *Global Spine J*, 2024, 14(5):1640-1649.
- [21] 孙亚平, 张征宇, 孙广亚. 中国医疗卫生体制改革的健康及医疗服务效应: 基于综合医改试点的考察 [J]. *系统工程理论与实践*, 2024, 44(9): 3006-3022.
- [22] 韩扬阳, 郭蕊, 袁梦, 等. 线上线下一体化医疗模式下中医诊疗质量主观评价差异研究 [J]. *中国医院管理*, 2025, 45(4): 12-15.
- [23] 徐叶, 贾玉龙, 谭婧, 等. 临床研究中重复测量定量资料的样本量估计思路与方法: 单个时间点的组间差异比较 [J]. *中国循证医学杂志*, 2025, 25(9): 1092-1096.
- [24] Liang J, Hu Z, Zhan C, et al. Using propensity score matching to balance the baseline characteristics [J]. *J Thorac Oncol*, 2021, 16(6): e45-e46.
- [25] Elze M C, Gregson J, Baber U, et al. Comparison of propensity score methods and covariate adjustment: evaluation in 4 cardiovascular studies [J]. *J Am Coll Cardiol*, 2017, 69(3): 345-357.
- [26] Tchetgen Tchetgen E J, Park C, Richardson D B. Universal difference-in-differences for causal inference in epidemiology [J]. *Epidemiology*, 2024, 35(1): 16-22.
- [27] Feng S, Ganguli I, Lee Y, et al. Difference-in-differences for health policy and practice: a review of modern methods [J]. *Stat Med*, 2025, 44(23/24): e70247.
- [28] Zeldow B, Hatfield L A. Confounding and regression adjustment in difference-in-differences studies [J]. *Health Serv Res*, 2021, 56(5): 932-941.

- [29] Friedman A S. A Difference-in-differences analysis of youth smoking and a ban on sales of flavored tobacco products in San Francisco, California[J]. JAMA Pediatr, 2021, 175(8):863–865.
- [30] Pescarini J M, Craig P, Allik M, et al. Evaluating the impact of the Bolsa Familia conditional cash transfer program on premature cardiovascular and all-cause mortality using the 100 million Brazilian cohort: a natural experiment study protocol [J]. BMJ Open, 2020, 10(11): e039658.
- [31] Hamad R, Nguyen T T, Bhattacharya J, et al. Educational attainment and cardiovascular disease in the United States: a quasi-experimental instrumental variables analysis[J]. PLoS Med, 2019, 16(6):e1002834.
- [32] Bärnighausen T, Tugwell P, Røttingen J A, et al. Quasi-experimental study designs series-paper 4: uses and value [J]. J Clin Epidemiol, 2017, 89: 21–29. DOI: 10.1016/j.jclinepi. 2017. 03. 012.
- [33] Zhang J, Zhang Z M. The challenges of ethical review in clinical research of traditional Chinese Medicine [J]. Evid Based Complement Alternat Med, 2021: 6754985. DOI: 10.1155/2021/6754985.
- [34] Jacobson E, Conboy L, Tsering D, et al. Experimental studies of inter-rater agreement in traditional Chinese medicine: a systematic review[J]. J Altern Complement Med, 2019, 25(11):1085–1096.

Advantages, Challenges, and Implementation Pathways of Natural Experiments in Evaluating the Efficacy of Traditional Chinese Medicine

LIU Rihe¹, LIU Tingxiu¹, LIU Kai¹, SUN Yuan¹, ZHANG Xiaohan¹, LIU Jianping^{2,3}, LIU Zhaolan¹

1. Center for Evidence-Based Medicine, Beijing University of Chinese Medicine, Beijing, 100029; 2. Center for Evidence-Based Medicine, Hospital of Chengdu University of Traditional Chinese Medicine; 3. Sichuan Provincial Center for Evidence-Based Chinese Medicine

ABSTRACT A natural experiment is a causal inference study design grounded in real-world research. Its core concept is to use quasi-random variation in exposure generated by external factors to evaluate the effects of interventions or exposures on outcomes. It is particularly suitable for research settings in which randomized controlled trials (RCTs) are infeasible or ethically constrained, and it combines causal inference capacity with relatively strong external validity. This article provides an overview of the definition, key features, and current application status of natural experiments, and proposes the application scenarios for applying natural experiment strategies to the evaluation of therapeutic effects in traditional Chinese medicine (TCM), including policy implementation or policy change evaluation, assessment of variations in clinical practice within hospitals, and evaluation of differences in resource allocation across regions or institutions. The design pathway includes starting from clinical or policy-related TCM questions, systematically collecting relevant data, performing statistical modeling and effect estimation, and interpreting effect estimates in the context of clinical practice. It further analyzes the advantages and challenges of using natural experiments in TCM.

Keywords natural experiments; traditional Chinese medicine therapy; clinical efficacy evaluation; real-world study

(收稿日期: 2025-11-25; 录用日期: 2026-04-29; 修回日期: 2026-06-06)

[编辑: 焦 爽]