

《药物流行病学研究方法学指南（第2版）》系列 解读（14）：目标试验仿真研究的应用与案例



赵厚宇^{1, 2}, 程吟楚³, 孙 凤^{2, 4, 5, 6, 7}, 詹思延^{1, 2, 4, 8}

1. 北京大学第三医院临床流行病学研究中心（北京 100191）
2. 重大疾病流行病学教育部重点实验室（北京大学）（北京 100191）
3. 北京大学第三医院药学部（北京 100191）
4. 北京大学公共卫生学院流行病与卫生统计学系（北京 100191）
5. 北京大学第三医院眼科（北京 100191）
6. 新疆医科大学中医学院（乌鲁木齐 830017）
7. 新疆石河子大学公共卫生学院（新疆石河子 832000）
8. 北京大学人工智能研究院智慧公众健康研究中心（北京 100191）

【摘要】目标试验仿真（TTE）逐渐发展为药物流行病学研究因果推断的主流框架，通过仿真随机对照试验的设计要素，显著提升观察性研究的质量。本文基于《药物流行病学研究方法学指南（第2版）》，对TTE方法常用设计的基本原理进行解读，重点介绍活性药物对照新用药者设计、序贯试验、克隆-删失-加权三种常用的TTE设计的基本原理、适用场景、实施步骤，并分别结合典型研究案例，展示其在药物流行病学中的具体应用，最后对三种设计的基本特征进行比较，并就如何根据研究目的选择合适的设计类型提出了参考意见。

【关键词】目标试验仿真；药物流行病学；方法学；指南；活性药物对照新用药者设计；序贯试验；克隆-删失-加权

【中图分类号】 R181.3+5 **【文献标识码】** A

Guide on Methodological Standards in Pharmacoepidemiology (2nd edition) and their series interpretation (14): research designs and case studies of target trial emulation

ZHAO Houyu^{1,2}, CHENG Yinchu³, SUN Feng^{2,4,5,6,7}, ZHAN Siyan^{1,2,4,8}

1. Research Center of Clinical Epidemiology, Peking University Third Hospital, Beijing 100191, China
2. Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing 100191, China
3. Department of Pharmacy, Peking University Third Hospital, Beijing 100191, China
4. Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China
5. Department of Ophthalmology, Peking University Third Hospital, Beijing 100191, China
6. School of Traditional Chinese Medicine, Xinjiang Medical University, Urumqi 830017, China
7. School of Public Health, Shihezi University, Shihezi 832000, Xinjiang Uygur Autonomous Region, China

DOI: 10.12173/j.issn.1005-0698.202601123

基金项目：国家自然科学基金面上项目（82574225）；国家自然科学基金青年项目（C类）（82204157）；中国博士后科学基金（2025M780703）

通信作者：詹思延，博士，教授，博士研究生导师，Email: siyan-zhan@bjmu.edu.cn

孙凤，博士，研究员，博士研究生导师，Email: sunfeng@bjmu.edu.cn

8. Center for Intelligent Public Health, Institute for Artificial Intelligence, Peking University, Beijing 100191, China
Corresponding authors: ZHAN Siyan, Email: siyan-zhan@bjmu.edu.cn; SUN Feng, Email: sunfeng@bjmu.edu.cn

【Abstract】 Target trial emulation (TTE) has gradually developed into a mainstream framework for causal inference in pharmacoepidemiological research. By emulating the design elements of randomized controlled trials (RCTs), TTE significantly enhances the quality of observational studies. Based on the *Guide on Methodological Standards in Pharmacoepidemiology (2nd edition)*, this article interprets the fundamental principles of commonly used TTE designs. It focuses on introducing the basic principles, applicable scenarios, and implementation steps of the three commonly used TTE designs: the active comparator new-user (ACNU) design, sequential trials, and the clone-censor-weight (CCW) approach. Typical case studies are presented to illustrate the specific applications of each design in pharmacoepidemiology. Finally, this article compares the fundamental characteristics of these three designs and offers guidance on selecting an appropriate design type based on research objectives.

【Keywords】 Target trial emulation; Pharmacoepidemiology; Methodology; Guidelines; Active comparator new-user design; Sequential trial; Clone-censor-weight

在临床研究中，随机对照试验（randomized controlled trials, RCT）是评估干预措施疗效与安全性的金标准，但存在样本量有限、纳入排除标准严格、随访周期短、伦理限制等固有缺陷，难以完全反映真实临床场景中干预措施的实际效果^[1-2]。随着电子健康记录等真实世界数据（real-world data, RWD）的日益丰富，观察性研究成为 RCT 的重要补充，但传统观察性研究易受适应证混杂（confounding by indication）、非死亡时间偏倚（immortal-time bias）等影响，导致因果推断结果不可靠^[1-5]。对此，Hernán 等^[4, 6-7]提出的目标试验仿真（target trial emulation, TTE）框架，通过明确定义目标试验的七个核心要素（研究对象纳入排除、治疗策略、随机分配、随访起止、结局定义、因果对比、分析计划），并尽可能用观察数据仿真这些要素，使其成为连接观察性数据与因果推断的桥梁，有效降低了传统观察性研究的设计偏倚。

目标试验思想最早萌芽于 20 世纪中叶的统计学与计量经济学研究^[8]，此后经过半个多世纪的缓慢发展，直至 2016 年，Hernán 和 Robins 确立了 TTE 的基本框架和原则^[7]，此后 TTE 得到快速发展和应用。随着 TTE 方法日益受到学界重视，国内众多学者也开始关注和应用这一框架，并在中文期刊上进行了相关介绍和综述^[9-17]，在此基础上，国内专家于 2025 年进一步形成了关于 TTE 应用的中文版专家共识^[18]。鉴于 TTE 方法日趋成熟，应用日渐广泛，在药物流行病学研究中的重要性日益凸显，2025 年新发布的《药

物流行病学研究方法学指南（第 2 版）》（以下简称“指南第 2 版”）将 TTE 纳入方法学推荐体系。对于“target trial emulation”这一术语，目前国内文献的差异主要体现在对“emulation”的理解，存在“模拟”与“仿真”两种中文译法。指南第 2 版采用“目标试验仿真”的译法，以与统计学中基于概率模型生成随机数据的“模拟（simulation）”^[19]相区分。

目前，国内已有多篇文献对 TTE 的基本框架和实施要素做了介绍和解读^[9-18]，但对于实施 TTE 时不同设计的基本原理和设计类型选择尚缺少详细说明。然而，TTE 并非单一、刻板的设计模板，而是一个灵活的方法学框架。其核心挑战之一，是针对现实世界中复杂多样的研究问题与数据结构，如何利用观察性数据，在“零时点”（time zero）这一关键锚点上，实现对 RCT 中“随机化”时刻的仿真，因此“零时点”必须是研究对象资格评估、治疗策略分配和随访开始三者对齐的时间点^[4, 7]。根据研究问题的性质，尤其是在确定“治疗策略”和“零时点”时所面临的挑战不同，TTE 框架衍生出了三种主流设计策略，即活性药物对照新药者设计（active comparator new-user, ACNU）、序贯试验设计（sequential trial）和克隆-删失-加权设计（clone-censor-weight, CCW）^[3]。本文重点聚焦于上述三种常见的 TTE 设计类型，对指南第 2 版有关 TTE 内容进行补充，并通过文献解读对一些典型研究案例进行概述。此外，考虑到篇幅限制，本文主要阐释不同 TTE 设计的基本原理，而不深入展开具体的

统计分析细节, 因此内容侧重于流行病学设计思路, 而非统计方法本身。

1 TTE的3种常用设计介绍

1.1 ACNU设计

1.1.1 ACNU设计的适用场景

ACNU 设计通常要回答的问题是, 针对同一适应证, 患者使用一种药物 (或一种治疗方案) 相对于另一种药物 (或另一种治疗方案) 的效果或安全性的差异^[20-21]。在药物流行病学研究中, ACNU 是最重要的设计之一, 其核心在于通过选择与研究药物具有相同适应证的对照药物, 并将研究对象限定于新使用者, 从而在设计层面高度仿真头对头 RCT^[20]。具体而言, 活性对照 (active comparator) 部分需选择临床实践中可替代研究药物的其他药物作为对照, 并将研究对象限定为有治疗指征且无禁忌证的个体; 而新用药者 (new-user) 部分将研究对象进一步严格限定为经过充分洗脱 (如 12 个月前无研究药和对照药用药记录) 后首次开始治疗的新使用者, 并将随访起点统一设定为治疗开始时点, 以有效控制现用药者偏倚风险并确保协变量测量与暴露评估之间的时间顺序合理^[20-21]。该设计能有效控制由医生处方偏好和患者健康行为引起的适应证混杂与健康使用者偏倚 (healthy-user bias), 同时通过将随访起点锚定在相互比较的药物首次用药时间而有效控制非死亡时间偏倚^[20]。该设计可以很容易地扩展到 3 种及以上药物比较研究^[22-23]。但是相互比较的药物之间适应证应足够相似而使药物之间可以相互替代, 否则仍然可能引起适应证混杂、非死亡时间偏倚、滞后时间偏倚等偏倚风险^[20, 24]。

1.1.2 ACNU的实现步骤

本文总结了各类设计的关键环节, 旨在为研究者在构思 TTE 时提供思路参考, 而非实际开展 TTE 研究的具体操作步骤。读者可结合 TTE 的基本框架, 参考本文内容来规划具体的研究方案。

(1) 对照药物选择。在选择对照药物时, 关键之一在于控制由治疗指征及其他未观测的患者特征 (如衰弱状态等) 所导致的混杂偏倚。因此, 两种药物在临床上的可互换性越高, 其潜在的未测量混杂就越小。在实际研究中, 临床治疗指南通常是选择对照药物的首要依据。另一种可行的方法是, 系统比较初始使用研究药物的患者

与初始使用各候选对照药物患者在基线特征上的相似性, 并选择在年龄、性别、共病情况及合并用药等方面与研究药物人群最为接近的药物作为对照。若已测量的基线特征均无法强有力地预测患者会接受研究药物而非对照药物, 则可在一定程度上认为, 未测量的因素对治疗选择的影响也可能有限^[21]。

(2) 新用药者识别。新用药者指在一段明确的未用药期后启动治疗, 且处于治疗起始阶段的个体。在药物流行病学研究中仅纳入新用药者, 并以治疗启动为随访起点, 可有效描述和评估时变风险, 同时确保协变量评估的时间顺序合理性。理想情况下, 新用药者应为首次用药者。但基于电子数据库的药物研究通常缺乏患者的终身治疗数据, 因此, 实践中通常基于洗脱期 (washout period) 识别新用药者, 即首次处方 (治疗启动) 前的特定时间窗口内, 研究对象在此期间有就诊记录, 但未使用过研究药物和对照药物^[20-21]。实践中, 洗脱期通常设置为 6~12 个月, 但延长洗脱期的研究^[25]显示, 较短的洗脱期可能不足以排除现用药者 (prevalent user)。因此, 如果数据允许, 应考虑使用大于 12 个月的洗脱期, 并且针对洗脱期长度进行敏感性分析, 以检查其对研究结果的影响^[25]。

(3) 确定“零时点 (time zero)”, 即研究的索引日期 (index date)。在 ACNU 设计中, 通常将索引日期定义为新用药者的治疗起始时间, 即个体在满足洗脱期后首次使用药物的日期。其具体定义取决于数据来源, 如电子病历中的“首次处方开具日期”, 或药房数据库中的“首次药品调剂日期”。该时间点实现了研究对象纳排判定、治疗分配与随访起始三个关键事件的时间对齐, 是有效避免非死亡时间偏倚的基石。

(4) 混杂因素测量。为确保混杂、暴露与结局之间的正确时序关系, 混杂因素的测量必须在零时点之前。通常将测量窗口设定于零时点前的一段时间 (如前 1 年), 以避免纳入受暴露影响的变量。需要测量的混杂因素可涵盖人口学特征、临床合并症、既往用药史及实验室指标等。最终混杂因素集可结合因果图 (directed acyclic graph) 识别的必要变量与实际数据可获得性共同确定。

(5) 随访与删失处理。随访起始于零时点, 结束于最早发生的预先定义的终点事件, 包括目

标结局、全因死亡、失访（如医保退保、迁移等）和研究截止日期。根据目标估计量不同，还需决定是否将违背初始治疗方案（如停药、换药、加用其他治疗等）作为额外删失事件，这直接影响可估计的因果效应类型^[7, 26]。TTE 研究主要关注两种目标估计量，即意向性治疗（intention-to-treat, ITT）和符合方案（per-protocol, PP）效应^[7]。在 TTE 中，明确目标估计量是先于分析的关键步骤，它定义了研究所要估计的因果对比的具体内涵，并直接决定了随访和删失规则。ITT 效应估计的是启动某种治疗策略的效应，仿真 RCT 中随机化分组时的意向。在分析中，无论患者后续是否遵循其最初的治疗策略，均按初始分配的策略进行随访，不因用药改变而删失。PP 效应估计的是持续遵循初始治疗策略的效应，因此当患者偏离初始分配的策略（如停药、换用对照药）时，其后续随访数据被人工删失。为校正这种非随机删失导致的选择偏倚，通常需基于逆删失概率加权（inverse probability of censoring weighting, IPCW）等方法控制时依性混杂^[7]。

图 1 展示了 ACNU 的设计原理。图中涉及两种药物及 4 个研究对象。根据新用药者定义，2 号患者被剔除，因为其在服用 B 类药物前有 A 类药物的用药记录。对于纳入的 1、3、4 号患者，零时点（time zero）即为两种药物的首次记录日期。图中还展示了不同的终点事件，如果研究估计 PP 效应，则需要考虑随访中用药的变化，例如 3 号研究对象，在由 A 药切换为 B 药时删失。

1.1.3 研究案例：不同降糖药物的肾脏安全性

以 Xie 等^[23]对不同降糖药物肾脏安全性的比较研究来简要说明 ACNU 设计。该研究资料来源于美国退伍军人数据库（VA Health Care

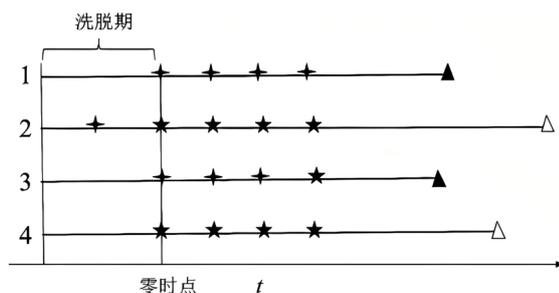


图1 ACNU设计的实施示意图

Figure 1. Illustration of ACNU design

注：四角星和五角星分别代表药物A和B；1、2、3、4代表4个研究对象；三角形为终点事件。

System），设计仿真了 4 种二线降糖药物的头对头目标试验。研究通过设置 1 年洗脱期，纳入 216 558 例 2 型糖尿病新用药者，比较 4 种降糖药物，即钠-葡萄糖协同转运蛋白-2（sodium-glucose cotransporter-2, SGLT-2）抑制剂、胰高血糖素样肽-1（glucagon-like peptide-1, GLP-1）受体激动剂、二肽基肽酶-4（dipeptidyl peptidase-4, DPP-4）抑制剂和磺脲类药物的复合肾脏结局（估算肾小球滤过率下降 > 50%、终末期肾病或全因死亡）风险。研究将患者首次处方日期设为索引日期并从此时间开始随访，通过从索引日期前 1 年内的电子健康记录中提取的高维变量构建广义倾向评分（propensity score, PS），采用逆概率加权仿真随机化分组，并辅以阴性/阳性对照结局分析检验未观测混杂，分别估计了 ITT 效应和 PP 效应，并在 PP 分析中采用 IPCW 控制人工删失引起的选择偏倚。结果显示，与磺脲类药物相比，SGLT-2 抑制剂、GLP-1 受体激动剂和 DPP-4 抑制剂均显著降低不良肾脏结局风险，风险比（hazard ratio, HR）及其 95% 置信区间（confidence interval, CI）分别为 0.68 [95%CI (0.63, 0.74)]、0.72 [95%CI (0.67, 0.77)] 和 0.90 [95%CI (0.86, 0.95)]，且 SGLT-2 抑制剂与 GLP-1 受体激动剂效果相当，均优于 DPP-4 抑制剂，两者相对于 DPP-4 抑制剂的 HR 分别为 0.76 [95%CI (0.70, 0.82)] 和 0.79 [95%CI (0.74, 0.85)]。该研究成功仿真了多组“头对头”比较的试验，关键在于选择了具有相同核心适应证（二线降糖）的对照药物，并将零时点锚定于首次处方日期。与多数 ACNU 设计的 TTE 研究不同，该研究目标试验比较组包含了 4 种药物，因此其在估计用药选择的倾向时使用了基于多分类 Logistic 回归的广义 PS。

1.2 序贯试验设计

1.2.1 序贯试验设计的适用场景

除了 ACNU 回答的针对同一适应证的不同药物或干预措施的相对效果或安全性问题，临床或公共卫生实践中，研究者有时可能更关心使用某种药物（或接受某种干预）相对于不使用该药物（或不接受干预）的效果和安全性^[20]。例如，他汀作为心血管疾病一级预防措施相比于不使用他汀的有效性和安全性^[27]，接种流感疫苗相比于不接种流感疫苗的有效性或安全性等问题^[28]。与 ACNU 的研究场景中对照组纳排评估和随访起始错

定于首次用药时间不同，此时非用药组没有明确的纳排、随访和治疗分配时间，在发生终点事件前其可能一直满足纳排条件。对此，很多传统的药物流行病学研究错误地忽视了上述三个关键事件的对齐，其后果是造成非死亡时间偏倚^[4, 29]。例如一种常用的传统设计是两组均从疾病（研究药物的适应证）诊断时间进行纳排和开始随访，根据随访中是否用药划分暴露组，此时用药组从疾病诊断到用药之间的时间即为非死亡时间（immortal time），研究对象在此期间不能发生终点事件^[4, 20, 29]。

序贯试验设计主要为应对上述问题而产生^[4]。与 ACNU 的静态框架不同，序贯试验设计将纵向观察数据离散化为一系列连续的时间单元（例如月、周、日），每个单元均被视作一个独立的试验期，从而形成随时间延伸、等间隔的序列“试验”，每个“试验”中独立进行纳排评估、干预分配与随访且三者锚定于“试验”所对应区间的起点（图2）^[4, 26, 30]，在每个试验期初重新评估纳排条件，将符合条件的个体虚拟分配至“开始启动”或“不治疗”策略，随访从每期分配日重新计算。在这种范式下，研究对象不再局限于单一时间点的状态，而是可以在多个试验期中重复符合条件并贡

献“人-试验”（person-trial）观察单位^[4, 26]。这种特性使得序贯试验设计尤其适用于纵向数据中治疗启动时间分散、任意时间点开始治疗的个体数量有限或结局事件较罕见的研究场景^[26, 31]。

1.2.2 序贯试验实现步骤

实践中可以根据以下过程设计基于序贯试验的 TTE（图2）。

（1）定义试验序列。将研究时期 T 划分为间隔为 t 的等间隔区间。间隔 t 的选取需要综合考虑疾病病程、研究时期 T 的长度、数据测量的频率与密集度、分析的便利性，以及运算能力等多个因素。在实践中多以月、周、日等作为间隔，例如研究期间跨度为 2011 年 1 月—2020 年 12 月，研究问题为长病程的慢性疾病用药，则可以月作为间隔，如此会产生 10 × 12 共 120 个“试验”。实践中还可能需要考虑设置洗脱期和观察期，则需要对观察期长度和试验的数量做相应调整，例如设置 1 年洗脱期，随访至少 1 年，则上述案例仿真的试验入组时期需调整为 2012 年 1 月—2019 年 12 月，共 96 个“试验”。Penning 等^[32]研究了存在时依性混杂时，混杂和暴露测量间隔对效应估计偏倚的影响，发现结果偏倚会随间隔延长

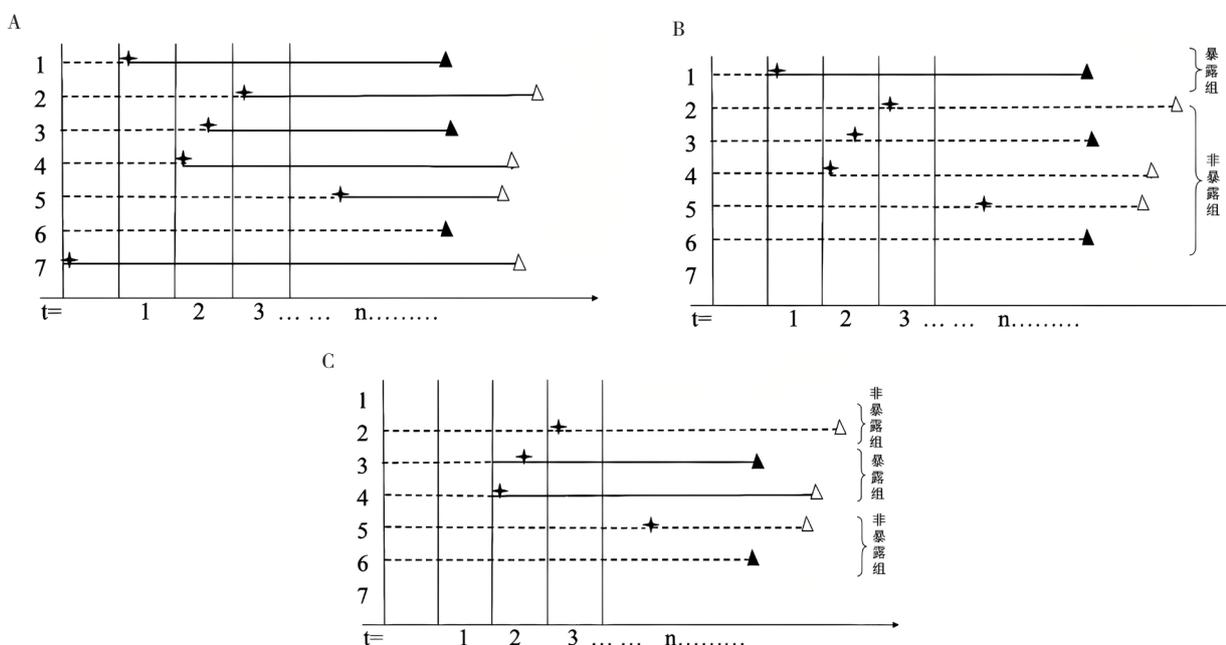


图2 序贯试验设计示意图

Figure 2. Illustration of sequential trial design

注：A. 7个研究对象观察性数据示例，图中四角星为药物处方，虚线表示非暴露时间，实线为暴露时间，三角形为终点事件，图中示意了3个等间隔的观察期，每一个对应一个“试验”。B. 基于图2-A的第一个试验，t=1为该“试验”的索引日期，7号患者在索引日期前有用药记录，因此从试验1中剔除，1~6号患者纳入试验1；t=1时，1号患者有药物使用记录，因此归入暴露组，2~6号患者无用药记录，因此归入非暴露组，随访从t=1开始至最早的终点事件。C. 基于图2-A的第2个试验，t=2为该“试验”的索引日期，1号和7号患者在索引日期前有用药记录，因此从试验2中剔除，2~6号患者纳入试验2；t=2时，3号和4号患者有药物使用记录，因此归入暴露组，2、5、6号患者无用药记录，因此归入非暴露组，随访从t=2开始至最早的终点事件。

而增大,因此在研究中应避免长间隔,且需根据暴露与协变量的变化速率合理设计测量频率,以防止因测量不足或测量时机不当而引入偏倚。

构建“人-试验”数据是序贯试验设计的方法学核心,每个符合条件的个体在其符合资格的时间段内,可重复进入多个连续的“试验”。这不是对数据的简单重复利用,而是一种解决时间相关性偏倚的机制,它确保了在每个试验的基线,暴露组与对照组的确定是同步的,从而避免了因以未来暴露状态回溯性定义基线而引入的非死亡时间偏倚。“人-试验”结构将复杂的纵向数据,转化为一系列截面化的、标准化的因果问题,是观察性研究模仿动态随机化过程的精巧体现。

(2) 确定每个试验的零时点。在目标试验中,随机化发生于每个试验的开始,因此将观察期离散化后,每个间隔的起点即为该间隔所对应的试验的零时点。需基于该时点确定研究对象纳排、干预分配和开始随访。

(3) 患者分配与随访。对于每一个“试验”,在该“试验”的零时点间隔内确定研究对象的暴露状态,该时间区间内使用了药物则为用药组,无药物记录则为对照组。

(4) 混杂因素测量。混杂测量原则同 ACNU 设计,但对每一个“试验”,索引日期不同,混杂测量的窗口随着试验序号而滑动。

(5) 随访与删失处理。随访原则同 ACNU 设计,同样需要基于每个“试验”独立的索引日期确定研究对象的随访终点。与 ACNU 设计的另一个明显不同在于,序贯试验设计下,同一研究对象可能重复多次入组,因此会造成观测数据量成倍增长,在提升研究功效的同时也增加计算量。由于重复纳入研究对象还会破坏观测间的独立性,导致传统的统计方法严重低估标准误,使 CI 变窄、*P* 值偏小,从而夸大统计显著性,增加第一类错误(假阳性)的风险。因此在实践中多采用稳健标准误或自助法估计效应值的 CI^[26,31]。

1.2.3 研究案例:他汀类药物与冠心病一级预防

Danaei 等^[26]的研究利用英国 THIN (The Health Improvement Network) 电子健康数据库评估 55~84 岁人群中使用时汀类药物相对于不使用该类药物治疗冠心病 (coronary heart disease, CHD)

的一级预防效果,研究观察期为 2000 年 1 月至 2006 年 12 月。该研究采用 TTE 框架,构建了 83 个序贯非随机“试验”(每月启动 1 个,间隔 1 个月)。研究纳入无 CHD 病史且至少有 2 年洗脱期(未使用他汀类药物)的个体。在每个试验的基线时,从符合条件的研究对象中判断其纳入试验的当月是否使用他汀类药物并据此进行干预分组,所有研究对象随访至 CHD、死亡、失访或研究结束。混杂测量采用动态策略,基线协变量取试验开始时最近记录值。为控制混杂和选择偏倚,研究运用逆概率加权估计 3 种目标估计量——ITT 效应(启动者 vs. 非启动者)、PP 效应(持续使用者 vs. 永不使用者)和实际接受治疗(as-treated, AT)效应(2 年累计暴露 vs. 无暴露)的风险比(HR)。最终纳入 74 806 例个体的 844 800 个“人-试验”观测,结果显示,调整混杂后 ITT 效应的 HR 为 0.89 [95%CI (0.73, 1.09)], IPCW 校正的 PP 效应 HR 为 0.84 [95%CI (0.54, 1.30)], 2 年累积暴露的 AT 效应的 HR 为 0.79 [95%CI (0.41, 1.41)], 提示在充分校正混杂后他汀类药物可能对 CHD 具有保护性效应,但研究统计效能不足,需要进一步研究对其结论加以验证。

与第一个研究案例的显著不同在于, Danaei 等^[26]研究中对照组是“不使用他汀”的人群,因此不适用于 ACNU 设计,通过构建以月为间隔的试验,为“使用他汀”相对于“不使用他汀”这一缺乏天然起点的对比创造了动态对齐的零时点。这篇文献是 TTE 发展早期应用序贯设计的代表性研究,其核心并非在于验证他汀类药物对 CHD 的保护性作用,而是通过案例展示 TTE 框架如何克服传统观察性研究的局限,它不仅系统阐述了序贯试验的构建流程、混杂控制策略(仿真随机化),以及效应估计方法等关键环节,还通过实际案例分析,对比和阐释了 3 种目标估计量 (ITT 效应、PP 效应和 AT 效应) 的估计方法及其对应的因果内涵。该研究为理解和应用序贯设计提供了清晰的方法学范本,是学习这一方法的重要参考文献。

1.3 CCW设计

1.3.1 CCW的适用场景

CCW 设计主要用于一些复杂干预策略下控制非死亡时间偏倚和混杂偏倚^[33-34],常用于基于宽限期 (grace period) 的干预比较(如癌症诊断后 6

个月内启动与不启动他汀类药物治疗)^[35]、治疗时长比较(如治疗 6 个月相对于治疗 12 个月的有效性或安全性)^[36]、中断治疗与维持治疗的比较^[37]，或基于生物标志物水平启动治疗的动态治疗策略(dynamic treatment strategy, 例如当 CD4 细胞计数低于 $200 \cdot \mu\text{L}^{-1}$ 、 $300 \cdot \mu\text{L}^{-1}$ ，或 $400 \cdot \mu\text{L}^{-1}$ 等不同水平时开始抗逆转录病毒治疗的研究)^[38]，以及药物联用与单用比较等问题(如在吉西他滨治疗开始后 12 周内加用厄洛替尼与单用吉西他滨治疗恶性肿瘤的效果比较的研究)^[39]。CCW 方法代表了 TTE 范式中最具技术创新的方向，通过“克隆-删失-加权”三步流程实现复杂治疗策略的因果解构。该范式的独特之处在于不依赖协变量调整即可实现组间基线特征均衡，即通过为每个受试者创建多个克隆，每个克隆被分配至不同的治疗策略，从而有效控制已观测与未观测的基线混杂^[34]。随后，当克隆体在随访过程中偏离其被分配的治疗策略时对其进行人工删失(图 3)，并通过 IPCW 控制人工删失导致的选择偏倚，恢复人群代表性^[34, 38]。CCW 方法的主要优点在于其能够清晰地嵌入 TTE 框架，通过克隆与删失的机制有效防止非死亡时间偏倚，同时允许估计与现实临床决策相符的、包含宽限期的治疗策略的因果

效应^[4]。该设计的有效性高度依赖于后续加权模型能否正确校正由人工删失(克隆偏离其被分配的治疗策略)所引入的时依性选择偏倚，并对暴露时间、结局时间的测量准确性敏感。此外，由于克隆导致数据膨胀，需采用稳健标准误或非参数自助法估计因果效应值的置信区间以考虑个体多次贡献观测数据的相关性^[37, 40]。

1.3.2 CCW的实现步骤

在 CCW 的三个关键步骤基础上，实践中可以根据以下过程开展 CCW 研究(图 3)。

(1) 定义干预策略和宽限期。干预策略直接来源于研究问题，即描述我们想要比较哪些治疗策略的相对有效性或安全性。如前所述，CCW 适合研究复杂干预策略，包括基于宽限期的干预定义、治疗中断、治疗时长、合并治疗、动态治疗策略等问题的研究^[34]。宽限期是指在随访开始后的一段时间窗口，在此期间个体如果按其被分配的方案接受治疗仍视为符合治疗方案^[4, 34]。这与 RCT 中受试者在随机分组后不能马上接受干预而需等待一定时间的情况相似^[4]。实践中宽限期的长度需要结合疾病特征、治疗指南、数据特性、处方模式和研究目标等综合考虑。并不存在一个适用于所有研究的、统一的经验性宽限期长度。

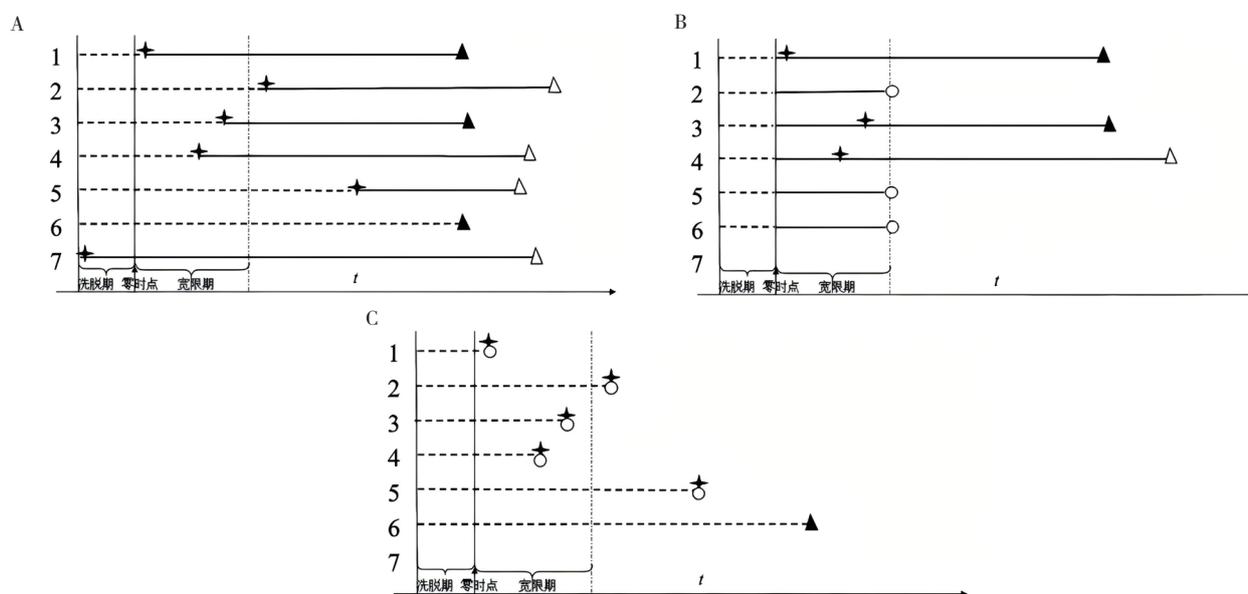


图3 CCW设计示意图

Figure 3. Illustration of CCW design

注：假定的研究场景是疾病诊断后宽限期(如6个月)内启动治疗与不治疗对结局的因果效应，图中四角星为药物处方，虚线表示非暴露时间，实线为暴露时间，三角形为终点事件，空心圆表示人工删失。A. 7个研究对象观察性数据示例。B. 图3-A中所有个体的克隆，全部研究对象均被分配到启动治疗组，在宽限期内启动治疗视为遵循治疗方案，图中1、3、4号研究对象在宽限期内启动了治疗，符合被分配的治疗方案，因此不进行人工删失；2、5、6号研究对象在宽限期结束时仍未启动治疗，违背被分配的治疗方案，因此在宽限期结束时人工删失，但是在宽限期内其仍然有可能在任何时间接受治疗，因此不在宽限期内人工删失。C. 图3-A中所有个体的克隆，全部研究对象均被分配到不治疗组，图中1~5号研究对象在随访中均违背了被分配的治疗方案，因此在启动治疗时人工删失，6号研究对象在观察期始终未接受治疗，符合被分配治疗方案，不进行人工删失，7号研究对象在洗脱期内有用药记录，从所有克隆中剔除。

(2) 确定零时点。在 CCW 设计中, 零时点的选择取决于试图仿真的试验特征, 不同的研究问题和干预策略比较可能产生不同的零时点设定^[41]。例如, 在针对某种疾病的用药策略比较时, 如果比较组之一是“不治疗”, 有些研究将零时点设定为疾病诊断日期^[35, 42]。当研究治疗持续时间^[43]或者不同联合用药策略比较^[39]等问题时, 可以将零时点设定为用药开始时间。因此, 在确定零时点时, 需结合具体研究问题与干预策略特征, 确保研究对象纳排、干预分配和随访开始在零时点对齐。

(3) 克隆 (clone)。为每个符合条件的个体创建与治疗方案数相等的克隆, 每个克隆被分配到一个不同的治疗策略 (例如“在 90 d 内开始治疗”和“从不治疗”), 这确保了基线时各策略组在已测量和未测量协变量上完全平衡, 因为所有克隆源于同一个体。

(4) 删失 (censor)。随访开始后每个克隆在其实际治疗偏离所分配策略时被人工删失。例如, 在比较治疗与不治疗时, 分配到治疗组的克隆若在宽限期结束时仍未开始治疗, 则在该时间点被删失; 分配到对照组的克隆若在随访期间开始治疗, 则在开始治疗时被删失。这一步骤确保了分析仅基于始终遵循初始治疗策略的个体-时间 (person-time)。

(5) 加权 (weight)。因为删失行为常与预后因素相关, 人工删失可能引入选择偏倚, 因此需要进行加权调整。通常使用 IPCW 方法, 其分母为个体在给定基线及时依协变量下保持未删失的条件概率, 分子可仅基于基线协变量计算以稳定权重^[34]。

1.3.3 研究案例: 早期手术治疗非小细胞肺癌的效果

Maringe 等^[44]采用 CCW 设计, 评估了老年非小细胞肺癌患者在诊断后早期接受手术的因素效应。该研究基于英格兰国家癌症登记数据 (National Cancer Registry in England), 纳入了 2012 年确诊、年龄 70~90 岁的非小细胞肺癌患者。研究设定观察期为诊断后 1 年, 以确诊日期作为随访起点, 并将干预策略设定为“在诊断后 6 个月内接受手术”, 即将宽限期设定为 6 个月以贴合临床实际等待时间。研究通过克隆患者构建了两个基线特征完全相同的队列; 随后, 依据患者实际治疗与所分配的治疗策略的一致性实施人工删失, 并采用 IPCW 控制由此引起的选择偏倚。

目标估计量为 PP 效应, 加权分析结果显示, 手术可使患者 1 年生存率提升 11.4% [95%CI (7.9%, 15.3%)], 通过估算限制性平均生存时间 (restricted mean survival time, RMST) 比较两组生存时间差异, 则相当于手术在诊断后第一年内平均延长 13 d [95%CI (8 d, 20 d)] 的生存时间。该研究中对组同样包含“不治疗” (即非手术组), 不适合采用 ACNU 设计, 同时宽限期的定义使 CCW 成为实施 TTE 时自然的设计选择。该研究提供了从设计到分析的完整 CCW 实现流程, 并附有 R 软件与 Stata 代码, 是学习和应用 CCW 方法的重要参考。

2 干预定义与策略选择

TTE 为适应不同研究问题及与之对应的目标试验而衍生出不同的设计类型, 共同致力于解决诸如非死亡时间偏倚、适应证混杂等一系列关键的方法学挑战^[3]。ACNU 设计高度依赖活性对照选择的恰当性, 若对照药物与研究药物适应证不完全重叠、临床可互换性低, 则可能无法有效控制适应证混杂。此外, 该设计不适用于无合适活性对照的药物研究 (如全新作用机制的药物), 也无法用于对比组中有“不治疗组”的研究。序贯试验是解决非治疗组在多个时间点满足纳入排除标准因而难以确定零时点问题的自然选择, 但该设计会使样本量大幅膨胀, 增加计算负担。由于同一个体贡献多个相关观测, 必须使用稳健方差估计或自助法以获取有效的 CI。此外, 间隔划分的长度 (如月、周) 可能引入时间离散化偏倚, 需通过敏感性分析评估。CCW 设计通过克隆有效控制基线混杂, 其估计结果对 IPCW 模型的正确设定敏感, 模型误设会导致残余混杂或选择偏倚。同时, 研究者在锚定时间点、权重类型等关键环节面临多样化的方法学选择, 目前缺乏统一标准, 可能影响结果的解释与推广。因此, 在应用 TTE 时, 应考虑不同设计类型的特征和适用场景, 同时了解其局限性以保证效应估计的有效性 (表 1), 根据具体的研究问题和目标试验框架选择恰当的 TTE 设计类型。良好的干预策略定义是匹配目标试验与恰当设计类型的桥梁, 否则将导致潜在结果模糊、一致性 (consistency) 条件不满足, 最终使因果效应估计失去意义^[6, 8]。例如, 在绝经后乳腺癌患者中, 研究“在确诊后 6 个月内

表1 3种常用TTE设计的特征比较

Table 1. Comparisons of three widely adopted TTE designs

| | ACNU设计 | 序贯试验设计 | CCW设计 |
|----------|---|--|--|
| 适用场景 | 仿真头对头RCT, 比较两种或多种具有相同适应证、相互可替代的药物的相对有效性或安全性 | 适用于暴露组人数少、需重复利用数据或干预开始时间不明确(如“不治疗”组无明确随访起点)的研究 | 设计灵活, 适用于基于宽限期的治疗策略比较以及其他复杂药物治疗定义, 例如中断治疗、动态治疗等问题 |
| 研究对象 | 新用药者, 无近期用药史, 常需设置洗脱期 | 新用药者, 无近期用药史, 常需设置洗脱期 | 在一些特殊问题研究中可以为现用药者, 如比较中断治疗与持续治疗 |
| 三要素对齐方式 | 将零时点锚定在相比较药物的首次用药时间 | 将每个“试验”的零时点锚定在其对应的时间间隔的起点 | 通过克隆将比较组的零时点锚定在相同时间(如开始用药时间、诊断时间等) |
| 目标估计量 | 可估计ITT和PP效应 | 可估计ITT和PP效应 | 通常用于估计PP效应 |
| 基线混杂控制 | 通过PS加权、匹配等方法控制已观测混杂, 仿真随机化分组 | 通过PS加权、匹配等方法控制已观测混杂, 仿真随机化分组 | 通过克隆使比较组间已观测和未观测混杂一致 |
| 时依性混杂控制 | 估计PP效应需控制时依性混杂, 常用IPCW方法 | 估计PP效应需控制时依性混杂, 常用IPCW方法 | 需控制时依性混杂, 常用IPCW方法 |
| 统计效率与功效 | 取决于相比较药物的临床使用量 | 通过重复利用数据提高效率, 尤其适用于暴露稀少或结局罕见研究。通常“不治疗”的对照数量远多于治疗组, 分析的数据量随“试验”数成倍增长, 增加计算负担 | 通过克隆扩大样本, 人工删失降低有效随访时间 |
| 局限性与误用风险 | 活性药物对照选择对于结果解释很关键, 不具有充分可替代性的药物之间的比较可能引入非死亡时间偏倚、滞后时间偏倚和适应证混杂。新用药者的识别依赖足够长的洗脱期, 设置不当引入现用药者偏倚 | 同一患者重复入组破坏独立性, 需校正标准误; 人为将时间离散化可能引入偏倚, 任意离散化不仅可能改变所估计的目标参数的定义, 还可能因无法充分控制时依性混杂或造成数据稀疏性而导致估计偏倚。实践中, 通常将所有“试验”的数据合并分析, 计算量较大 | 实施复杂且依赖于高质量纵向数据。方法涉及克隆、时依性权重建模等多重步骤, 计算负担大, 且权重估计可能不稳定。同时, 研究者在锚定时间点、权重类型等关键环节面临多样化的方法学选择, 目前缺乏统一标准, 可能影响结果的解释与推广。未来需进一步明确不同实施策略的优劣以形成最佳实践。克隆导致数据膨胀并破坏观测间独立性, 导致标准误被低估, 需采用稳健标准误或非参数自助法等方法估计CI |

启用芳香化酶抑制剂治疗”与“不启用该治疗”对生存的影响, 更有可能被认为是一个良好定义的策略, 其结果可直接指导临床“是否开始治疗”的决策。反之, 若比较“正在使用”与“从未使用”, 则可能是一个定义不佳的干预, 因为其结果既不能提示是否应该启用该药治疗, 也无法告诉患者是否应该停用该药治疗。在药物流行病学研究中, 干预定义模糊是导致研究结果异质性大、无法与 RCT 结果交叉验证的重要原因^[6]。定义良好的干预 (well-defined intervention) 本质是可清晰描述、可操作化实施、能与观察性数据准确映射, 并支持因果对比的干预方案^[6, 8, 45]。其核心内涵源于 TTE 的一致性条件 (consistency condition) —— 即对于接受干预水平 $A=a$ 的个体, 其潜在结果 Y^a 需与观察结果 Y 一致^[8]。这一内涵包含三个关键特征: 第一, 可明确锚定目标试验方案。干预定义需对应目标试验中具体的干预策略, 包括如何

实施干预、干预对象、干预持续时间等核心要素, 使观察性数据的分析能够复刻 RCT 的干预逻辑。第二, 具备可操作性与可重复性。干预的每一个关键维度都应是可执行、可验证的, 避免抽象描述。对于药物干预而言, 需明确起始时间、剂量、给药途径等可测量的指标, 确保不同研究者基于同一定义能从数据库中筛选出相同的研究人群。第三, 支持无歧义的因果对比。干预定义需明确“与什么相比” (即对照组), 且干预与对照定义的差异仅在于干预本身, 而非其他混杂因素。最后, 干预定义的每个要素都必须在所使用的观察性数据库中准确找到对应的变量或变量组合^[6]。若数据库无法提供定义所必需的详细信息, 应审慎修订目标试验方案以匹配数据可及性。

3 小结与展望

TTE 框架的 3 种设计类型具有各自的优势和

适用的研究场景, ACNU 是经典设计, 当研究问题为相同适应证的药物比较时, 其设计透明、易于实施, 已成为药物流行病学研究的标准方法^[20-21]。序贯试验特别适用于暴露频率低、结局事件罕见或启动时间高度可变的场景, 通过构建试验序列, 充分利用观察数据的纵向特征, 是回答治疗相对于不治疗问题的自然选择^[31]。CCW 通过克隆机制在观察性数据中构建反事实情景和植入随机化元素, 干预策略灵活, 是动态治疗等复杂治疗策略的理想研究方法^[34]。

然而, 3 种设计类型并非矛盾互斥, 当同一研究问题适用多种设计时, 不同设计下结果的一致性能证明结果的稳健性和可靠性。甚至在一些研究问题中, 可以将不同设计类型结合使用, 发挥多种设计类型的优势。例如, García-Albéniz 等^[41]在研究乳腺 X 线筛查是否降低乳腺癌死亡率时, 结合了序贯试验和 CCW 设计, 在分析中为 70~84 岁女性按年龄创建了 15 个“试验”, 在每个“试验”中应用 CCW。Kirkegård 等^[46]在比较手术相比于化疗对胰腺癌患者生存率的因果效应时, 结合了 ACNU 和 CCW 设计, 定义了手术和化疗两个新治疗组, 在克隆后将宽限期(胰腺癌诊断后 16 周)内未接受手术和化疗的患者进行人工删失。

TTE 的设计仍在不断发展, 一些传统的流行病学设计与 TTE 结合, 能够克服关键设计因素处理不当造成的结果偏倚。例如, Dickerman 等^[47]率先将 TTE 框架引入巢式病例对照设计。其核心是明确界定每个观察对象的零时点, 并在此时间点评估治疗策略和混杂因素, 而非在病例发病或对照抽样的时点进行评估。该方法有效避免了传统病例对照分析中常见的现用者偏倚, 并防止因调整受既往治疗影响的变量而引入的偏倚。随后 Rasouli 等^[48]进一步细化了该方法, 允许个体在符合条件的不同时间点重复进入多轮“试验”, 提高了统计效率。研究表明, 遵循该框架的病例对照分析所得效应估计与同源队列分析结果一致, 且更接近随机试验的证据, 而偏离该框架的传统方法则可能得出严重有偏的估计。该框架的优势在于, 病例对照设计允许在资源有限条件下集中验证关键结局或暴露变量(如通过病历回顾确认心肌梗死诊断), 在保证数据质量的同时, 通过结构化的 TTE 最大限度地减少因时依性混杂、不恰当协变量调整 and 选择偏倚所导致

的结果偏倚。同时, 在计算上, 分析仅需处理一个比全队列规模小得多的数据集, 显著降低了数据管理与模型拟合的负担, 而通过发病密度抽样和恰当的对照比例, 所得效应估计仍是全队列风险比的有效近似, 实现了科学严谨性与研究可行性之间的平衡^[48]。此外, 随着机器学习方法的发展, TTE 的不同设计与机器学习结合, 能够克服传统方法的一些局限性。例如, 因果机器学习方法(如因果森林、元学习器、双重机器学习等)通过灵活建模 PS 与结局, 能够有效估计目标效应, 从而识别药物治疗效应在不同亚组或个体中的差异^[49]。

总之, TTE 为观察性研究提供了严谨的因果推断框架。理解不同设计类型的特征、优势及适用场景, 有助于研究者根据具体的研究目标、数据条件和科学问题, 选择并实施最恰当的研究设计与分析方法, 从而产生更可靠、更接近随机试验证据的真实世界研究结果。

利益冲突声明: 作者声明本研究不存在任何经济或非经济利益冲突。

参考文献

- Lambourg E. Improving the quality of pharmacoepidemiological studies using the target trial emulation framework[J]. *Nat Rev Nephrol*, 2024, 20(12): 769. DOI: [10.1038/s41581-024-00884-4](https://doi.org/10.1038/s41581-024-00884-4).
- Chiodo S, Tailor L, Platt RW, et al. Emulating a target trial in perinatal pharmacoepidemiology: challenges and methodological approaches[J]. *Curr Epidemiol Rep*, 2023, 10(4): 275-285. DOI: [10.1007/s40471-023-00339-7](https://doi.org/10.1007/s40471-023-00339-7).
- Fu EL. Target trial emulation to improve causal inference from observational data: what, why, and how?[J]. *J Am Soc Nephrol*, 2023, 34(8): 1305-1314. DOI: [10.1681/ASN.000000000000152](https://doi.org/10.1681/ASN.000000000000152).
- Hernán MA, Sauer BC, Hernández-Díaz S, et al. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses[J]. *J Clin Epidemiol*, 2016, 79: 70-75. DOI: [10.1016/j.jclinepi.2016.04.014](https://doi.org/10.1016/j.jclinepi.2016.04.014).
- Hubbard RA, Gatsonis CA, Hogan JW, et al. "Target trial emulation" for observational studies—potential and pitfalls[J]. *N Engl J Med*, 2024, 391(21): 1975-1977. DOI: [10.1056/NEJMp2407586](https://doi.org/10.1056/NEJMp2407586).
- Hernán MA, Dahabreh IJ, Dickerman BA, et al. The target trial framework for causal inference from observational data: why and when is it helpful?[J]. *Ann Intern Med*, 2025, 178(3): 402-407. DOI: [10.7326/ANNALS-24-01871](https://doi.org/10.7326/ANNALS-24-01871).
- Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available[J]. *Am J Epidemiol*, 2016, 183(8): 758-764. DOI: [10.1093/aje/kwv254](https://doi.org/10.1093/aje/kwv254).
- Hernán MA. Does water kill? a call for less casual causal

- inferences[J]. *Ann Epidemiol*, 2016, 26(10): 674–680. DOI: [10.1016/j.annepidem.2016.08.016](https://doi.org/10.1016/j.annepidem.2016.08.016).
- 9 陶立元, 刘珏. 目标仿真试验的基本原理、设计要素及其优缺点[J]. *中华流行病学杂志*, 2023, 44(12): 1963–1969. [Tao LY, Liu J. Basic principles, design elements, advantages and challenges of emulated target trial[J]. *Chinese Journal of Epidemiology*, 2023, 44(12): 1963–1969.] DOI: [10.3760/cma.j.cn112338-20230515-00300](https://doi.org/10.3760/cma.j.cn112338-20230515-00300).
 - 10 周佳薇, 黄丽红, 尤东方, 等. 利用现实世界数据仿真临床试验: 目标试验的发展与应用[J]. *中华流行病学杂志*, 2024, 49(2): 279–285. [Zhou JW, Huang LH, You DF, et al. The emulation of clinical trials with real-world data: development and application of target trial[J]. *Chinese Journal of Epidemiology*, 2024, 49(2): 279–285.] DOI: [10.3760/cma.j.cn112338-20230821-00081](https://doi.org/10.3760/cma.j.cn112338-20230821-00081).
 - 11 贾玉龙, 孙鑫, 谭婧, 等. 基于真实世界数据的中西药联用疗效评价目标仿真试验方法学探讨[J]. *中华流行病学杂志*, 2025, 46(11): 2040–2046. [Jia YL, Sun X, Tan J, et al. Methodological discussion on target trial emulation for efficacy evaluation of combination of traditional Chinese medicine and western medicine with real-world data[J]. *Chinese Journal of Epidemiology*, 2025, 46(11): 2040–2046.] DOI: [10.3760/cma.j.cn112338-20250328-00205](https://doi.org/10.3760/cma.j.cn112338-20250328-00205).
 - 12 马高亭, 左颖婷, 雷少元, 等. 模仿目标临床试验在脑血管疾病领域的应用现状及展望[J]. *中华脑血管病杂志(电子版)*, 2024, 18(3): 277–280. [Ma GT, Zuo YT, Lei SY, et al. Current applications and future prospects of target trial emulation in the field of cerebrovascular diseases[J]. *Chinese Journal of Cerebrovascular Diseases (Electronic Edition)*, 2024, 18(3): 277–280.] DOI: [10.11817/j.jssn.1673-9248.2024.03.014](https://doi.org/10.11817/j.jssn.1673-9248.2024.03.014).
 - 13 卢存存, 陈子佳, 王志飞. 基于真实世界数据的观察性因果推断研究新框架(目标试验模拟)及其在中医药领域中的应用展望[J]. *协和医学杂志*, 2024, 15(2): 422–428. [Lu CC, Chen ZJ, Wang ZF. Novel framework (target trial emulation) in observational causal inference research based on real world data and its application prospects in traditional Chinese medicine[J]. *Medical Journal of Peking Union Medical College Hospital*, 2024, 15(2): 422–428.] DOI: [10.12290/hyxxz.2023-0398](https://doi.org/10.12290/hyxxz.2023-0398).
 - 14 石舒原, 赵厚宇, 尹旭辉, 等. 目标试验仿真研究概念及其应用价值的再认识[J]. *中华流行病学杂志*, 2025, 46(12): 2240–2246. [Shi SY, Zhao HY, Yin XH, et al. Re-understanding of the concept and application value of target trial emulation research[J]. *Chinese Journal of Epidemiology*, 2025, 46(12): 2240–2246.] DOI: [10.3760/cma.j.cn112338-20250407-00215](https://doi.org/10.3760/cma.j.cn112338-20250407-00215).
 - 15 龙子临, 赵厚宇, 刘佐相, 等. 目标试验仿真的全球进展与趋势: 文献计量分析与可视化[J]. *中华流行病学杂志*, 2025, 46(7): 1262–1268. [Long ZL, Zhao HY, Liu ZX, et al. Global progress and trend in research of target trial emulation: a bibliometric and visualization analysis[J]. *Chinese Journal of Epidemiology*, 2025, 46(7): 1262–1268.] DOI: [10.3760/cma.j.cn112338-20241105-00694](https://doi.org/10.3760/cma.j.cn112338-20241105-00694).
 - 16 刘佐相, 龙子临, 杨智荣, 等. 临床干预措施效力-效果差距弥合的方法学研究进展(一): 改善真实世界研究的效果估计[J]. *中华流行病学杂志*, 2024, 45(2): 286–293. [Liu ZX, Long ZL, Yang ZR, et al. Progress in methodological research on bridging the efficacy-effectiveness gap of clinical interventions (I): to improve the validity of real-world evidence[J]. *Chinese Journal of Epidemiology*, 2024, 45(2): 286–293.] DOI: [10.3760/cma.j.cn112338-20230925-00189](https://doi.org/10.3760/cma.j.cn112338-20230925-00189).
 - 17 刘佐相, 龙子临, 赵厚宇, 等. 基于真实世界数据的随机对照试验模拟研究设计、评价与实例分析[J]. *中国食品药品监管*, 2023(10): 124–131. [Liu ZX, Long ZL, Zhao HY, et al. Design framework, evaluation and case application of randomized controlled trial emulation research[J]. *China Food Drug Administration*, 2023(10): 124–131.] DOI: [10.3969/j.issn.1673-5390.2023.10.014](https://doi.org/10.3969/j.issn.1673-5390.2023.10.014).
 - 18 中国康复医学会康复大数据工作委员会. 目标仿真试验临床研究应用规范专家共识(2025版)[J]. *中华医学杂志*, 2025, 105(46): 4239–4250. [Expert Group for the Development of the Expert Consensus on the Standardized Application of Emulated Target Trials. Expert consensus on the standardized application of emulated target trials in clinical research (2025 edition)[J]. *National Medical Journal of China*, 2025, 105(46): 4239–4250.] DOI: [10.3760/cma.j.cn112137-20250722-01821](https://doi.org/10.3760/cma.j.cn112137-20250722-01821).
 - 19 Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods[J]. *Stat Med*, 2019, 38(11): 2074–2102. DOI: [10.1002/sim.8086](https://doi.org/10.1002/sim.8086).
 - 20 Her QL, Rouette J, Young JC, et al. Core concepts in pharmacoepidemiology: new-user designs[J]. *Pharmacoepidemiol Drug Saf*, 2024, 33(12): e70048. DOI: [10.1002/pds.70048](https://doi.org/10.1002/pds.70048).
 - 21 Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application[J]. *Curr Epidemiol Rep*, 2015, 2(4): 221–228. DOI: [10.1007/s40471-015-0053-5](https://doi.org/10.1007/s40471-015-0053-5).
 - 22 Xie Y, Bowe B, Xian H, et al. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of major adverse cardiovascular events: emulation of a randomised target trial using electronic health records[J]. *Lancet Diabetes Endocrinol*, 2023, 11(9): 644–656. DOI: [10.1016/S2213-8587\(23\)00171-7](https://doi.org/10.1016/S2213-8587(23)00171-7).
 - 23 Xie Y, Bowe B, Gibson AK, et al. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of kidney outcomes: emulation of a target trial using Health Care Databases[J]. *Diabetes Care*, 2020, 43(11): 2859–2869. DOI: [10.2337/dc20-1890](https://doi.org/10.2337/dc20-1890).
 - 24 Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology[J]. *Pharmacoepidemiol Drug Saf*, 2020, 29(9): 1101–1110. DOI: [10.1002/pds.5083](https://doi.org/10.1002/pds.5083).
 - 25 Roberts AW, Dusetzina SB, Farley JF. Revisiting the washout period in the incident user study design: why 6–12 months may not be sufficient[J]. *J Comp Eff Res*, 2015, 4(1): 27–35. DOI: [10.2217/ceer.14.53](https://doi.org/10.2217/ceer.14.53).
 - 26 Danaei G, Rodríguez LAG, Cantero OF, et al. Observational data for comparative effectiveness research: an emulation of

- randomised trials of statins and primary prevention of coronary heart disease[J]. *Stat Methods Med Res*, 2013, 22(1): 70–96. DOI: [10.1177/0962280211403603](https://doi.org/10.1177/0962280211403603).
- 27 Xu W, Yau YK, Pan Y, et al. Effectiveness and safety of using statin therapy for the primary prevention of cardiovascular diseases in older patients with chronic kidney disease who are hypercholesterolemic: a target trial emulation study[J]. *Lancet Healthy Longev*, 2025, 6(3): 100683. DOI: [10.1016/j.lanhl.2025.100683](https://doi.org/10.1016/j.lanhl.2025.100683).
- 28 Zhuo L, Zhang B, Huang Y, et al. Association of influenza vaccination with acute kidney injury: a prospective target trial emulation study[J]. *Am J Kidney Dis*, 2026, 87(1): 75–86. e1. DOI: [10.1053/j.ajkd.2025.09.005](https://doi.org/10.1053/j.ajkd.2025.09.005).
- 29 Suissa S. Immortal time bias in pharmaco-epidemiology[J]. *Am J Epidemiol*, 2008, 167(4): 492–499. DOI: [10.1093/aje/kwm324](https://doi.org/10.1093/aje/kwm324).
- 30 Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments[J]. *Epidemiology*, 2008, 19(6): 766–779. DOI: [10.1097/EDE.0b013e3181875e61](https://doi.org/10.1097/EDE.0b013e3181875e61).
- 31 Li S, Rezvani R, Seaman SR, et al. TrialEmulation: an R package to emulate target trials for causal analysis of observational time-to-event data[EB/OL]. (2024–02–19) [2026–01–30]. <https://doi.org/10.48550/arXiv.2402.12083>.
- 32 Penning De Vries BBL, Groenwold RHH. Bias of time-varying exposure effects due to time-varying covariate measurement strategies[J]. *Pharmacoepidemiol Drug Saf*, 2022, 31(1): 22–27. DOI: [10.1002/pds.5328](https://doi.org/10.1002/pds.5328).
- 33 Zhao SS, Lyu H, Yoshida K. Versatility of the clone-censor-weight approach: response to "trial emulation in the presence of immortal-time bias"[J]. *Int J Epidemiol*, 2021, 50(2): 694–695. DOI: [10.1093/ije/dyaa223](https://doi.org/10.1093/ije/dyaa223).
- 34 Gaber CE, Hanson KA, Kim S, et al. The clone-censor-weight method in pharmacoepidemiologic research: foundations and methodological implementation[J]. *Curr Epidemiol Rep*, 2024, 11(3): 164–174. DOI: [10.1007/s40471-024-00346-2](https://doi.org/10.1007/s40471-024-00346-2).
- 35 Emilsson L, García-Albéniz X, Logan RW, et al. Examining bias in studies of statin treatment and survival in patients with cancer[J]. *JAMA Oncol*, 2018, 4(1): 63–70. DOI: [10.1001/jamaoncol.2017.2752](https://doi.org/10.1001/jamaoncol.2017.2752).
- 36 Hernán MA. How to estimate the effect of treatment duration on survival outcomes using observational data[J]. *BMJ*, 2018, 360: k182. DOI: [10.1136/bmj.k182](https://doi.org/10.1136/bmj.k182).
- 37 Fu EL, Evans M, Clase CM, et al. Stopping renin-angiotensin system inhibitors in patients with advanced CKD and risk of adverse outcomes: a nationwide study[J]. *J Am Soc Nephrol*, 2021, 32(2): 424–435. DOI: [10.1681/ASN.2020050682](https://doi.org/10.1681/ASN.2020050682).
- 38 Cain LE, Robins JM, Lanoy E, et al. When to start treatment? a systematic approach to the comparison of dynamic regimes using observational data[J]. *Int J Biostat*, 2010, 6(2): 1–24. DOI: [10.2202/1557-4679.1212](https://doi.org/10.2202/1557-4679.1212).
- 39 Petito LC, García-Albéniz X, Logan RW, et al. Estimates of overall survival in patients with cancer receiving different treatment regimens: emulating hypothetical target trials in the Surveillance, Epidemiology, and End Results (SEER)-Medicare Linked Database[J]. *JAMA Netw Open*, 2020, 3(3): e200452. DOI: [10.1001/jamanetworkopen.2020.0452](https://doi.org/10.1001/jamanetworkopen.2020.0452).
- 40 Sjolander A, Nyren O, Bellocco R, et al. Comparing different strategies for timing of dialysis initiation through inverse probability weighting[J]. *Am J Epidemiol*, 2011, 174(10): 1204–1210. DOI: [10.1093/aje/kwr249](https://doi.org/10.1093/aje/kwr249).
- 41 García-Albéniz X, Hernán MA, Logan RW, et al. Continuation of annual screening mammography and breast cancer mortality in women older than 70 years[J]. *Ann Intern Med*, 2020, 172(6): 381–389. DOI: [10.7326/M18-1199](https://doi.org/10.7326/M18-1199).
- 42 Chen A, Ju C, Mackenzie IS, et al. Impact of beta-blockers on mortality and cardiovascular disease outcomes in patients with obstructive sleep apnoea: a population-based cohort study in target trial emulation framework[J]. *Lancet Reg Health Eur*, 2023, 33: 100715. DOI: [10.1016/j.lanepe.2023.100715](https://doi.org/10.1016/j.lanepe.2023.100715).
- 43 Boyne DJ, Cheung WY, Hilsden RJ, et al. Association of a shortened duration of adjuvant chemotherapy with overall survival among individuals with stage III colon cancer[J]. *JAMA Network Open*, 2021, 4(3): e213587. DOI: [10.1001/jamanetworkopen.2021.3587](https://doi.org/10.1001/jamanetworkopen.2021.3587).
- 44 Maringe C, Benitez Majano S, Exarchakou A, et al. Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data[J]. *Int J Epidemiol*, 2020, 49(5): 1719–1729. DOI: [10.1093/ije/dyaa057](https://doi.org/10.1093/ije/dyaa057).
- 45 Murray EJ, Marshall BDL, Buchanan AL. Emulating target trials to improve causal inference from agent-based models[J]. *Am J Epidemiol*, 2021, 190(8): 1652–1658. DOI: [10.1093/aje/kwab040](https://doi.org/10.1093/aje/kwab040).
- 46 Kirkegård J, Gaber C, Heide-Jørgensen U, et al. Effect of surgery versus chemotherapy in pancreatic cancer patients: a target trial emulation[J]. *J Natl Cancer Inst*, 2024, 116(7): 1072–1079. DOI: [10.1093/jnci/djae024](https://doi.org/10.1093/jnci/djae024).
- 47 Dickerman BA, García-Albéniz X, Logan RW, et al. Emulating a target trial in case-control designs: an application to statins and colorectal cancer[J]. *Int J Epidemiol*, 2020, 49(5): 1637–1646. DOI: [10.1093/ije/dyaa144](https://doi.org/10.1093/ije/dyaa144).
- 48 Rasouli B, Chubak J, Floyd JS, et al. Combining high quality data with rigorous methods: emulation of a target trial using electronic health records and a nested case-control design[J]. *BMJ*, 2023; e72346. DOI: [10.1136/bmj-2022-072346](https://doi.org/10.1136/bmj-2022-072346).
- 49 Wang X, Zou Y, Wang Y, et al. Estimating heterogeneous treatment effects with target trial emulation: a checklist of causal machine learning for observational data[J]. *Chest*, 2025, 168(4): 856–859. DOI: [10.1016/j.chest.2025.05.028](https://doi.org/10.1016/j.chest.2025.05.028).

收稿日期: 2026年01月30日 修回日期: 2026年02月09日
 本文编辑: 洗静怡 杨燕