arXiv:2603.04544v1 [stat.ME] 4 Mar 2026

含外部对照试验的近端学习：一项HIV预防案例研究

Yilin Song[1],Yinxiang Wu[2], Raphael J. Landovitz[3], Susan Buchbinder[4], Srilatha Edupuganti[5], Lydia Soto-Torres[6], Kendrick Li[7], Xu Shi[8], FeiGao[9], Deborah Donnell[9],Holly Janes[8],和 Ting Ye[2]*

[1]哥伦比亚大学生物统计学系[2]华盛顿大学生物统计学系[3]加州大学洛杉矶分校大卫·格芬医学院[4]旧金山公共卫生局Bridge HIV[5]埃默里大学医学系[6]美国国家过敏症和传染病研究所艾滋病司[7]圣裘德儿童研究医院[8]密歇根大学生物统计学系[9]弗雷德·哈钦森癌症研究中心疫苗与传染病部

摘要

随着有效的暴露前预防药物问世，主动对照HIV预防试验已成为常见的研究设计。尽管如此，估计相对于安慰剂的绝对疗效仍然重要。本文介绍近端因果推断方法的一项新应用：利用来自入选标准相似的安慰剂对照试验的外部对照数据，估计卡博特韦主动对照试验参与者在安慰剂条件下的反事实HIV累积发病率。我们分别将基线性传播感染状态和地理地区作为负对照结局和暴露变量。我们应对两项关键挑战：各试验之间HIV风险的未测量差异，以及两项研究中HIV发病率较低所导致的统计困难。为克服这些挑战，我们开发了两种近端推断方法：（1）半参数逆删失概率加权估计量；（2）针对低事件率情境的基于回归的两阶段策略。理论与数值研究表明，这些方法能够得到安慰剂条件下反事实一年累积HIV发生率的可靠估计值，并为卡博特韦相较于安慰剂的优越疗效提供有力证据。这些发现凸显了近端推断方法借助外部对照，在单臂试验和主动对照试验中估计安慰剂对照效应的潜力。

关键词：主动对照试验、因果推断、删失、数据整合

*请联系 tingye1@uw.edu。

1 HIV预防临床试验

1.1 HPTN 083作为主要关注的临床试验

HPTN 083（ClinicalTrials.gov：NCT02720094）是一项 2b/3期随机临床试验，旨在评估长效注射用卡博特韦与每日口服富马酸替诺福韦二吡呋酯/恩曲他滨（TDF/FTC）的安全性和疗效。后者是首个经证实可通过暴露前预防（PrEP）预防人类免疫缺陷病毒感染的生物医学药物（Landovitz等，2021）。该临床试验于2016年12月至2020年3月期间在美国、拉丁美洲、亚洲和非洲招募参与者。意向治疗（ITT）人群包括4,566名顺性别男性和与男性发生性行为的跨性别女性，其中2,282名参与者被随机分配至卡博特韦组，2,284名被分配至TDF/FTC组。

在意向治疗人群中，事后发现有5名参与者（卡博特韦组2名，TDF/FTC组3名）在基线时已感染人类免疫缺陷病毒，另有71名参与者入组后未接受任何随访访视。所有这些参与者均被排除在主要疗效分析之外。临床试验显示，卡博特韦在预防HIV感染方面优于TDF/FTC。卡博特韦组的一年HIV累积发病率为0.41%（95% CI：0.20至0.70），而TDF/FTC组为1.22%（95% CI：0.90至1.70）。

1.2 估计安慰剂条件下的反事实HIV发病率

随着 TDF/FTC 等有效暴露前预防药物问世，主动对照HIV预防试验（如 HPTN 083）已成为常见的设计（Mayer等，2020）。鉴于这些药物的疗效已得到充分确立，开展需停用已证实有效干预措施的安慰剂对照试验，必须经过审慎的伦理论证（联合国艾滋病规划署（UNAIDS）和世界卫生组织（WHO），2021）。此外，HIV预防策略的快速发展也带来了额外挑战。分配至安慰剂组的参与者必须在试验中获得备择预防措施，许多人还可能在试验之外寻求其他预防方案，这使得

难以单独确定试验药物的效果。

尽管没有同期安慰剂组，估计新干预措施的绝对疗效仍然高度相关。这涉及将新干预措施下观察到的结局与明确定义的安慰剂条件下预期的结局进行比较——即反事实安慰剂发生率（Glidden 2019，Glidden等2020，Temple和Ellenberg 2000）。进行此类估计有两个主要动机。首先，它使患者和临床医生能够清楚了解不同干预措施所提供的真实保护水平，从而做出知情决策。其次，随着暴露前预防药物变得更加有效，且HIV发病率持续下降，以HIV发病率为主要终点开展主动对照试验变得越来越不切实际和/或不可行。

在HIV预防领域，为了在没有安慰剂组的情况下估计干预措施的绝对疗效，人们提出了多种研究设计策略。 Donnell等 (2024年)介绍了六种潜在的研究设计策略，并阐述了利用以下数据时各自的优势与局限：(1) 注册队列，(2) 近期感染检测，(3) 历史外部安慰剂组，(4) HIV发生率生物标志物，(5) 药物浓度，以及(6) 免疫生物标志物。其中一些策略已应用于实际临床试验。例如，PrEPVacc研究利用试验前登记队列估计反事实安慰剂发生率（PrEPVacc2024，Dunn等，2018年，Kansiime等，2025年）。近期感染检测方法根据横断面样本中接受临床试验参与筛查者的近期感染比例，估计基线HIV发生率；该方法已用于吉利德的PURPOSE1和PURPOSE2临床试验（PURPOSE12024，PURPOSE2 2024）。Donnell等 (2023年)采用外部安慰剂组方法，通过直接标准化对观测协变量进行调整。美国食品药品监督管理局（FDA）的指导意见草案也讨论了外部安慰剂组方法的相关考量（美国食品药品监督管理局，2023年）。HIV发生率生物标志物方法的代表性研究包括 Mullick和Murray (2020年) 以及 Zhu等

(2024年)，该方法通过拟合历史数据中的模型，建立研究层面的淋病和HIV发病率之间的关联，并据此预测当前主要研究中的反事实安慰剂HIV发病率。这种方法面临复杂因素的影响：人群中抗逆转录病毒治疗的预防作用可能使性传播感染（STI）的传播与HIV传播脱钩；此外，鉴于多西环素暴露后预防已被证实有效，该方法目前对某些人群已不再适用。研究人员还探索了其他主动对照试验设计（Prudden等，2024年，Gao等，2025年）。总之，这些研究设计策略为估计反事实安慰剂发生率以及在无法纳入安慰剂组时设计HIV预防试验，提供了日益丰富的选择。

在本研究中，我们利用一项安慰剂对照研究的外部对照数据，重点估计HPTN 083试验参与者在安慰剂条件下的反事实HIV累积发病率。将这些反事实估计值与HPTN 083试验卡博特韦组中观察到的HIV累积发病率进行比较后，我们得到了绝对疗效的估计值。然而，这种方法存在方法学挑战，需要进一步发展统计方法。

1.3 HVTN704/HPTN 085作为外部对照数据集

抗体介导预防（AMP）研究（也称为HVTN 704/HPTN 085，ClinicalTrials.gov：NCT02716675）于2016年4月至2018年10月在美国、秘鲁、巴西和瑞士开展。AMP研究的入选标准与HPTN 083相似，并招募了有感染风险的顺性别男性和跨性别者，以评估被动输注VRC01广谱中和抗体（bnAb）用于HIV预防的安全性、耐受性和疗效。

AMP研究包含三个组：（1）高剂量组，接受 30 mg/kg 的 VRC01；（2）低剂量组，接受 10 mg/kg 的 VRC01；（3）安慰剂组，接受无菌生理盐水输注。低剂量组、高剂量组和安慰剂组截至 1 年时的 HIV累积发病率（%）分别为 2.5（95% CI：[1.71，3.53]）、2.2（95% CI：[1.46，3.18]）和 2.98（95% CI：[2.11，4.09]）。我们将安慰剂组作为阴性对照数据，并

排除了 $n = 11$ 随机分组后未进行任何随访访视的参与者。数据处理步骤的详细信息见图1。

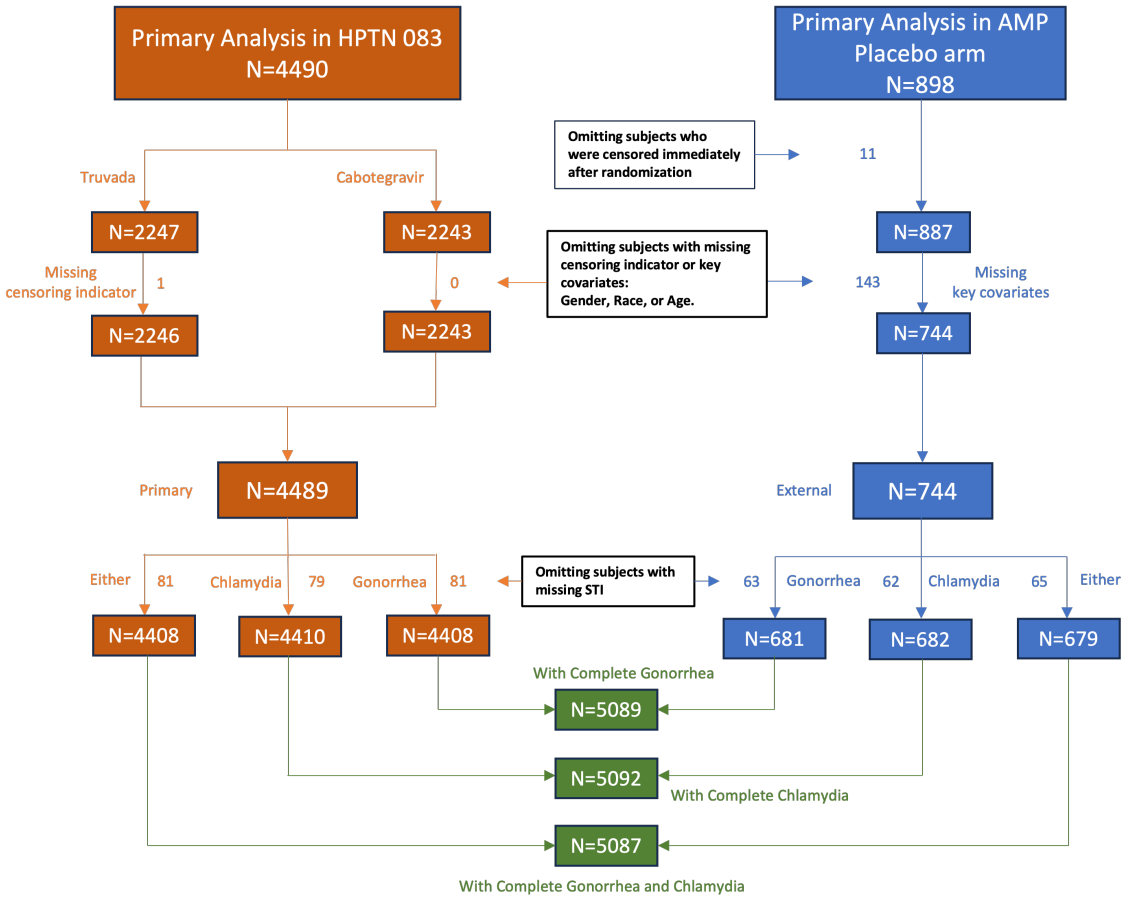


图1：数据处理流程。我们首先排除了关键协变量缺失的观测，包括性别、种族和年龄。随后，我们分别通过排除基线淋病缺失、基线衣原体感染缺失，或淋病或衣原体感染数据缺失的个体，构建了三个独立的分析数据集。基线性传播感染指标被用作负对照结局。

为了正确解读我们针对 HPTN 083 估算的反事实安慰剂发生率，明确界定 AMP研究中的安慰剂条件十分重要。在 AMP研究中，随访数据显示，39.0%（95% CI：[35.9, 42.0]）的人年检测到* TDF/FTC，28.9%（95% CI：[26.4，31.6]）的人年检测到有效的† TDF/FTC，

*定义为检测值达到或超过定量下限

†定义为检测值至少为 700 fmol/冲切样本

这些比例在各治疗组之间均匀分布。这体现了现代 HIV预防试验中的现实情况：很少有安慰剂对照研究是在未同时使用生物医学预防的情况下开展的。相比之下，在 HPTN 083 的 TDF/FTC 组中，随访数据显示，72.3% 的样本检测到的 TDF/FTC 浓度与每周至少服用四剂 TDF-FTC 相符。因此，AMP研究中的安慰剂条件反映了额外生物医学预防的使用；可以说，在这种情况下，这是对安慰剂具有临床相关性的定义。

为评估两项临床试验的可比性，我们考察了HPTN 083参与者以及AMP安慰剂组参与者的基线特征（表1）。由于不同临床试验中测量的变量很少完全相同，即使是在同一网络内、相近时期内开展的试验亦是如此，我们通过与各试验方案团队的讨论，对变量定义进行了协调统一。如表1所示，AMP安慰剂组中白人参与者比例更高，且30岁以上个体所占百分比更大。相比之下，HPTN 083纳入了更多黑人参与者以及更多顺性别男性。基线性传播感染（包括直肠淋病和衣原体）的流行率在HPTN 083中也更高。这些变量在分析中需要进行校正，因为它们可能与HIV感染风险相关（《柳叶刀·HIV》2020，Ward& Rönn 2010）。

这一比较凸显了使用外部对照时的一项核心挑战。试验人群之间的系统性差异可能引入混杂偏倚；即便对观测协变量进行严格校正，来自未测量变量的残余混杂仍可能持续存在。例如，区域HIV流行率、伴侣的病毒载量以及当地性网络密度都是影响个体感染风险的重要情境因子，但它们难以测量，且两项研究均未记录这些信息。由于两项临床试验从不同地理区域招募参与者，当地HIV传播动态中未被测量的差异可能导致安慰剂条件下研究间HIV发病率的差异，而这些差异无法完全由已测量变量解释。

表1：HPTN 083 和 AMP 安慰剂组参与者的基线特征。人口统计学特征通过基线问卷收集。基线淋病和衣原体感染情况通过直肠拭子的实验室检测确定。

	HPTN 083 ($n = 4490$)	AMP 安慰剂组 ($n = 898$)
年龄, n (%)		
18–20 岁	634 (14.1)	95（10.6）
21–30 岁	2557（56.9）	493（54.9）
31–40 岁	907（20.2）	230 (25.6)
$\geq$ 40年	392 (8.7)	80 (8.9)
种族, n (%)		
白人	1231 (27.4)	285 (31.7)
黑人	1074 (23.9)	131 (14.6)
其他	2185 (48.7)	482 (53.7)
性别, n (%)		
女性或跨性别女性	480 (10.7)	33 (3.7)
Male	3899 (86.8)	688 (76.6)
其他	111 (2.5)	29 (3.2)
缺失	0 (0.0)	148 (16.5)
地区, n (%)		
非拉丁美洲	2562 (57.1)	470 (52.3)
拉丁美洲	1928 (42.9)	428 (47.6)
族裔, n （%）		
西班牙裔/拉丁裔	2065 (46.0)	529 (58.9)
非西班牙裔/拉丁裔	2424 (54.0)	369 (41.1)
缺失	1 (0.0)	0 (0.0)
教育程度, n (%)		
大专及以上	3431 (76.4)	575 (64.0)
高中及以下	1059 (23.5)	175 (19.5)
缺失	0 (0.0)	148 (16.5)
直肠淋病（%）		
Yes	293 (6.5)	33 (3.7)
No	4116 (91.7)	654 (72.8)
缺失	81 (1.8)	211 (23.5)
直肠衣原体感染（%）		
Yes	495 (11.0)	65 (7.2)
No	3916 (87.2)	623 (69.4)
缺失	79 (1.8)	210 (23.4)
直肠淋病或衣原体（%）		
Yes	679 (15.1)	82 (9.1)
No	3730 (83.1)	603 (67.1)
缺失	81 (1.8)	213 (23.7)

1.4 现有外部对照统计方法的不足

评估历史对照或外部对照可接受性的一般标准由Pocock（1976年）概述。这些标准强调，确保主要试验与外部数据来源在患者群体、研究设计和试验实施方面具有可比性至关重要。这些基本原则为后续用于纳入外部对照的统计方法奠定了基础。

在不存在未测量混杂的假设下，已发展出多种纳入外部对照数据的方法。该假设要求主要数据集和外部数据集中的参与者在观测变量的每一水平上都可交换——这是一个无法凭经验检验的强假设。当该假设成立时，研究者可以应用标准因果推断方法，包括倾向评分匹配（Rubin 1974年，Dehejia与Wahba 2002年，Li 等人 2001年，Li 2013年，Chen等人 2022年，Stuart与Rubin 2008年）、逆概率加权（Joffe等人 2004，Chesnaye等人 2021，Thorlund等人 2020）、G计算（Snowden等人 2011年，Robins 1986年，Hernán与Robins 2020），以及双重/去偏机器学习和半参数理论（Chernozhukov等人 2018年，Belloni等人 2017年，Liu 等人 2025）。此外，还有若干专门针对外部对照场景开发的方法（Li 等人 2023年，Valancius等人 2024年）。所有这些方法都利用观测协变量来提高主要人群与外部人群之间的可比性，并校正潜在混杂。

未测量混杂因素（例如，当地的 HIV 流行率、性伴侣的 HIV 病毒载量、当地的性网络密度，以及当地在 HIV 预防方面的行为和文化规范）可能同时影响 HIV 感染风险和参与研究的可能性，尤其是因为这两项临床试验招募的参与者来自不同的地理区域。在这种情况下，依赖于不存在未测量混杂这一假设的方法可能会产生有偏估计。如果主要数据集和外部数据集中都有安慰剂数据，研究人员便可以在合并数据来源之前正式评估其可比性（Viele 等人，2014，Liu 等人，2022年）。敏感性分析框架，例如Yi 等人 (2023年) 提出的方法，可以

帮助量化未测量混杂可能造成的影响。

当有多个历史或外部试验可用时，可使用荟萃分析技术对研究间的异质性进行建模，并将研究结果外推至目标人群（Dias 等人 2013年）。贝叶斯借用方法提供了一种互补策略，可根据外部数据与主要试验人群的相似程度，对外部数据的贡献进行加权。此类方法包括幂先验、弹性先验和相称先验（Ibrahim& Chen 2000、Jiang 等人 2023年、Hobbs 等人 2012、Schmidli 等人 2017年）。然而，贝叶斯框架往往对模型假设较为敏感，且可能需要大量领域知识来指导先验设定。

因此，在利用外部对照时，开发能够处理未测量混杂的统计方法仍存在空缺，尤其是在主要试验缺少安慰剂组的情况下。较低的人类免疫缺陷病毒发病率可能带来额外的统计挑战，包括观察到的事件数量有限导致精度降低，以及估计不稳定；在这种情况下，点估计及其置信区间都可能落在合理范围之外，例如，人类免疫缺陷病毒累积发病率的合理范围为（0，100%）。最后，较低的人类免疫缺陷病毒发病率可能会降低标准统计分析的功效，并带来额外的统计挑战，因此需要采用效率更高的方法

1.5 近端因果推断方法综述

近端因果推断方法通过利用负对照暴露（NCEs）和负对照结局（NCOs），旨在解决观察性研究中的未测量混杂（Miao 等，2018，Shi、Miao、Nelson 和 Tchetgen Tchetgen 2020）。负对照结局是指不受所关注治疗因果影响的变量，而负对照暴露是指不会对所关注结局产生因果影响的变量。重要的是，负对照结局和负对照暴露都必须与影响主要暴露-结局关系的未测量混杂因素相关。

该领域的关键基础性和综述性文章包括 Tchetgen Tchetgen 等人（2024年）、Shi、Miao、Nelson 和 Tchetgen Tchetgen （2020）以及 Shi、Miao 和 TchetgenTchetgen

（2020）。在这项奠基性工作的基础上，Cui 等人 （2024年）提出了一个半参数框架，用于识别和估计平均处理效应（ATE）和处理组平均处理效应（ATT）。Ying 等人 （2022年）将 ATE 的估计扩展至事件时间结局。近期，研究人员开发了基于回归的方法进行近端推断，并采用广义线性模型进行结局建模（Liu 等人 2024年、Li 等人 2025）。

与我们的研究最相关的是，Su 等人（2025）提出了一种近端间接比较框架，适用于目标治疗在主要随机对照试验中不可用、但在外部随机对照试验中可用的情形。他们的方法利用代理变量和桥接函数，调整数据集之间未测量且发生偏移的效应修饰因子，从而得到一个双重稳健且渐近正态的估计量。然而，他们的识别策略要求两项临床试验共享一个共同的第三治疗组，而我们的研究并非如此。

1.6 方法开发与 HPTN 083 应用

HPTN 083 和 AMP研究收集的丰富基线协变量，使我们能够识别适用于近端推断的合理负对照变量。我们将当地 HIV 传播环境风险定义为一项综合指标，涵盖当地 HIV流行率、伴侣的病毒载量、当地性网络密度，以及有关 HIV预防的行为和文化规范。在跨临床试验进行比较时，它是我们主要关注的未测量混杂因素，而负对照变量则充当该因子的代理变量。基线时诊断出的性传播感染（例如淋病和衣原体）是合适的负对照结局，因为它们与 HIV 感染具有关键的行为和生物学风险因素，且与当地 HIV 传播环境风险密切相关。此外，除通过其与共同混杂因素的关联外，性传播感染预计不会直接影响个体参加哪项研究——两项临床试验的性传播感染资格标准相似，性传播感染不会影响临床试验覆盖区域的选择，而且通常在调整其他协变量后，不会直接影响参与意愿。地理区域是合理的负对照暴露，因为它与当地 HIV 传播环境风险相关

但不会直接影响个体的 HIV风险。这些强代理变量的存在，支持在此情境中应用近端推断方法（见第2.1 节和图2）。

我们的目标是利用 AMP研究的外部对照数据，估计 HPTN 083试验中安慰剂组的反事实HIV累积发病率。为解决两项试验之间未测量的HIV风险差异，我们将Cui 等人(2024年) 和Ying 等人(2022年) 提出的半参数近端推断方法扩展到可处理外部对照数据与事件时间结局的整合。此外，为克服两项研究中 HIV发病率较低所带来的统计挑战，我们开发了一种新颖的两阶段基于回归的近端推断方法，该方法基于 Cox比例风险模型。

本文其余部分安排如下。第2节介绍符号和研究设定，第2.1节随后详细阐述所提出的近端识别和推断程序。第2.2节。在第3节，我们对 HPTN 083数据集进行分析，并讨论统计推断结果。最后，我们在第4节进行讨论并总结全文。

2 方法

2.1 符号与设定

我们将含有 J 个活性治疗组的主要数据集记为 $R=0$ ，将外部对照数据集记为 $R=1$。令 A 表示治疗分配，其中 $A=0$ 表示安慰剂， $A=1, \ldots\ldots, J$ 表示活性治疗组。除非另有说明，治疗分配部分由数据来源决定：具体而言， $R=1$ 对应于安慰剂组（$A=0$），$R=0$ 对应于活性处理（$A=1, \ldots\ldots, J$）。在我们的应用中，主要数据集（HPTN 083）包含 $J=2$ 个活性治疗组：卡博特韦（$A=1$）和TDF/FTC（$A=2$）。令 $T=\min(T^*, C)$表示观察随访时间，其中 T^* 为HIV感染时间， C 为删失时间。我们将事件指示变量定义为 $\Delta=I(T^* \leq C)$ ，其中 $\Delta=1$ 表示观察到HIV感染， $\Delta=0$ 表示右删失。我们用 $\boldsymbol{X}$ 表示基线协变量向量。

此外，我们在主要数据集和外部数据集的基线均观察到负对照结局 W。进一步地，对于两阶段回归方法，在两个数据集中均可观察到负对照暴露 Z，但半参数方法并不要求主要数据集中具备该暴露。对于规模为 n的合并数据集中的每个个体，我们有（R，T，Δ，$\boldsymbol{X}$，Z，W）。

令 $T^*(a)$ 表示在治疗组 $a \in \{0, 1, \ldots, J\}$ 下 HIV感染时间的反事实。我们假设稳定单元处理值假设（SUTVA）（Rubin 1980），该假设意味着标准的一致性假设 $T^* = T^*(A)$。我们的主要目标估计量是 HPTN 083 中的反事实安慰剂风险，定义为在安慰剂条件下截至时间 t发生 HIV感染的概率：$P(T^*(0) \leq t \mid R = 0)$。当 t对应一年时，也可将其解释为一年HIV累积发病率。我们还关注 HPTN 083 人群中的平均处理效应，其定义为对比 $P(T^*(a) \leq t \mid R = 0) - P(T^*(0) \leq t \mid R = 0)$，其中 $a = 1, \cdots, J$。在 HPTN 083 内进行随机分组的条件下，治疗组的潜在结局是可识别的，因此对于 $a = 1, \ldots, J$，$P(T^*(a) \leq t \mid R = 0) = P(T^* \leq t \mid R = 0, A = a)$。

为应对未测量混杂可能导致的偏差，我们假定存在一个未测量变量 U，使得主要参与者与外部参与者在同时以 U 和 $\boldsymbol{X}$ 为条件时具有可比性（Rosenbaum& Rubin 1983*a*,*b*, Imbens 2003）。U 的存在意味着，仅通过调整 $\boldsymbol{X}$ 无法实现可比性。例如，HPTN 083与AMP研究中的参与者可能在当地 HIV 传播环境风险方面存在差异；该因素会影响HIV风险，但往往未能以足够的空间粒度或与试验时期的时间对齐进行测量。因此，即使在以 $\boldsymbol{X}$ 为条件之后，安慰剂下的反事实结局分布在不同数据集之间仍可能不同：$P(T^*(0) \leq t \mid \boldsymbol{X}, R = 0) \neq P(T^*(0) \leq t \mid \boldsymbol{X}, R = 1)$。这种可交换性的缺失表明，需要能够处理未测量混杂的方法，例如近端推断。

近端推断利用负对照暴露 Z 和负对照结局

W 满足以下关键假设（Tchetgen Tchetgen 等，2024年）：**假设1**（负对照）。
$(Z,\ R) \perp\!\!\!\perp (T^*(0)\ W) \mid \boldsymbol{X},\ U$。

假设1指出，存在一个未测量混杂因素 U ，在同时给定 U 和观测协变量 $\boldsymbol{X}$的条件下，来自主要临床试验和外部临床试验的参与者具有可比性（Rosenbaum & Rubin1983a，b，Imbens 2003）。此外，该假设要求负对照暴露 Z 不会直接影响主要结局 $T^*(0)$ 或负对照结局 W，且 W不会直接影响试验参与 R。这些条件使得 Z 和 W 能够在近端推断框架中作为未测量混杂因素 U 的有效代理变量。

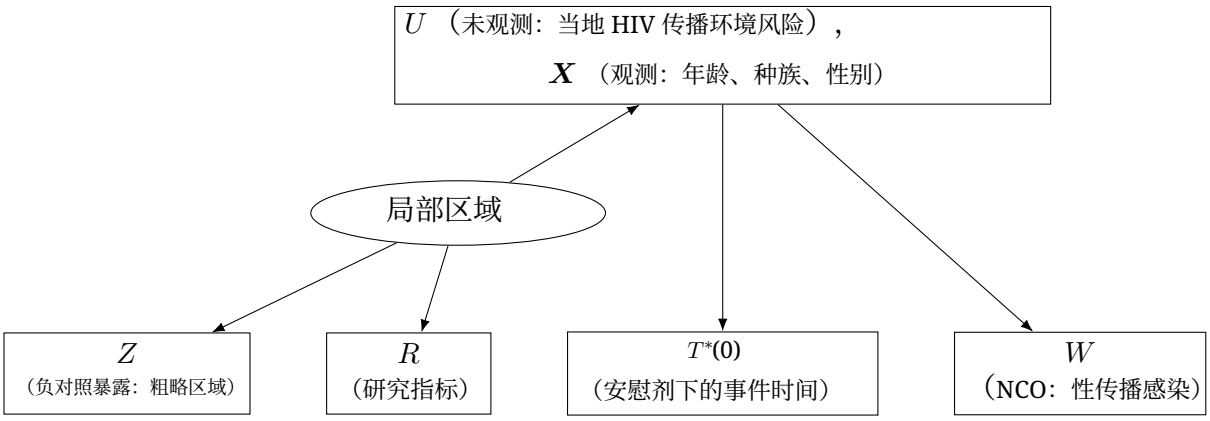


图 2: 描绘支撑假设的因果关系的因果有向无环图 n

1。未测量混杂因素 U（例如当地 HIV 传播环境风险）与观测协变量 $\boldsymbol{X}$ 共同影响安慰剂条件下的潜在 HIV 结局 $T^*(0)$、负对照结局 W （例如性传播感染状态）以及研究指标 R。粗略区域变量 Z （例如拉丁美洲与非拉丁美洲）通过依赖于局部区域而充当U 的代理变量。在这种结构下， Z 和 W 分别满足负对照暴露和负对照结局有效性的条件。

在我们的应用中，变量之间假定的因果关系由图2中的有向无环图（有向无环图）表示。一个关键的未测量混杂因素 U 是精细地理层级上的当地 HIV 传播环境风险。这一精细

地理位置变量（在有向无环图中标记为“局部区域”）会影响粗略区域变量 Z（例如，拉丁美洲与非拉丁美洲）和未测量的 HIV 流行率U，因此 Z可作为 U的代理变量。地理位置也决定研究可及性，因为入组通常取决于研究地点。因此，未测量的 HIV 流行率 U 通过地理区域与研究分配 R 相关，同时也会直接影响个体感染 HIV 的风险，使其成为未测量混杂的来源。重要的是，粗略区域变量 Z 不会在其通过U 及协变量 $\boldsymbol{X}$产生的关联之外，直接影响 HIV 风险；这些协变量包括个体层面的 HIV 风险因素。此外，负对照结局 W（例如基线性传播感染）可以与 HIV 感染具有相似的风险因素（U， $\boldsymbol{X}$），但不会直接影响研究指标 R ，因为它不会影响临床试验覆盖区域的选择，不属于任一临床试验的纳入与排除标准，并且在调整其他协变量后，通常不会直接影响参与意愿。

2.2 一种近端因果推断方法

2.2.1 方法1：半参数逆概率删失加权估计量

本节将在近端因果推断框架下，介绍一种基于逆概率删失加权（IPCW）的半参数识别与估计策略。首先，我们将介绍两个关键假设——正值性和完备性。这两个假设在近端推断文献中十分常见（Miao 等，2018、Cui 等，2024），并且可以直接推广到我们的研究情境中。

假设2 （正值性）。对于某个常数 $c > 0$, $c < P(R = 1 \mid \boldsymbol{X}, U) < 1 - c$成立$P_{U,\ \boldsymbol{X}}(\cdot|R = 0)$-几乎必然。

此条件确保，对于主要数据集（$R = 0$）中每个具有正概率的（$\boldsymbol{X}$， U）组合，在外部数据集（$R = 1$）中观察到该（$\boldsymbol{X}$， U）组合的概率也为正。请注意，应根据需要对协变量 $\boldsymbol{X}$ 进行协调（例如，通过重新分类），以确保主要数据集中的 $\boldsymbol{X}$ 所有水平也存在于外部数据集中。

假设3（完备性）。 对于任意平方可积函数 g 以及所有 $\boldsymbol{x}$，满足：***(a)*** 对所有 u，只要 $E[g(U) \mid Z = z, R = 1, \boldsymbol{X} = \boldsymbol{x}] = 0$，则 $g(u) = 0$；***(b)*** 对所有 u，只要 $E[g(U) \mid W = w, R = 1, \boldsymbol{X} = \boldsymbol{x}] = 0$，则 $g(u) = 0$，对所有 w。

近端推断方法的有效性在很大程度上取决于所识别的负对照变量是否能充分提供有关未测量混杂因素的信息。假设3为定理1中的识别提供了一个充分条件。当 U、 Z和 W 为分类变量时，这一假设更容易理解。在这种情况下，假设3要求 Z 和 W 的类别数至少与 U一样多，并且可以根据 Z 和 W的变化还原 U 的变化。更具体地说，在$R = 1$ 中，对于 $\boldsymbol{x}$的每个可能取值， U 的不同取值都必须分别对应于 W 和 Z 的不同分布（Miao 等，2018；Shi、Miao与Tchetgen Tchetgen（2020））。当 U 为连续变量时，假设3要求在外部数据集中，对于 $\boldsymbol{X}$的任一水平， U 的任何函数都不能与 W 和 Z所提供的信息正交（Olivas-Martinez等（2025））。换言之， W 和 Z 需要能够充分预测 U、 U^2、 U^3，以及更一般而言， U的所有可能变化。关于满足这一假设的各种参数模型、半参数模型和非参数模型，参见Miao等（2018）的讨论。关于放宽这一假设的策略，请读者参阅Bennett等（2022）。

为了充分控制连续 U 所导致的未测量混杂（例如，当地 HIV 传播环境风险），半参数方法需要与 U高度相关的连续负对照变量。目前我们的数据中缺少此类变量，因为 Z （地区）和 W （基线性传播感染）均为二元变量。因此，我们对假设1 作出修改，假设控制当地 HIV 传播环境风险的二分版本（高或低）足以解决未测量混杂问题。**假设 1***（负对照与二元未测量混杂因素）。（Z， R）$\perp\!\!\!\perp$（T^*（0） W）$|\boldsymbol{X}$， U_b，其中 U_b 是 U的二分版本。

在补充材料中 ement S3，我们通过解析方法评估偏差的大小，并展示模拟

在二元 U_b 和观测协变量 $\boldsymbol{X}$ 未能充分调整 (Z, R) 与 $(T^*(0)\ W)$之间混杂、从而违反假设 1*的情形下所得的结果。在这种情况下，我们发现，更强的代理变量 Z 用于未测量混杂因素 U 可降低偏差并减小标准误。

此外，由于主要终点是 HIV感染时间，我们还考虑了右删失。删失来源有两个：（1）大多数病例在研究结束时被行政删失，即参与者在研究期间结束时尚未发生事件；（2）其余病例则是失访所致（例如退出）。我们假设外部研究中的删失是随机的，即删失机制可以依赖于观测协变量 $\boldsymbol{X}$ 和负对照暴露 Z，但不依赖于未测量混杂因素U 或负对照结局 W。这一假设是标准假设；当 $\boldsymbol{X}$ 包含足够丰富的参与者特征时，预计该假设近似成立。

假设 4 (删失)。 令 $\lambda_C(t \mid \cdot)$ 表示时间 t处删失的条件风险函数。我们假设 $\lambda_C(t \mid T > t, Z, \boldsymbol{X}, U, W, T^*, R = 1) = \lambda_C(t \mid T > t, Z, \boldsymbol{X}, R = 1)$。

定理1 给出了所提出的半参数逆概率删失加权方法的识别结果，并将Cui 等人（2024年）的工作扩展至考虑右删失的情形。

定理 1（半参数IPCW识别）。假设 *1*, 2-4*，且存在平方可积函数 h 和 q ，它们几乎处处满足以下积分方程：

$$E\left[\frac{\Delta I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R = 1)} - h(W, \boldsymbol{X}) \mid Z, \boldsymbol{X}, R = 1\right] = 0, \tag{1}$$

$$\frac{P(R = 0 \mid W, \boldsymbol{X})}{P(R = 1 \mid W, \boldsymbol{X})} = E[q(Z, \boldsymbol{X}) \mid W, \boldsymbol{X}, R = 1]. \tag{2}$$

那么， $P(T^*(0) \le t \mid R = 0)$ 是非参数识别的，并且可以用以下两种形式之一表示：

$$P(T^*(0) \le t \mid R = 0) = E[h(W, \boldsymbol{X}) \mid R = 0], \tag{3}$$

$$P(T^*(0) \le t \mid R = 0) = \frac{1}{P(R = 0)} E\left[\frac{Rq(Z, \boldsymbol{X})\Delta I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R = 1)}\right]. \tag{4}$$

此外，$P(T^*(0) \leq t \mid R=0)$ 可以通过以下增广形式加以识别：

$$P(T^*(0) \leq t \mid R=0) = \frac{1}{P(R=0)} E\left[Rq(Z, \boldsymbol{X}) \left\{ \frac{\Delta I(T^* \leq t)}{P(C > T^* \mid Z, \boldsymbol{X}, R=1)} - h(W, \boldsymbol{X}) \right\} + (1-R)h(W, X) \right], \tag{5}$$

只要以下任一条件成立，便仍然有效：h 满足(1) 或 q 满足(2)。

在近端推断文献中，函数 h 和 q 分别称为结果桥函数和治疗桥函数。在定理1中，我们不要求 h 和 q 具有唯一性，这放宽了Cui 等人(2024年中的假设。任何满足（(1)–(2)）的函数对都能识别出相同的目标参数。

直接求解方程（(1)）和（(2)）颇具挑战，因为它们是积分方程，其解可能不唯一且数值不稳定。如何妥善应对这些挑战仍是一个活跃的研究领域；我们建议感兴趣的读者参阅 Kress(1989)、Dikkala 等人(2020)、Ghassami 等人(2022年)以及Bennett 等人(2023年)。本节余下部分将简要介绍定理 h 和 q 所蕴含的估计程序，并假设二者均采用参数形式。

为了估计 h，我们首先需要估计条件生存概率 $P(C > T^* \mid Z, \boldsymbol{X}, R=1)$。这可以通过使用外部数据集拟合删失时间的 Cox 比例风险模型来实现，将删失视为事件，并将 $\hat{}Z$, $\boldsymbol{X}$ 作为协变量。估计的生存概率为：$P(C > T^* | Z, \boldsymbol{X}, R=1) = \exp\left\{ -\hat{}H(T^*;\, Z, \boldsymbol{X}, R=1) \right\}$，其中 $\hat{}H(T^*;\, \cdot)$ 是截至时间 T^* 的估计累积风险函数。注意，当 $\Delta = 1$ 时，我们有观测时间 $T = T^*$。接下来，我们将公式 (1) 重新表述为：

$$E\left[g_h(Z, \boldsymbol{X}) \left[\frac{\Delta I(T^* \leq t)}{P(C > T^* \mid Z, \boldsymbol{X}, R=1)} - h(W, \boldsymbol{X}) \right] \mid R=1 \right] = 0 \tag{6}$$

对于任意函数 $g_h(Z, \boldsymbol{X})$，我们将 $h(W, \boldsymbol{X})$ 参数化，并选择 $g_h(Z, \boldsymbol{X})$ 作为 Z 和 $\boldsymbol{X}$ 的变换集合，使得（(6)）中的矩方程数量与 $h(W, \boldsymbol{X})$中的参数数量相匹配。这确保方程组恰好可识别，并且可以使用外部数据集求解。

估计 q 时也采用类似的策略。将（2）乘以任意函数$g_q(W, \boldsymbol{X})$，可得

$$E\left[g_q(W, \boldsymbol{X})\left[q(Z, \boldsymbol{X}) - \frac{P(R=0 \mid W, \boldsymbol{X})}{P(R=1 \mid W, \boldsymbol{X})}\right] \mid R=1\right] = 0,$$

其中，$P(R = 1 \mid W, X)$ 可通过逻辑回归估计。我们对 $q(Z, \boldsymbol{X})$ 进行参数化，并选择 $g_q(W, \boldsymbol{X})$ 作为 W 和 $\boldsymbol{X}$ 的一组变换，以求解 $q(Z, \boldsymbol{X})$ 中的参数。

一旦估计出 h 或 q，我们就应用定理 1中的识别结果，通过结局桥接形式（3）、治疗桥接形式（4）或增广（双重稳健）形式（5）（该形式结合了两个桥接函数）来估计 $P(T^*(0) \leq t|R = 0)$。需要注意的是，当事件发生率较低或代理变量较弱时，逆概率删失加权估计可能落在 (0,1) 范围之外。如果参数删失模型（例如指数分布）适用，则可通过非参数自助法或夹心方差估计量获得标准误和置信区间。为改善有限样本表现，我们应用 $\log(-\log(\cdot))$ 变换，在变换后的尺度上构造置信区间，再反变换回原始尺度。

补充材料S5提供了关于半参数方法的更多结果。我们表明，如果结局桥接函数和处理桥函数均被非参数地估计——并且治疗桥接函数是通过求解 $E\left[g_q(W, \boldsymbol{X})\left[I(R=1)q(Z, \boldsymbol{X}) - I(R=0)\right]\right]$ =0而非使用公式(2)来估计的——则由两个桥接函数得到的估计量是相同的。然而，当使用参数模型时，这种等价性通常并不成立。我们还讨论了当 Z 或 W 与 U无关时的影响，如$Z \perp\!\!\!\perp W \mid \boldsymbol{X}$所示。为防范这种情况，我们建议在进行主要分析之前，评估在给定 $\boldsymbol{X}$ 的条件下 Z 与 W 之间的条件关联。

2.2.2 方法2：基于回归的两阶段估计量

上述半参数估计量对底层数据生成过程的假设较少，但在人类免疫缺陷病毒预防试验中，事件发生率通常较低，因此效率更高的基于回归的方法可能更为实用。

Liu 等人此前提出了基于回归的近端因果推断方法，Liu 等人（2024年），用于广义线性模型；Li 等人（2025）则提出了适用于加性风险框架下右删失生存结局的方法。由于 Cox 比例风险模型（Cox 1972）在人类免疫缺陷病毒预防试验中更为常用，我们开发了一种基于 Cox 模型的新型基于回归的近端推断方法，专门针对低发病率结局情境。

对于基于回归的方法，我们引入了一项新假设，以反映预防试验中通常观察到的 HIV 发病率较低这一特点。这一罕见事件假设有助于解决 Cox 模型的不可折叠性问题。

假设5 (罕见事件**)**。 对于某个预先指定的时间点 t_0，安慰剂组的无事件概率满足 $P(T^*(0) > t_0|U, \boldsymbol{X}) \approx 1$，这反映了事件发生率较低。

在假设 5下，安慰剂组的风险函数可以用密度函数近似，而密度函数具有可折叠性（Tchetgen Tchetgen 等，2015）。具体而言，对于 $t \leq t_0$，$\lambda^{(0)}(t \mid U, \boldsymbol{X}) \approx f^{(0)}(t \mid U, \boldsymbol{X})$，其中 $\lambda^{(0)}(t \mid \cdot)$ 和 $f^{(0)}(t \mid \cdot)$ 分别表示安慰剂组的条件风险函数和密度函数。在我们的应用中，我们将 $t_0 = 1$ 年设为时间单位，这是 HIV 预防试验中自然的时间单位，其依据是安慰剂组的1年累积发生率通常较低。例如，AMP研究报告称，安慰剂组的一年累积发生率为2.98%，对应于无 HIV感染的一年概率为0.97。

接下来，我们为未测量混杂因素 U 指定一个分布假设。

假设 6 (未测量混杂因素**)**。 我们假设 $U = E(U|R, Z, \boldsymbol{X}) + \epsilon$，其中$\epsilon \perp\!\!\!\perp (R, Z, \boldsymbol{X})$、$E(\epsilon) = 0$，且 $E(U \mid R = 1, Z, \boldsymbol{X}) \neq E(U \mid R = 1, \boldsymbol{X})$ 几乎处处成立。

假设6 适用于 U 为连续变量的情形，因为这意味着误差结构均值独立。当 U 为二元变量时，该假设不成立。我们还基于一个替代模型，探索了针对二元 U 的假设修改版本，但发现由于涉及的参数数量较多，估计会变得不稳定。因此，应用两阶段方法时，应假定 U 为连续变量。

请注意，潜在的数据生成过程可能同时满足 IPCW 方法（其中 U为二元变量）和两阶段方法（其中U为连续变量）的识别假设；我们在补充材料S7中给出了一个示例。此外，我们假设在 $\boldsymbol{X}$ 条件下，负对照暴露 Z 与 U 相关，且这一关系适用于外部研究。值得注意的是，两阶段方法不需要完备性假设（假设3），而是隐含地以广义线性模型中的标准满秩条件取而代之，因此，即使 U 为连续变量，我们也可以使用二元 Z 和 W 来处理未测量混杂。

接下来，我们为二元负对照结局 W 指定一个对数线性模型，并为反事实事件时间 T^*(0) 指定一个 Cox 模型。

假设 7（W 和 $T^*(0)$ 的模型）。*(a)* 我们假设负对照结局W服从对数线性模型：$E(W|U, \boldsymbol{X}) = exp(\beta_{W0} + \beta_{Wx}\boldsymbol{X} + \beta_{Wu}U)$，其中 β_{W0}、β_{Wx} 和 β_{Wu}是模型参数，且 $\beta_{Wu} \neq 0$。*(b)* 我们假设反事实事件时间 $T^*(0)$ 服从 *Cox* 模型：$\lambda^{(0)}(t \mid U, \boldsymbol{X}) = \lambda_0(t)\, exp(\beta_{Tu}U + \beta_{Tx}\boldsymbol{X})$，其中 $\lambda_0(t)$ 是基线风险函数，β_{Tu} 和 β_{Tx} 是模型参数。

最后，我们还对负对照施加假设 1，并对外部数据中的随机删失施加假设 4。定理 2 随后给出了基于回归的方法的主要识别结果。

定理 2. 在假设 *1* 和 *4-7*下，对于任意 $t \leq t_0$，我们有

$$\lambda^{(0)}(t \mid R, Z, \boldsymbol{X}) \approx \tilde{\lambda}_0(t) \exp\left[\left\{\tilde{\beta}_{Tx}\boldsymbol{X} + \tilde{\beta}_{Tw} \log\left(E(W \mid R, Z, \boldsymbol{X})\right)\right\}\right], \tag{7}$$

其中 $\tilde{\lambda}_0(t)$、$\tilde{\beta}_{Tx}$ 和 $\tilde{\beta}_{Tw}$ 是可利用外部研究（$R = 1$）的数据识别的参数。因此，主要研究的反事实风险函数可识别为 $\lambda^{(0)}(t|R = 0, Z, \boldsymbol{X}) = \tilde{\lambda}_0(t) \exp\left[\left\{\tilde{\beta}_{Tx}\boldsymbol{X} + \tilde{\beta}_{Tw} \log\left(E(W \mid R = 0, Z, \boldsymbol{X})\right)\right\}\right]$。

我们方法背后的直觉可通过以下推导来说明。令 $t \leq t_0$。在假设 1、5 和 7下，我们有

$$f^{(0)}(t \mid R, Z, U, \boldsymbol{X}) \approx \lambda^{(0)}(t \mid R, Z, U, \boldsymbol{X}) = \lambda_0(t) \exp(\beta_{Tu}U + \beta_{Tx}\boldsymbol{X}),$$

$$E(W \mid R, Z, U, \boldsymbol{X}) = \exp(\beta_{W0} + \beta_{Wx}\boldsymbol{X} + \beta_{Wu}U).$$

接下来，对给定 U条件下的分布进行边际化$(R,\ Z,\ \boldsymbol{X})$，并应用假设 6，得到:

$$f^{(0)}(t \mid R, Z, \boldsymbol{X}) = E\{\lambda_0(t)\exp(\beta_{Tu}\epsilon)\}\exp(\beta_{Tx}\boldsymbol{X} + \beta_{Tu}E(U \mid R, Z, \boldsymbol{X}))$$

$$E(W \mid R, Z, \boldsymbol{X}) = E\{\exp(\beta_{W0} + \beta_{Wu}\epsilon)\}\exp(\beta_{Wx}\boldsymbol{X} + \beta_{Wu}E(U \mid R, Z, \boldsymbol{X}))$$

因此，存在 β_{T0}、β_{Tu}、β_{Tx}，它们在主要数据集和外部数据集中是共享的，使得 $\lambda^{(0)}(t \mid R, Z, \boldsymbol{X}) \approx f^{(0)}(t \mid R, Z, \boldsymbol{X}) = \tilde{\lambda}_0(t)\exp\left\{\tilde{\beta}_{Tx}\boldsymbol{X} + \tilde{\beta}_{Tw}\log(E(W|R, Z, \boldsymbol{X}))\right\}$，其中近似是通过 $P(T^*(0) > t_0|R, Z, \boldsymbol{X}) = E\{P(T^*(0) > t_0|U, \boldsymbol{X}) \mid R, Z, \boldsymbol{X}\} \approx 1$得到的。至此，定理 2 的论证完成。

在实施时，我们首先使用主要研究和外部研究的数据估计 $E(W \mid R, Z, \boldsymbol{X})$。由于只有外部数据集包含安慰剂条件下的观测，且模型系数在各数据集中共享，因此我们使用外部数据拟合方程 (7) 中的 Cox 模型。具体而言，我们在 Cox 回归中将观测到的事件时间对 $\boldsymbol{X}$ 和 $\log\ E(W \mid R = 1, Z, \boldsymbol{X})$进行回归，并将其作为回归变量。由此得到参数 $\lambda_0(t)$ β_{Tx}, β_{Tw}的估计值。这些参数是可识别的，因为在假设条件下，$E(W \mid R = 1, Z, \boldsymbol{X})$随 Z变化，从而为估计提供了足够的变异性。参数估计完成后，我们可以根据定理 2 计算 $\lambda^{(0)}(t \mid R = 0, Z, \boldsymbol{X})$。进而得到 $P(T^*(0) \leq t|R = 0, Z, \boldsymbol{X})$，从而得到 $P(T^*(0) \leq t|R = 0) = E\{P(T^*(0) \leq t|R = 0, Z, \boldsymbol{X}) \mid R = 0\}$。当假设 $\lambda_0(t)$不随时间变化时，可以将所有模型参数——包括关注参数和干扰参数——合并起来，并通过估计方程来表述其估计过程。

进行统计推断时，可以使用非参数自助法。或者，在假设$\lambda_0(t)$不随时间变化时，可以使用夹心方差估计量。同样，使用 $\log(-\log(\cdot))$ 对感兴趣的参数进行变换，以提高渐近近似的准确性。在我们的实证结果中，当事件较为罕见时，夹心方法得出的方差估计值比自助法更精确。

我们通过补充材料中展示的全面模拟研究，评估所有拟议近端推断估计量的有限样本性能S7。

3 案例研究：在 HIV 预防试验中的应用

3.1 估计安慰剂下的反事实HIV发病率

我们将该方法应用于 HPTN 083，使用 AMP 研究的安慰剂组作为外部对照数据，估计 HPTN 083 中安慰剂下的反事实一年累积HIV发生率。鉴于 AMP 研究三个组的 HIV 发病率相近，我们还在补充部分 S1.1.

调整集中的协变量包括年龄（18–20岁、21–30岁）、性别（男性、其他）和种族（黑人、白人、其他）。请注意，Cox 模型仅显示年龄和种族与结局之间存在显著关联，但由于文献表明性别与HIV感染相关，我们在本主要分析中也对性别进行了调整。地理区域被用作NCE，定义为拉丁美洲（$Z = 0$）与非拉丁美洲（$Z = 1$）。根据实验室检测确定的基线性传播感染（STI）状态被用作NCO，其中 $W = 1$ 表示STI检测结果呈阳性（表1）。我们评估了三种NCO选项：（1）基线直肠淋病感染；（2）基线直肠衣原体感染；（3）基线直肠淋病或衣原体感染。为评估NCO的有效性，我们拟合了以 Z 和 $\boldsymbol{X}$为自变量、以 W 为因变量的逻辑回归。使用基线直肠衣原体作为NCO时，在给定 $\boldsymbol{X}$ 的条件下，W 与 Z 之间的关联无统计学显著性（OR $= 0.846$，$p = 0.068$），这表明它并非未测量混杂因素的合适代理变量。基线直肠淋病和综合STI指标与地理区域在给定 $\boldsymbol{X}$ 的条件下均存在统计学显著关联（OR$= 0.605$，$p < 0.001$ 和OR $= 0.740$，$p < 0.001$）。然而，鉴于两个数据集中基线淋病感染的患病率均较低（HPTN 083数据集中为6.5%，AMP中为3.7%），我们将综合STI指标作为NCO用于主要分析。各STI组成部分的结果见补充材料S1.2。

我们采用两种拟议方法估计安慰剂下的反事实HIV发病率。第一种是半参数逆概率删失加权方法（2.2.1节），其中使用线性结局桥接函数 $h(W, \boldsymbol{X}) = \beta_{h0} + \beta_{hW} W + \beta_{hX} \boldsymbol{X}$ 和线性治疗桥接函数 $q(Z, \boldsymbol{X}) = \beta_{q0} + \beta_{qZ} Z + \beta_{qX} \boldsymbol{X}$。随后，通过结合这两个桥接函数构建了双重稳健估计量。第二种是基于回归的两阶段方法（2.2.2节），该方法假设基线风险随时间不变，并应用于完整随访数据以及截断数据集；在该数据集中，所有参与者均在一年时接受行政删失，对应定理 2中的 $t_0 = 1$ 年，以满足假设5。具体而言，我们假设外部数据（$R = 1$）服从指数分布，并据此估计生存函数；随后使用拟合函数预测主要HPTN 083数据集中的生存概率。作为比较，我们报告了朴素回归模型的结果，这些模型仅调整 $\boldsymbol{X}$ ，或同时调整 $\boldsymbol{X}$、 Z和 W 。所有方法的估计值和置信区间均通过R中的 `gmm` 软件包获得。为改善渐近近似，我们对目标参数进行了变换，使用log(–log(·))。

在我们的主要分析中，NCO值缺失的参与者（每个NCO的缺失比例均低于10%）被排除，如图1所示，结果见图3。我们还在补充材料S1.3中探索了其他协变量调整集：（1）年龄和种族；（2）年龄、种族、性别、族裔和教育水平（大专及以上 vs. 高中及以下）。调整较少协变量得到相似的估计值，而调整更多协变量则因低事件率而产生更大的估计值和更宽的置信区间。

在所有近端因果推断方法中，安慰剂下的反事实一年累积HIV发生率是一致的。三种半参数逆概率删失加权估计量（结局桥接函数、治疗桥接函数和双重稳健）以及两种两阶段回归方法给出的点估计范围为每100人年4.3至5.5。正如预期，应用于完整随访数据的两阶段回归方法比一年分析产生更窄的置信区间，这是因为观察到的事件更多。

相比之下，两个朴素回归模型得出的反事实估计值较低

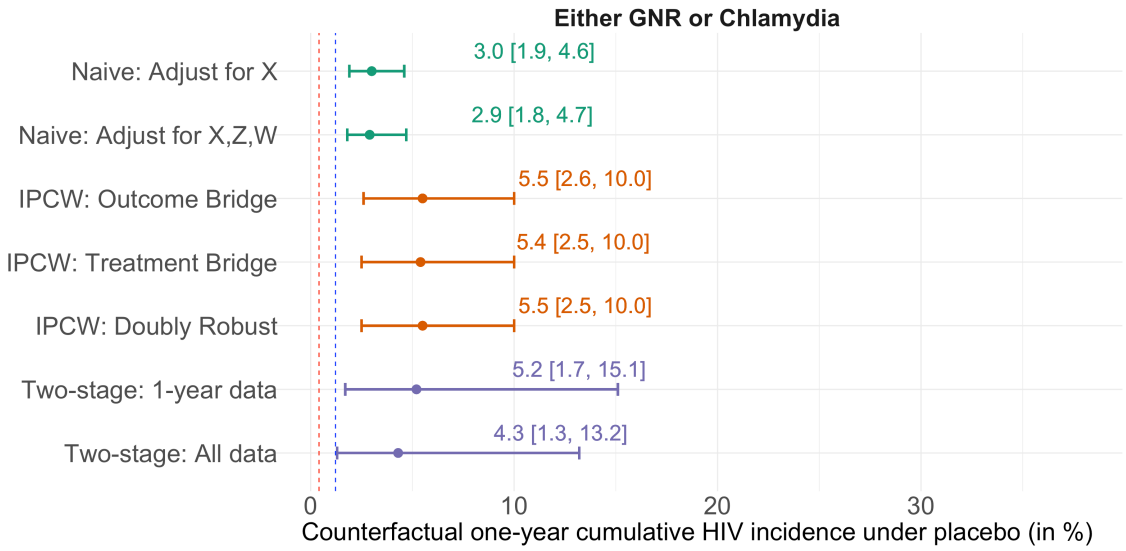


图3：采用不同方法，并以基线淋病或衣原体感染作为NCO时，HPTN 083中安慰剂组一年反事实HIV累积发病率的估计值（%）。红色虚线表示卡博特韦组观察到的一年累积发病率（0.41%），蓝色虚线表示TDF/FTC组观察到的一年累积发病率（1.22%）。

安慰剂组的HIV累积发病率，这很可能反映了未测量混杂造成的残余偏倚。仅针对 $\boldsymbol{X}$ 调整的朴素模型，以及同时针对 $\boldsymbol{X}$ 和负对照变量 Z 、 W调整的朴素模型，分别得出了每100人年3.0和2.9的点估计。此外，朴素回归模型得出的置信区间较窄，因为它们假设不存在未测量混杂，并且基于外部对照数据中观测协变量的结局模型可以直接应用于主要数据集。相比之下，近端学习认识到，由于未测量混杂，这些结局模型在不同数据集之间可能不兼容，因此使用负对照来处理这一问题；这会引入额外的不确定性，从而导致置信区间更宽。

3.2 绝对疗效的统计推断

接下来，我们使用近端因果推断方法估计的安慰剂组反事实一年HIV累积发病率，评估HPTN 083中两种处理的绝对疗效。作为比较，HPTN 083中观察到的卡博特韦组一年HIV累积发病率为0.41%（95%置信区间： $[0.2, 0.7]$），TDF/FTC组为1.22%（95%置信区间： $[0.9, 1.7]$）。

为评估统计显著性，我们计算了 Wald 检验统计量，以确定各活性治疗组与估计的安慰剂组之间的一年期 HIV累积发病率是否存在显著差异。所有方法的 Wald 检验统计量均按以下方式计算：

$$T_{\mathrm{CAB}} = \frac{\hat{\beta}_{\mathrm{CAB.log}} - \hat{\beta}_{\mathrm{placebo.log}}}{SE(\hat{\beta}_{\mathrm{CAB.log}} - \hat{\beta}_{\mathrm{placebo.log}})}, \qquad T_{\mathrm{TDF}} = \frac{\hat{\beta}_{\mathrm{TDF.log}} - \hat{\beta}_{\mathrm{placebo.log}}}{SE(\hat{\beta}_{\mathrm{TDF.log}} - \hat{\beta}_{\mathrm{placebo.log}})}.$$

在上述公式中， $\beta_{*.\mathrm{log}} = \log(-\log(\beta_*))$ 表示 HPTN 083 中各治疗组（$* =$ TDF 或 CAB）一年累积发生率经互补对数-对数变换后的估计值， $\beta_{\mathrm{placebo.log}} = \log(-\log(\beta_{\mathrm{placebo}}))$ 表示通过逆概率删失加权或两阶段回归方法得到的安慰剂反事实估计值。请注意， $\beta_{*.\mathrm{log}}$ 和 $\beta_{\mathrm{placebo.log}}$ 相关，因为 $\beta_{\mathrm{placebo.log}}$ 是利用两个治疗组的参与者估计的。为考虑这种相关性，我们直接使用逆概率删失加权和两阶段

基于回归的方法。

结果见表2。我们的结果表明，在所有近端推断方法中，卡博特韦相对于反事实安慰剂组均具有统计学显著的优效性（所有p值 ≤ 0.001）。所有方法也均显示，TDF/FTC相对于安慰剂具有统计学显著的优效性，估计相对疗效为71.6%–77.8%，但p值较大。这一模式符合预期，因为AMP研究安慰剂组仍保持了不可忽视的暴露前预防用药（如第1.3节所述，AMP安慰剂组中39.0%的样本可检测到TDF/FTC浓度，而HPTN 083 TDF/FTC组中这一比例为72.3%）。这些发现为卡博特韦的疗效提供了有力证据，也为与反事实安慰剂条件相比的TDF/FTC疗效提供了中等程度的证据。尽管存在局限性，结果表明，所提出的方法可作为一种可行的分析框架，用于评估未来没有同期安慰剂对照组的临床试验中的安慰剂对照疗效，尤其适用于疗效与卡博特韦相当的研究性药物。

表2：估计相对有效性（1减去活性治疗组相对于安慰剂的一年累积HIV感染发生率之比）、估计绝对疗效（一年HIV累积发病率之差）、Wald检验统计量，以及比较活性治疗组与估计安慰剂组的p值。针对不同分析方法，呈现了卡博特韦与安慰剂以及TDF/FTC与安慰剂的统计比较。

NCO	方法	设定	卡博特韦				TDF/FTC			
			相对疗效	绝对疗效	检验统计量	p值	相对疗效	绝对疗效	检验统计量	p值
	朴素	针对 X 调整后	0.863	0.026	-5.593	< 0.001	0.593	0.017	-3.111	0.002
任一种性传播感染	朴素	针对 X，Z，W 调整后	0.859	0.025	-5.205	< 0.001	0.579	0.017	-2.933	0.003
	IPCW	结局桥接函数	0.925	0.051	-4.960	< 0.001	0.778	0.043	-3.360	0.001
	IPCW	治疗桥接函数	0.924	0.050	-4.870	< 0.001	0.774	0.042	-3.284	0.001
	IPCW	双重稳健	0.925	0.051	-4.945	< 0.001	0.778	0.043	-3.346	0.001
	两阶段	1年数据	0.921	0.048	-3.692	< 0.001	0.765	0.040	-2.252	0.023
	两阶段	所有数据	0.905	0.039	-3.731	< 0.001	0.716	0.031	-2.202	0.028

逆概率删失加权 = 逆概率删失加权； X = 基线协变量； Z = 负对照暴露； W = NCO。

4 讨论

我们提出两种近端推断方法，利用存在未测量混杂的外部对照数据，估计安慰剂条件下的反事实HIV发病率。这些方法旨在应对HIV预防试验中的关键方法学挑战，包括右删失和低事件率。具体而言，我们（1）扩展了Cui 等人（2024年）和Ying 等人（2022年）的半参数框架，以适应右删失并整合外部对照，并且（2）开发了一种针对低发病率结局的新型两阶段回归方法。我们的方法在HIV预防试验之外具有广泛适用性。两种方法均适用于低发病率情形，而逆删失概率加权法可扩展至中至高发病率结局。表3 比较了半参数逆概率删失加权方法与两阶段回归方法，包括其假设、优势和局限性。半参数逆概率删失加权方法所需的建模假设更少，且不要求 Z 存在于主要数据集中，但估计值和置信区间可能不在（0，1）的支撑内。相比之下，两阶段回归方法假设稀有事件设定，且不易推广至离散的未测量混杂因素。在实际应用中，必须审慎考虑这些方法学差异。

我们将这些方法应用于 HPTN 083试验，以估计安慰剂条件下反事实一年HIV累积发病率。在不同的分析设定下，半参数逆概率删失加权和两阶段回归方法得出的估计值均一致，范围为 4.3% 至 5.5%。所有近端推断方法均显示，观察到的卡博特韦发病率和TDF/FTC发病率与估计的安慰剂发病率之间存在显著差异。这些方法表明，在没有同期安慰剂组的临床试验中评估绝对疗效是可行的，尤其适用于疗效与卡博特韦相当的药物。

识别可作为未测量混杂因素强代理变量的有效阴性对照 Z 和 W，对于在近端

学习框架中识别目标因果参数至关重要，并直接影响估计值的精确度。在我们的 HIV 预防试验应用中，我们将基线直肠淋病或衣原体感染作为 NCO，并将地理区域作为负对照暴露，以处理一个重要的未测量混杂因素：当地 HIV 传播环境风险。实践中，可能需要考虑多个未测量混杂因素，因此需要利用近端学习方法针对每个混杂因素识别有效的负对照暴露与 NCO 配对。未来研究应在范围更广的 HIV 预防试验中应用这些方法，以评估其在不同环境下的稳健性，并考察最优的负对照选择策略，从而获得可靠且高效的反事实 HIV 发病率估计值。

使用外部对照方法（包括我们提出的方法）估计的绝对疗效，应结合外部数据集中的具体安慰剂条件来解读。例如，我们的安慰剂条件应相对于 AMP研究安慰剂组中的暴露前预防用药情况来解读；这代表了一种具有临床相关性的安慰剂定义。未来可以开发统计方法，利用 AMP 安慰剂组中暴露前预防用药模式的差异，估计相对于不使用暴露前预防用药的“纯”安慰剂的绝对疗效。

表 3：所提出的半参数逆概率删失加权与基于两阶段回归方法的比较

方法

	半参数逆概率删失加权	基于两阶段回归的方法
假设	A1. 未测量混杂因素和有效阴性对照 A2. 正值性 A3. 完备性 A4. 外部研究中的随机删失项	A1*. 连续 未测量 con-创始人与有效阴性对照 A4. 外部研究中的随机删失-外部研究 A5. 低事件发生率（罕见事件） A6. 未测量 混杂因素 均值-独立误差结构 A7. NCO 的对数线性模型和 Cox 事件时间模型
优点	• 所需的建模假设更少 • 可根据未测量混杂因素的类型进行处理，具体取决于负对照变量的类型 • 负对照暴露不必存在于主要数据集中	•通常可得到更精确的估计—计值具有更窄的置信区-间，在存在罕见事件的情况下
局限性	• 对罕见结局的效率较低，且置信区间更宽 •估计值和置信区间可能不在（0，1）内	• 不易推广到离散未测量混杂因素 • 需要更强的建模假设 • 假设处于稀有事件设定

参考文献

Belloni, A., Chernozhukov, V., Fernández-Val, I. 与 Hansen, C. (2017年), ‘高维数据下的项目评估与因果推断’, *Econometrica* **85**(1), 233–298.

Bennett, A. 等 (2022年), ‘弱识别函数的强识别泛函推断’,*arXiv*预印本 *arXiv:2208.08291* .

Bennett, A. 等 (2023年), ‘逆问题泛函的源条件双稳健推断’,*arXiv*预印本 *arXiv:2307.13793* .

Chen, J. W. 等 (2022年), ‘医学研究中倾向得分方法的最佳实践指南：对理论、实施与报告的考量。综述’,《关节镜》：关节镜及相关手术杂志**38**(2), 632–642.

Chernozhukov, V. 等 (2018年), ‘治疗与结构参数的双重/去偏机器学习’, 《计量经济学杂志》 **21**(1), C1–C68.

Chesnaye, N. C. 等（2021）, “观察性研究中治疗逆概率加权简介”, 《临床肾脏病学杂志》**15**(1), 14–20.

Cox, D. R.（1972）, “回归模型与生命表”, 《皇家统计学会杂志：*B*辑（方法学）》 **34**(2), 187–202.

Cui, Y. 等（2024）, “半参数近端因果推断”, 《*Journal of the American Statistical Association*》**119**(546), 1348–1359.

Dehejia, R. H. & Wahba, S.（2002）, “非实验性因果研究中的倾向评分匹配方法”, 《经济学与统计学评论》**84**(1), 151–161.

Dias, S., Sutton, A. J., Ades, A. E. & Welton, N. J. (2013年), ‘用于决策的证据综合 2：随机对照试验成对和网络荟萃分析的广义线性建模框架’, 《医学决策》**33**(5),607–617. PMID: 23104435.

Dikkala, N., Lewis, G., Mackey, L. & Syrgkanis, V. (2020), ‘条件矩模型的极小极大估计’, 神经信息处理系统进展 **33**,12248– 12262.

Donnell, D. 等（2023），《使用外部试验数据估计新型暴露前预防药物相对于安慰剂的反事实疗效：注射用卡博特韦与女性口服暴露前预防的案例》，《国际艾滋病学会杂志》**26**(6), e26118。

Donnell, D. 等（2024），《未来人类免疫缺陷病毒预防主动对照临床试验的研究设计方法》，《传染病统计通讯》**15**(1), 20230002。

Dunn, D. T., Glidden, D. V., Stirrup, O. T. & McCormack, S.（2018），《避免感染比率：实验性人类免疫缺陷病毒暴露前预防药物有效性的新指标》，《柳叶刀·艾滋病病毒》**5**(6), e329–e334。

Gao, F., Janes, H., Buchbinder, S. & Donnell, D.（2025），《采用反事实安慰剂的人类免疫缺陷病毒预防临床试验主动对照试验设计》，《医学统计学》**44**(6), e70022。

Ghassami, A., Ying, A., Shpitser, I. & Tchetgen Tchetgen, E. (2022年), 面向一类双重稳健泛函的极小极大核机器学习及其在近端因果推断中的应用，载于《第25届国际人工智能与统计会议论文集》，《机器学习研究论文集》第151卷，PMLR，第7210–7239页。

Glidden, D. V. (2019), ‘推进新型暴露前预防产品——非劣效性的备择方案’，《传染病统计通讯》**11**(1), 20190011。

Glidden, D. V., Stirrup, O. T. & Dunn, D. T. (2020), ‘针对人类免疫缺陷病毒感染数量较少的暴露前预防试验的贝叶斯避免感染框架：DISCOVER试验结果的应用’，《柳叶刀·*HIV*》 **7**(11), e791–e796。

Hernán, M. A. 与 Robins, J. M.（2020），《因果推断：如果……会怎样》，Chapman & Hall/CRC。可于以下网址获取：https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/。

Hobbs, B. P.、Sargent, D. J. 与 Carlin, B. P.（2012），‘在临床试验中使用一般线性模型与广义线性模型纳入历史信息的相称先验’，贝叶斯分析 **7**(3),639–674。

Ibrahim, J. G. & Chen, M.-H. (2000), ‘回归模型的幂先验分布’,*Statistical Science***15**(1), 46–60.

Imbens, G. W. (2003), ‘项目评估中对外生性假设的敏感性’, 《美国经济评论》**93**, 126–132.

Jiang, L., Nie, L. & Yuan, Y. (2023年), ‘在临床试验中利用弹性先验动态借用历史数据中的信息’, *Biometrics* **79**(1),49–60.

Joffe, M. M., Have, T.R. T., Feldman, H. I.& Kimmel, S. E. (2004), ‘模型选择、混杂因素控制与边际结构模型’, 《美国统计学家》**58**(4), 272– 279.

联合国艾滋病规划署 (UNAIDS) & 世界卫生组织 (WHO) (2021), HIV预防试验中的伦理考量，指导文件 UN- AIDS/JC3010E，联合国艾滋病规划署 (UNAIDS) 与世界卫生组织 (WHO)，日内瓦。另有 WHO 电子版：ISBN 978-92-4-002094-8。许可：CC BY-NC-SA 3.0 IGO。

Kansiime, S. 等（2025），《基于登记队列估计反事实安慰剂 HIV发病率的挑战：PrEPVacc试验》，临床试验**22**(3)，289–300。在线发表于 2024-12-31。

Kress, R.（1989），线性积分方程，第 82 卷，Springer。

Landovitz, R. J. 等（2021），《顺性别男性和跨性别女性的卡博特韦用于HIV预防》，《新英格兰医学杂志》 **385**(7), 595–608.

Li, K. Q., Linderman, G. C., Shi, X. & TchetgenTchetgen, E. J.（2025），《基于回归的近端因果推断：用于右删失生存时间数据》，《*Epidemiology*》**36**(5), 694–704.

Li, M.（2013年），‘使用倾向评分法估计因果效应：综述与实践指南’，《*Organizational Research Methods*》**16**(2)，188–226。

Li, X., Miao, W., Lu, F. & Zhou, X.-H.（2023年），‘提高使用外部对照数据的临床试验中推断的效率’，*Biometrics* **79**(1)，394–403。

Li, Y. P., Propert, K. J. & Rosenbaum, P. R.（2001年），‘平衡风险集匹配’，《*Journal of the American Statistical Association*》**96**(455)，870–882。

Liu, J., Park, C., Li, K. & Tchetgen Tchetgen, E. J.（2024年），‘基于回归的近端因果推断’，《*American Journal of Epidemiology*》。提前发表文章。

Liu, Y. 等（2022年），‘使用真实世界数据为随机对照临床试验的对照组扩充样本的匹配设计’，《*Journal of Biopharmaceutical Statistics*》**32**(1)，124–140。

Liu, Y. 等（2025），《分布偏移下因果生存分析的靶向数据融合》，*arXiv* 预印本 *arXiv:2501.18798* 。

Mayer, K. H. 等（2020），《恩曲他滨和丙酚替诺福韦与恩曲他滨和富马酸替诺福韦二吡呋酯用于人类免疫缺陷病毒暴露前预防（DISCOVER）：一项随机、双盲、多中心、主动对照、3期非劣效性临床试验的主要结果》，《柳叶刀》**396**(10246)，239–254。

Miao, W.、Geng, Z. 和 TchetgenTchetgen, E.（2018），《利用未测量混杂因素的代理变量识别因果效应》，*Biometrika* **105**(4)，987–993。

Mullick, C. & Murray, J. (2020), ‘男男性行为者中人类免疫缺陷病毒（HIV）感染与直肠淋病发病率之间的相关性：对未来 HIV暴露前预防临床试验的启示——

示’，《传染病杂志》**221**(2), 214–217.

Olivas-Martinez, A., Gilbert, P. B. & Rotnitzky, A. (2025), ‘针对修正治疗策略的近端因果推断’, *arXiv* 预印本 *arXiv:2512.12038* 。

Pocock, S. (1976), ‘临床试验中随机对照与历史对照的结合’,《慢性病杂志》**29**(3), 175–188。

PrEPVacc (2024年), ‘ClinicalTrials.gov：在非洲开展的一项疗效研究，评估两种DNA-MVA-Env蛋白或DNA-Env蛋白人类免疫缺陷病毒-1疫苗方案与暴露前预防的组合（PrEPVacc）’,https://classic.clinicaltrials.gov/ct2/show/NCT04066881。 [访问日期: 6月21日]。

Prudden, H. J. et al. (2024年), ‘关于HIV预防效力试验设计方法的观点’, 《艾滋病研究与人类逆转录病毒》 **40**(5), 301–307。PMID: 37392020。

PURPOSE1（2024年），《ClinicalTrials.gov：针对有HIV感染风险的青春期女孩和年轻女性开展的来那帕韦与恩曲他滨/丙酚替诺福韦暴露前预防研究（PURPOSE1）》，https://clinicaltrials.gov/study/NCT04994509。访问日期：2025- 05-08。

PURPOSE2（2024年），《ClinicalTrials.gov：针对有HIV感染风险人群开展的来那帕韦暴露前预防研究（PURPOSE2）》，https://clinicaltrials.gov/study/NCT04925752。访问日期：2025-05-08。

Robins, J. M.（1986），《在持续暴露期死亡率研究中进行因果推断的新方法——应用于控制健康工人幸存者效应》，《数学建模》**7**（9-12），1393–1512。

Rosenbaum, P. R. & Rubin, D. B. (1983*a*)，《评估观察性研究中未观测二元协变量对二元结局的敏感性》，*Journal of Royal StatisticalSociety, Series B* **45**，212–218。

Rosenbaum, P. R. & Rubin, D. B. (1983*b*), 《倾向评分在因果效应观察性研究中的核心作用》, *Biometrika* **70**(1), 41–55。

Rubin, D. B. (1974), '估计随机与非随机研究中处理的因果效应。', 《教育心理学杂志》 **66**(5),688.

Rubin, D. B. (1980), '实验数据的随机化分析：关于 Fisher 随机化检验的评论', 《*Journal of the American Statistical Association*》 **75**, 591–593.

Schmidli, H., Neuenschwander, B. & Friede, T. (2017年), '在临床试验的设计与分析中使用历史方差数据进行元分析预测', 《计算统计与数据分析》 **113**,100–110.

Shi, X., Miao, W., Nelson, J.C. &TchetgenTchetgen, E.J. (2020), '利用双阴性对照调整对分类变量未测量混杂因素进行多重稳健因果推断', 英国皇家统计学会会刊*B*辑 **82**(2), 521–540.

Shi, X., Miao, W. & Tchetgen Tchetgen, E. (2020), '流行病学中阴性对照方法的选择性综述', 《当代流行病学报告》**7**(4), 190–202.

Snowden, J. M., Rose, S. & Mortimer, K. M. (2011), '在模拟数据集上实施 G计算：因果推断技术示例', 《*American Journal of Epidemiology*》**173**(7), 731–738.

Stuart, E. A. & Rubin, D. B. (2008), '针对组间差异进行调整的多个对照组匹配', 《教育与行为统计学杂志》 **33**(3), 279–306.

Su, Z., Rytgaard, H. C., Ravn, H. & Eriksson, F. (2025), '近端间接比较',*Biometrika* p.asaf044.

Tchetgen Tchetgen, E. J. 等（2015）, "生存情境中的工具变量估计", 《*Epidemiology*》 **26**(3), 402–410。

Tchetgen Tchetgen, E. J. 等（2024年），“近端因果推断简介”，*Statistical Science***39**(3)，375–390。

Temple, R. 和 Ellenberg, S. S.（2000），《新处理评价中的安慰剂对照试验与阳性对照试验。第1部分：伦理与科学问题》，*Annals of Internal Medicine* **133**(6)，455–463。

《柳叶刀·HIV》（2020），《人类免疫缺陷病毒中的种族不平等》，《柳叶刀·*HIV*》**7**(7)，e449。PMID: 32621870。

Thorlund, K.、Dron, L.、Park, J. J. H. 和 Mills, E. J.（2020），《临床试验中的合成对照与外部对照：研究者入门》，*Clinical Epidemiology* **12**，457–467。

美国食品药品监督管理局（2023年），《药品和生物制品外部对照临床试验的设计与实施考量：行业指南》。马里兰州银泉市：药品审评与研究中心、生物制品审评与研究中心、肿瘤学卓越中心。

Valancius, M. 等（2024年），《在混合临床试验中利用外部对照的因果推断框架》，*Biometrics* **80**(4)，ujae095。

Viele, K. 等（2014），《使用历史对照数据评估临床试验中的治疗效果》，*Pharmaceutical statistics* **13**(1)，41–54。

Ward, H. 与 Rönn, M.（2010），《性传播感染对人类免疫缺陷病毒性传播的贡献》，*Current Opinion in HIV and AIDS* **5**(4)，305–310。

Yi, Y.、Zhang, Y.、Du, Y. 与 Ye, T.（2023年），《在临床试验中利用内部和外部对照两次检验治疗效果》，*Journal of Causal Inference* **11**(1)，20220018。

Ying, A., Cui, Y. & Tchetgen Tchetgen, E. J.（2022年），‘边际反事实生存曲线的近端因果推断’，*arXiv*预印本 *arxiv:2204.13144* 。

Zhu, Y. 等（2024年），‘基于HIV暴露标志物估计无安慰剂组的HIV预防试验中的反事实安慰剂HIV发病率’，临床试验**21**(1)，114– 123。PMID:37877356。

补充材料

S1 补充分析

S1.1 将 AMP 的所有组合作为外部数据集

如第1.3节所述，AMP 中截至 1 年的 HIV 累积发病率（%）分别为低剂量组 2.5（95% CI：$[1.71, 3.53]$）、高剂量组 2.2（95% CI：$[1.46, 3.18]$）和安慰剂组 2.98（95% CI：$[2.11, 4.09]$）。该临床试验未显示 VRC01 组与安慰剂组之间的 HIV 发病率存在统计学显著差异。因此，作为补充分析，我们将 AMP 的所有组合作为外部数据集，并考察将直肠淋病、直肠衣原体感染以及任一种性传播感染作为 NCO 的情况。为评估各 NCO 的有效性，我们拟合了 W 对 Z 和 X 的逻辑回归。将直肠淋病、衣原体感染或任一种性传播感染作为 NCO 时，我们发现，在调整 Z 后，W 与 $\boldsymbol{X}$ 之间存在统计学显著关联（$p<0.001$，$p=0.023$ 和 $p<0.001$，依次对应），支持将其用作更合适的代理变量。所有结果见图S1 和表S1。

我们发现，各方法所估计的1年反事实 HIV 累积发生率均较低。这符合预期，因为尽管与安慰剂组相比无统计学显著性，AMP中低剂量组和高剂量组的发生率仍然较低。由于合并所有组后样本量更大且事件数更多，置信区间也更小。然而，如前所述，反事实 HIV 累积发生率需依据外部数据中“安慰剂”的定义加以解释。由于在本分析中，所有组（包括安慰剂组和活性治疗组）均伴有混合的暴露前预防用药，并被合并作为外部数据，结果可能更难解释。

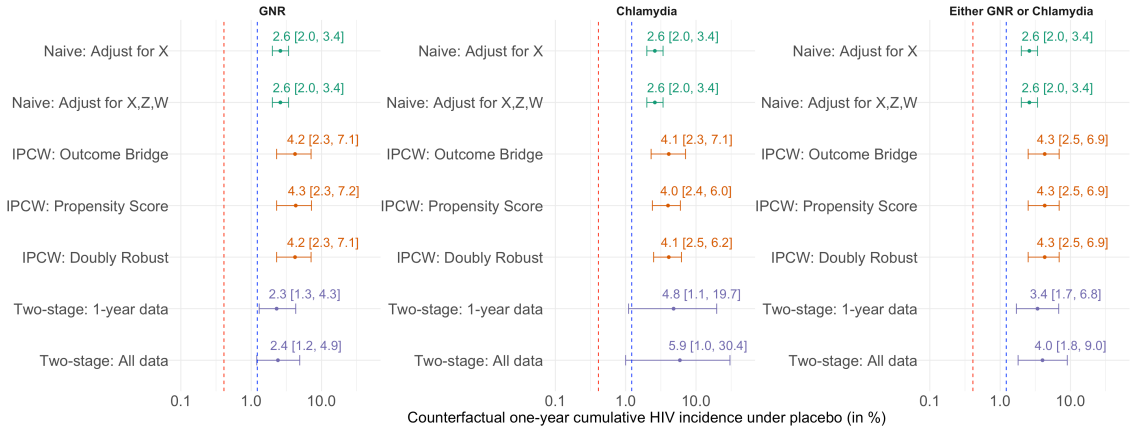


图 S1：使用不同方法和负对照结局估计的HPTN 083 1年反事实 HIV 累积发生率（%）。AMP数据中的所有组被合并为外部对照数据集。红色虚线表示卡博特韦组观察到的发生率（每100人年0.41），蓝色虚线表示TDF/FTC组观察到的发生率（每100人年1.22）。在所有模型中均对**年龄、种族和性别** 进行了调整。AMP数据中的**所有组** 被合并为外部数据。为便于可视化，x轴采用log()变换绘制。

表 S1：比较活性治疗组与估计的安慰剂组 1 年内 HIV 累积发病率的 Wald 检验统计量。按不同分析方法，分别比较卡博特韦与安慰剂，以及恩曲他滨/富马酸替诺福韦二吡呋酯（TDF/FTC）与安慰剂。所有模型均调整了**年龄、种族和性别**。AMP数据中的**所有组** 均合并为外部数据。

NCO	方法	设定	卡博特韦				TDF/FTC			
			相对疗效	绝对疗效	检验统计量	p值	相对疗效	绝对疗效	检验统计量	p值
	朴素	调整 X	0.842	0.022	-6.193	< 0.001	0.531	0.014	-3.885	< 0.001
GNR	朴素	调整 X，Z，W	0.842	0.022	-6.013	< 0.001	0.531	0.014	-3.746	< 0.001
	IPCW	结局桥接函数	0.902	0.038	-5.336	< 0.001	0.710	0.030	-3.383	0.001
	IPCW	倾向评分	0.905	0.039	-5.321	< 0.001	0.716	0.031	-3.393	0.001
	IPCW	双重稳健	0.902	0.038	-5.340	< 0.001	0.710	0.030	-3.388	0.001
	两阶段	1年	0.822	0.019	-4.847	< 0.001	0.470	0.011	-2.898	0.004
	两阶段	所有数据	0.829	0.020	-4.992	< 0.001	0.492	0.012	-2.957	0.003
衣原体	朴素	调整 X，Z，W	0.842	0.022	-6.034	< 0.001	0.531	0.014	-3.774	< 0.001
	IPCW	结局桥接函数	0.900	0.037	-5.336	< 0.001	0.702	0.029	-3.383	0.001
	IPCW	倾向评分	0.898	0.036	-6.127	< 0.001	0.695	0.028	-3.896	< 0.001
	IPCW	双重稳健	0.900	0.037	-6.179	< 0.001	0.702	0.029	-3.971	< 0.001
	两阶段	1年	0.915	0.044	-1.990	0.047	0.746	0.036	-1.385	0.166
	两阶段	所有数据	0.931	0.055	-2.939	0.003	0.793	0.047	-1.801	0.072
GNR 或衣原体	朴素	调整 X，Z，W	0.842	0.022	-6.017	< 0.001	0.531	0.014	-3.749	< 0.001
	IPCW	结局桥接函数	0.905	0.039	-5.964	< 0.001	0.716	0.031	-3.873	< 0.001
	IPCW	倾向评分	0.905	0.039	-5.933	< 0.001	0.716	0.031	-3.836	< 0.001
	IPCW	双重稳健	0.905	0.039	-5.340	< 0.001	0.716	0.031	-3.873	< 0.001
	两阶段	1年	0.879	0.030	-4.477	< 0.001	0.641	0.022	-2.642	0.008
	两阶段	所有数据	0.898	0.036	-4.639	< 0.001	0.695	0.028	-2.920	0.004

S1.2 使用淋病和衣原体作为NCO

本节展示将直肠淋病和直肠衣原体感染作为 NCO 时的结果。主要分析未纳入直肠衣原体感染的结果，因为其逻辑回归分析 $W \sim Z + \boldsymbol{X}$ 表明它并非未测量混杂因素 U的良好代理变量；主要分析也未纳入直肠淋病的结果，因为其在 HPTN 083 和 AMP 中的患病率均较低。结果见图S2 和表 S2。将直肠淋病作为 NCO 时，估计的 1 年内反事实 HIV 累积发生率与主要分析的结果相似，但由于直肠淋病患病率较低，置信区间更宽。

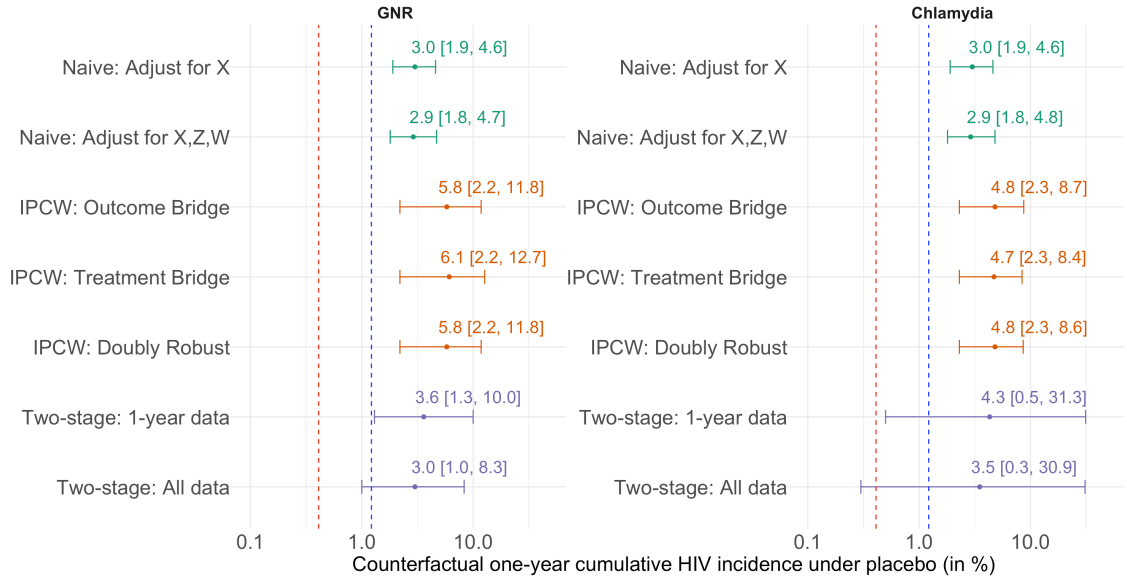


图 S2：采用不同方法，并以直肠淋病和衣原体作为 NCO，估计 HPTN 083 中截至 1 年的反事实 HIV 累积发生率（%）。红色虚线表示卡博特韦组的观察发生率（每 100 人年 0.41），蓝色虚线表示 TDF/FTC 组的观察发生率（每 100 人年 1.22）。**年龄、种族和性别** 均在所有模型中进行了调整。**AMP 数据中仅使用安慰剂组** 作为外部数据。为便于观察，横轴采用 log() 转换绘制。

表 S2：比较活性治疗组与估计的安慰剂组截至 1 年的 HIV 累积发病率的 Wald 检验统计量。表中展示了不同分析方法下卡博特韦与安慰剂、恩曲他滨/富马酸替诺福韦二吡呋酯（TDF/FTC）与安慰剂的统计学比较。**年龄、种族和性别** 均在所有模型中进行了调整。**AMP 数据中仅使用安慰剂组** 作为外部数据。

NCO	方法	设定	卡博特韦				TDF/FTC			
			相对疗效	绝对疗效	检验统计量	p值	相对疗效	绝对疗效	检验统计量	p值
	朴素	调整 X，Z，W	0.863	0.026	-5.593	< 0.001	0.593	0.018	-3.292	0.001
GNR	朴素	调整 X，Z，W	0.859	0.025	-5.382	< 0.001	0.579	0.017	-3.111	0.002
	IPCW	结局桥接函数	0.929	0.054	-4.212	< 0.001	0.790	0.046	-2.859	0.004
	IPCW	倾向评分	0.933	0.057	-4.149	< 0.001	0.800	0.049	-2.853	0.004
	IPCW	双重稳健	0.929	0.054	-4.198	< 0.001	0.790	0.046	-2.849	0.004
	两阶段	1年	0.886	0.032	-2.853	0.004	0.661	0.024	-1.877	0.061
	两阶段	所有数据	0.863	0.026	-3.094	0.002	0.593	0.018	-2.228	0.026
衣原体	朴素	调整 X，Z，W	0.859	0.025	-5.177	< 0.001	0.579	0.017	-2.900	0.004
	IPCW	结局桥接函数	0.915	0.044	-4.913	< 0.001	0.746	0.036	-3.217	0.001
	IPCW	倾向评分	0.913	0.043	-4.927	< 0.001	0.740	0.035	-3.202	0.001
	IPCW	双重稳健	0.915	0.044	-4.923	< 0.001	0.746	0.036	-3.222	0.001
	两阶段	1年	0.905	0.039	-1.858	0.063	0.716	0.031	-0.947	0.344
	两阶段	所有数据	0.883	0.031	-1.817	0.069	0.651	0.023	-0.931	0.352

S1.3 其他协变量集

在主要分析中，我们在模型中将年龄、种族、性别、族裔和教育程度作为协变量进行调整 $\boldsymbol{X}$。在本节中，我们探讨其他调整集：（1）年龄、种族（图S3 和表S3）；（2）年龄、种族、性别、族裔、教育程度（图S4 和表S4）。请注意，正如第3节所述，Cox 模型仅显示年龄和种族与结局之间存在显著关联。我们首先进行了回归 $W \sim Z + \boldsymbol{X}$，以评估负对照结局的强度。当使用协变量集（1）作为 $\boldsymbol{X}$时，使用直肠衣原体感染作为负对照结局时，Z 的p值为0。143；而使用直肠淋病或任一种性传播感染作为负对照结局时，p值均小于0.001。这表明，与其他负对照结局相比，直肠衣原体感染作为未测量混杂因素代理变量的可靠性较低。同样，当使用协变量集（2）作为 $\boldsymbol{X}$时，使用直肠淋病、直肠衣原体感染和任一种性传播感染作为负对照结局时，p值分别为0。029、0。138和0。018 。此外，在 HPTN 083 和 AMP 中，直肠淋病的患病率均较低，分别为6.5%和3.7%。因此，与主要分析一样，将任一种性传播感染合并使用是更好的负对照结局。

在调整年龄和种族（即集合 (2)）时，以任一性传播感染作为 NCO 的估计结果与主要分析相似，置信区间略小。在调整更多协变量（即集合 (2)）时，所有方法对 1 年反事实 HIV 累积发生率的估计值均更大，且置信区间更宽。这可能是由于我们的事件率较低，且在纳入 5 个协变量时估计变得不稳定。在所有协变量集合中，以任一性传播感染合并作为 NCO 时，我们均观察到卡博特韦相对于安慰剂组具有一致的统计学显著优势，而基于回归的两阶段方法中 TDF/FTC 的结果则不一致。

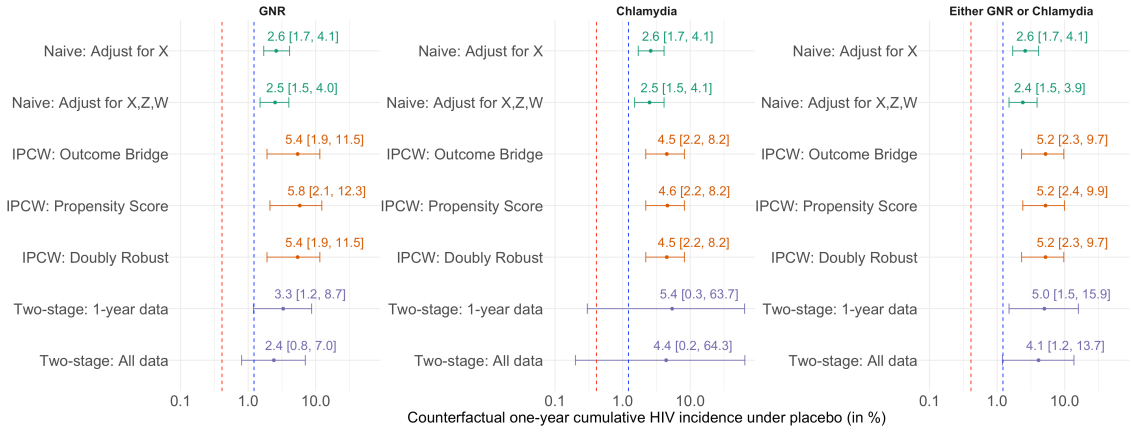


图 S3：使用不同方法和 NCO 定义估计的 HPTN 083 的 1 年反事实 HIV 累积发生率。红色虚线表示卡博特韦组中观察到的发生率（每 100 人年 0.41），蓝色虚线表示 TDF/FTC 组中观察到的发生率（每 100 人年 1.22）。所有模型均对**年龄和种族** 进行了调整。**仅安慰剂组** 在 AMP 中被用作外部数据。为便于可视化，x 轴采用 log() 变换绘制。

表 S3：比较活性治疗组与估计的安慰剂组截至 1 年的 HIV 累积发病率的 Wald 检验统计量。统计比较展示了采用不同分析方法时卡博特韦与安慰剂，以及恩曲他滨/富马酸替诺福韦二吡呋酯（TDF/FTC）与安慰剂之间的比较。**所有模型均调整了年龄和种族。外部数据仅使用 AMP数据中的安慰剂组。**

NCO	方法	设定	卡博特韦				TDF/FTC			
			相对疗效（%）	绝对疗效（%）	检验统计量	p值	相对疗效（%）	绝对疗效（%）	检验统计量	p值
	朴素	调整 X	0.842	0.022	-5.252	< 0.001	0.531	0.014	-2.842	0.004
GNR	朴素	调整 X，Z，W	0.836	0.021	-4.874	< 0.001	0.512	0.013	-2.480	0.013
	IPCW	结局桥接函数	0.924	0.050	-3.922	< 0.001	0.774	0.042	-2.609	0.009
	IPCW	倾向评分	0.929	0.054	-4.003	< 0.001	0.790	0.046	-2.715	0.007
	IPCW	双重稳健	0.924	0.050	-3.919	< 0.001	0.774	0.042	-2.606	0.009
	两阶段	1年	0.876	0.029	-2.964	0.003	0.630	0.021	-1.716	0.086
	两阶段	所有数据	0.829	0.020	-3.118	0.002	0.492	0.012	-1.860	0.063
衣原体	朴素	调整 X，Z，W	0.836	0.021	-4.811	< 0.001	0.512	0.013	-2.442	0.015
	IPCW	结局桥接函数	0.909	0.041	-4.788	< 0.001	0.729	0.033	-3.077	0.002
	IPCW	倾向评分	0.911	0.042	-4.864	< 0.001	0.735	0.034	-3.134	0.002
	IPCW	双重稳健	0.909	0.041	-4.810	< 0.001	0.729	0.033	-3.091	0.002
	两阶段	1年	0.924	0.050	-1.605	0.108	0.774	0.042	-0.888	0.375
	两阶段	所有数据	0.907	0.040	-1.492	0.136	0.723	0.032	-0.860	0.390
GNR 或衣原体	朴素	调整 X，Z，W	0.829	0.020	-4.780	< 0.001	0.492	0.012	-2.381	0.017
	IPCW	结局桥接函数	0.921	0.048	-4.703	< 0.001	0.765	0.040	-3.129	0.002
	IPCW	倾向评分	0.921	0.048	-4.710	< 0.001	0.765	0.040	-3.146	0.002
	IPCW	双重稳健	0.921	0.048	-4.702	< 0.001	0.765	0.040	-3.125	0.002
	两阶段	1年	0.918	0.046	-3.676	< 0.001	0.756	0.038	-2.116	0.034
	两阶段	所有数据	0.900	0.037	-3.509	< 0.001	0.702	0.029	-2.055	0.040

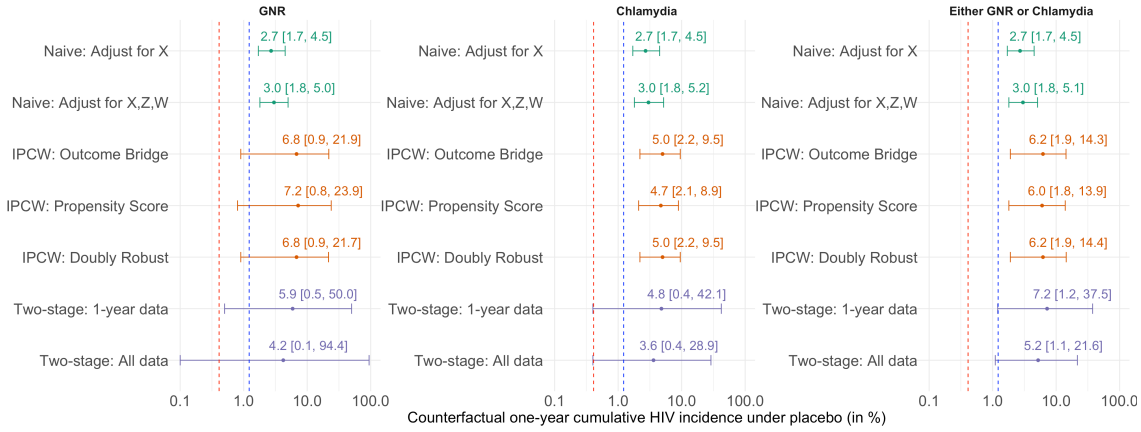


图 S4：使用不同方法和 NCO 定义估计的 HPTN083 反事实 HIV 累积发生率截至 1 年的数值（%）。红色虚线表示卡博特韦组观察到的发生率（每 100 人年 0.41），蓝色虚线表示 TDF/FTC 组观察到的发生率（每 100 人年 1.22 ）。**所有模型均调整了年龄、种族、性别、族裔和教育程度。外部数据仅使用 AMP中的安慰剂组**。x 轴采用 log() 变换绘制，以便更清晰地显示。

表 S4：比较活性治疗组与估计的安慰剂组截至 1 年的 HIV累积发病率的 Wald检验统计量。针对不同的分析方法，分别呈现卡博特韦与安慰剂、恩曲他滨/富马酸替诺福韦二吡呋酯（TDF/FTC）与安慰剂之间的统计学比较。**所有模型均调整了年龄、种族、性别、族裔和教育程度。外部数据仅采用 AMP数据中的安慰剂组。**

NCO	方法	设定	卡博特韦				TDF/FTC			
			相对疗效 (%)	绝对疗效 (%)	检验统计量	p值	相对疗效 (%)	绝对疗效 (%)	检验统计量	p值
	朴素	调整 X	0.848	0.023	-5.091	< 0.001	0.548	0.015	-2.737	0.006
GNR	朴素	调整 X，Z，W	0.863	0.026	-5.328	< 0.001	0.593	0.018	-3.088	0.002
	IPCW	结局桥接函数	0.940	0.064	-2.438	0.015	0.821	0.056	-1.696	0.090
	IPCW	倾向评分	0.943	0.068	-2.360	0.018	0.831	0.060	-1.663	0.096
	IPCW	双重稳健	0.940	0.064	-2.456	0.014	0.821	0.056	-1.710	0.087
	两阶段	1年	0.931	0.055	-1.462	0.144	0.793	0.047	-0.855	0.393
	两阶段	所有数据	0.902	0.038	-2.628	0.009	0.710	0.030	-1.554	0.120
衣原体	朴素	调整 X，Z，W	0.863	0.026	-5.097	< 0.001	0.593	0.018	-2.872	0.004
	IPCW	结局桥接函数	0.918	0.046	-4.526	< 0.001	0.756	0.038	-2.970	0.003
	IPCW	倾向评分	0.913	0.043	-4.505	< 0.001	0.740	0.035	-2.904	0.004
	IPCW	双重稳健	0.918	0.046	-4.534	< 0.001	0.756	0.038	-2.976	0.003
	两阶段	1年	0.915	0.044	-2.031	0.042	0.746	0.036	-1.053	0.292
	两阶段	所有数据	0.886	0.032	-1.942	0.052	0.661	0.024	-1.028	0.304
GNR 或衣原体	朴素	调整 X, Z, W	0.863	0.026	-5.104	< 0.001	0.593	0.018	-2.882	0.004
	IPCW	结局桥接函数	0.934	0.058	-3.640	< 0.001	0.803	0.050	-2.503	0.012
	IPCW	倾向评分	0.932	0.056	-3.587	< 0.001	0.797	0.048	-2.443	0.015
	IPCW	双重稳健	0.934	0.058	-3.606	< 0.001	0.803	0.050	-2.480	0.013
	两阶段	1年	0.943	0.068	-3.389	0.001	0.831	0.060	-2.091	0.037
	两阶段	所有数据	0.921	0.048	-3.105	0.002	0.765	0.040	-1.932	0.053

S2 技术证明

S2.1 定理1的证明

证明。(a) 我们首先使用结果桥函数 h 展示识别公式。

首先，我们注意到

$$
\begin{aligned}
&E\left[\frac{\Delta I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R=1)} \mid Z, \boldsymbol{X}, R=1\right] \\
=&E\left[\frac{E[\Delta \mid T^*, Z, \boldsymbol{X}, R=1] I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R=1)} \mid Z, \boldsymbol{X}, R=1\right] \\
=&E\left[\frac{P(C > T^* \mid T^*, Z, \boldsymbol{X}, R=1) I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R=1)} \mid Z, \boldsymbol{X}, R=1\right] \\
=&E\left[\frac{P(C > T^* \mid Z, \boldsymbol{X}, R=1) I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R=1)} \mid Z, \boldsymbol{X}, R=1\right] \\
=&E[I(T^* \le t) \mid Z, \boldsymbol{X}, R=1] \\
=&E[I(T^*(0) \le t) \mid Z, \boldsymbol{X}, R=1]
\end{aligned}
$$

第一个等式依据期望的塔式性质；第二个等式依据定义；第三个等式依据假设 4；第四个等式依据代数运算；最后一个等式依据一致性。由定理1中的方程（1）可得

$$
E\left[\frac{\Delta I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R=1)} \mid Z, \boldsymbol{X}, R=1\right] = \int h(w, \boldsymbol{x}) dP(w \mid Z=z, R=1, \boldsymbol{X}=\boldsymbol{x}).
$$

因此，

$$
E[I(T^*(0) \le t) \mid Z=z, R=1, \boldsymbol{X}=\boldsymbol{x}] = \int h(w, \boldsymbol{x}) dP(w \mid Z=z, R=1, \boldsymbol{X}=\boldsymbol{x}),
$$

对每个 z， $\boldsymbol{x}$ 都成立。注意，左侧等于

$$
\begin{aligned}
&E[I(T^*(0) \le t) \mid Z=z, R=1, \boldsymbol{X}=\boldsymbol{x}] \\
&= \int E[I(T^*(0) \le t) \mid U_b=u, Z=z, R=1, \boldsymbol{X}=\boldsymbol{x}] dP(u \mid Z=z, R=1, \boldsymbol{X}=\boldsymbol{x}),
\end{aligned}
$$

且右侧等于

$$
\int h(w, \boldsymbol{x}) dP(w \mid Z=z, R=1, \boldsymbol{X}=\boldsymbol{x})
$$

$$= \int \int h(w, \boldsymbol{x}) dP(w \mid U_b = u, Z = z, R = 1, \boldsymbol{X} = \boldsymbol{x}) dP(u \mid Z = z, R = 1, \boldsymbol{X} = \boldsymbol{x})$$
$$= \int E[h(W, \boldsymbol{x}) \mid U_b = u, Z = z, R = 1, \boldsymbol{X} = \boldsymbol{x}] dP(u \mid Z = z, R = 1, \boldsymbol{X} = \boldsymbol{x}).$$

根据完备性假设 3(a)，对于在 $P(u, \boldsymbol{x}, z \mid R = 1)$ 中具有正密度的任意 $u, \boldsymbol{x}, z$，我们有 $E[I(T^*(0) \leq t) \mid U_b = u, Z = z, R = 1, \boldsymbol{X} = \boldsymbol{x}] = E[h(W, \boldsymbol{x}) \mid U_b = u, Z = z, R = 1, \boldsymbol{X} = \boldsymbol{x}]$。

于是，由 $(T^*(0)\ W) \perp\!\!\!\perp (Z, R) \mid U_b, \boldsymbol{X}$，我们有 $E[I(T^*(0) \leq t) \mid U_b = u, R = 0, \boldsymbol{X} = \boldsymbol{x}] = E[h(W, \boldsymbol{x}) \mid U_b = u, R = 0, \boldsymbol{X} = \boldsymbol{x}]$ 对任意在 $P(x \mid R = 1)$ 中具有正密度的 $u, \boldsymbol{x}$。由假设2，$P(u, \boldsymbol{x} \mid R = 0) > 0$ 蕴含 $P(u, \boldsymbol{x} \mid R = 1) > 0$。因此，

$$\begin{aligned}
&P(T^*(0) \leq t) \mid R = 0] \\
&= \int E[I(T^*(0) \leq t) \mid U_b = u, R = 0, \boldsymbol{X} = \boldsymbol{x}] dP(u, \boldsymbol{x} \mid R = 0) \\
&= \int E[h(W, \boldsymbol{x}) \mid U_b = u, R = 0, \boldsymbol{X} = \boldsymbol{x}] dP(u, \boldsymbol{x} \mid R = 0) \\
&= E[h(W, \boldsymbol{X}) \mid R = 0]
\end{aligned}$$

接下来，我们利用暴露桥函数证明识别公式 q。由$W \perp\!\!\!\perp (Z,\ R) \mid U_b,\ \boldsymbol{X}$ 以及方程（2）在定理1中，我们有

$$\frac{P(R = 0 \mid W = w, \boldsymbol{X} = \boldsymbol{x})}{P(R = 1 \mid W = w, \boldsymbol{X} = \boldsymbol{x})} = \int q(z, \boldsymbol{x}) dP(z \mid W = w, R = 1, \boldsymbol{X} = \boldsymbol{x})$$

对每一个 $w,\ \boldsymbol{x}$ 成立。在该方程中，左端等于

$$\begin{aligned}
&\frac{P(R = 0 \mid W = w, \boldsymbol{X} = \boldsymbol{x})}{P(R = 1 \mid W = w, \boldsymbol{X} = \boldsymbol{x})} \\
&= \frac{1}{P(R = 1 \mid W = w, \boldsymbol{X} = \boldsymbol{x})} - 1 \\
&= \int \frac{1}{P(R = 1 \mid U_b = u, W = w, \boldsymbol{X} = \boldsymbol{x})} dP(u \mid W = w, \boldsymbol{X} = \boldsymbol{x}, R = 1) - 1 \\
&= \int \left\{ \frac{1}{P(R = 1 \mid U_b = u, \boldsymbol{X} = \boldsymbol{x})} - 1 \right\} dP(u \mid W = w, \boldsymbol{X} = \boldsymbol{x}, R = 1),
\end{aligned}$$

且右侧等于

$$\int q(z, \boldsymbol{x}) dP(z \mid W = w, R = 1, \boldsymbol{X} = \boldsymbol{x})$$

$$= \int \int q(z, \boldsymbol{x}) dP(z \mid U_b = u, R = 1, \boldsymbol{X} = \boldsymbol{x}) dP(u \mid W = w, \boldsymbol{X} = \boldsymbol{x}, R = 1).$$

根据完备性假设3（b），我们有

$$\frac{P(R = 0 \mid U_b, \boldsymbol{X} = \boldsymbol{x})}{P(R = 1 \mid U_b, \boldsymbol{X} = \boldsymbol{x})} = \int q(z, \boldsymbol{x}) dP(z \mid U_b, R = 1, \boldsymbol{X} = \boldsymbol{x})$$

almost surely. Similarly from the above outcome bridge derivation, we can also show that $E[RI(T^*(0) \le t)q(Z, X)] = E\left[\frac{Rq(Z,X)\Delta I(T^* \le t)}{P(C > T^* | Z, X, R=1)}\right]$. Thus, by $T^*(0) \perp\!\!\!\perp (R, Z) \mid (U_b, \boldsymbol{X})$,

$$\begin{aligned}
&\frac{1}{P(R = 0)} E\left[\frac{Rq(Z, X)\Delta I(T^* \le t)}{P(C > T^* \mid Z, X, R = 1)}\right] \\
&= \frac{1}{P(R = 0)} E[RI(T^*(0) \le t)q(Z, \boldsymbol{X})] \\
&= \frac{1}{P(R = 0)} \int E[RI(T^*(0) \le t)q(Z, \boldsymbol{X}) \mid \boldsymbol{X} = \boldsymbol{x}] dP(\boldsymbol{x}) \\
&= \frac{1}{P(R = 0)} \int E[E[RI(T^*(0) \le t)q(Z, \boldsymbol{X}) \mid U_b, \boldsymbol{X}] \mid \boldsymbol{X} = \boldsymbol{x}] dP(\boldsymbol{x}) \\
&= \frac{1}{P(R = 0)} \int E[E[I(T^*(0) \le t) \mid U_b, \boldsymbol{X}] E[Rq(Z, \boldsymbol{X}) \mid U_b, \boldsymbol{X}] \mid \boldsymbol{X} = \boldsymbol{x}] dP(\boldsymbol{x}) \\
&= \frac{1}{P(R = 0)} \int E[E[I(T^*(0) \le t) \mid U_b, \boldsymbol{X}] P(R = 1 \mid U_b, \boldsymbol{X}) E[q(Z, \boldsymbol{X}) \mid U_b, \boldsymbol{X}, R = 1] \mid \boldsymbol{X} = \boldsymbol{x}] dP(\boldsymbol{x}) \\
&= \frac{1}{P(R = 0)} \int E\left[E[I(T^*(0) \le t) \mid U_b, \boldsymbol{X}] P(R = 0 \mid U_b, \boldsymbol{X}) \mid \boldsymbol{X} = \boldsymbol{x}\right] dP(\boldsymbol{x}) \\
&= \frac{1}{P(R = 0)} \int E\left[E[(1 - R)I(T^*(0) \le t) \mid U_b, \boldsymbol{X}] \mid \boldsymbol{X} = \boldsymbol{x}\right] dP(\boldsymbol{x}) \\
&= \frac{1}{P(R = 0)} \int E\left[(1 - R)I(T^*(0) \le t) \mid \boldsymbol{X} = \boldsymbol{x}\right] dP(\boldsymbol{x}) \\
&= E\left[I(T^*(0) \le t) \mid R = 0\right].
\end{aligned}$$

(b) 接下来，我们证明识别公式（5）的双重稳健性质。

首先，假设 $h(W, \boldsymbol{X})$ 设定正确，但 $q(Z, \boldsymbol{X})$ 可能有误。方程（5）的右侧可写为

$$\underbrace{\frac{1}{P(R = 0)} E\left[Rq(Z, \boldsymbol{X})\left(\frac{\Delta I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R = 1)} - h(W, \boldsymbol{X})\right)\right]}_{A} + \underbrace{\frac{1}{P(R = 0)} E\left[(1 - R)h(W, \boldsymbol{X})\right]}_{B}.$$

注意，根据 (a)，我们有 $B = E[h(W, \boldsymbol{X}) \mid R = 0) = P(T^*(0) \le t \mid R = 0)$。然后，i t 尚需证明 $A = 0$，这可由定理1中的方程（1）推出。我们有

$$E\left[Rq(Z, \boldsymbol{X})\left(\frac{\Delta I(T^* \le t)}{P(C > T^* \mid Z, \boldsymbol{X}, R = 1)} - h(W, \boldsymbol{X})\right)\right]$$

$$
\begin{aligned}
&= E\left[E\left[Rq(Z,\boldsymbol{X})\left(\frac{\Delta I(T^*\le t)}{P(C>T^*\mid Z,\boldsymbol{X},R=1)}-h(W,\boldsymbol{X})\right)\mid Z,X,R=1\right]\right]\\
&= E\left[q(Z,\boldsymbol{X})E\left[\left(\frac{\Delta I(T^*\le t)}{P(C>T^*\mid Z,\boldsymbol{X},R=1)}-h(W,\boldsymbol{X})\right)\mid Z,X,R=1\right]\right]\\
&= 0
\end{aligned}
$$

无论 $q(Z,\ \boldsymbol{X})$ 是否设定正确，因为

$$
E\left[\left(\frac{\Delta I(T^*\le t)}{P(C>T^*\mid Z,\boldsymbol{X},R=1)}-h(W,\boldsymbol{X})\right)\mid Z,X,R=1\right]=0
$$

由定理1中的方程（1）。

那么，假设 q（$Z,\ \boldsymbol{X}$）是正确的，但 h（$W,\ \boldsymbol{X}$）可能是错误的。定理1中方程（5）的右侧可写为

$$
\underbrace{\frac{1}{P(R=0)}E\left[Rq(Z,\boldsymbol{X})\frac{\Delta I(T^*\le t)}{P(C>T^*\mid Z,\boldsymbol{X},R=1)}\right]}_{A'}+\underbrace{\frac{1}{P(R=0)}E\left[-Rq(Z,\boldsymbol{X})h(W,\boldsymbol{X})+(1-R)h(W,\boldsymbol{X})\right]}_{B'}.
$$

注意，根据本证明的（a）部分，我们有 $A'=P$（T^*（0）$\le t\mid R=0$）。然后，只需证明 $B'=0$，这可由定理1中的方程（2）推出。我们有

$$
\begin{aligned}
&E[Rq(Z,\boldsymbol{X})h(W,\boldsymbol{X})]\\
=&E\left[E\left[Rq(Z,\boldsymbol{X})h(W,\boldsymbol{X})\mid W,\boldsymbol{X}\right]\right]\\
=&E[h(W,\boldsymbol{X})E[Rq(Z,\boldsymbol{X})\mid W,\boldsymbol{X}]]\\
=&E[h(W,\boldsymbol{X})P(R=1\mid R,\boldsymbol{X})E[q(Z,\boldsymbol{X})\mid W,\boldsymbol{X},R=1]]\\
=&E\left[h(W,\boldsymbol{X})P(R=1\mid W,\boldsymbol{X})\frac{P(R=0\mid W,\boldsymbol{X})}{P(R=1\mid W,\boldsymbol{X})}\right]\\
=&E[h(W,\boldsymbol{X})P(R=0\mid W,\boldsymbol{X})]\\
=&E[(1-R)h(W,\boldsymbol{X})],
\end{aligned}
$$

其中第四个等式由定理1中的方程（2）得出，因此 $B'=$ 0.

□

S2.2 定理2的证明

证明。 我们从模型开始

$$\lambda(t|R,Z,X,U) = \lambda_0(t)exp(\beta_{Tx}X + \beta_{Tu}U)$$

当事件率较低时，对于所有 t，我们有 $S(t|R,\ U,\ X) \approx 1$。

Then, the density function $f(t|R,U,X) \approx \lambda(t|R,U,X)$.

注意

$$\begin{aligned} f(t|R,Z,X) &= E[f(t|R,U,X,Z)|R,Z,X] \\ &= E[f(t|R,U,X)|R,Z,X] \\ &\approx E[\lambda(t|R,U,X)|R,Z,X] \\ &= \lambda_0(t)exp(\beta_{Tx}X)E[exp(\beta_{Tu}U)|R,Z,X] \end{aligned}$$

假设 $U = E[U|R,Z,X] + \epsilon$，其中 $\epsilon \perp\!\!\!\perp (R,\ Z,\ X)$，且 $E(\epsilon) = 0$。

于是，$f(t|R,\ Z,\ X) \approx \lambda_0(t)exp(\beta_{Tx}X+\beta_{Tu}E[U|R,Z,X])$，其中 $\lambda_0(t) = \lambda_0(t)E[exp(\beta_{Tu}\epsilon)]$。接下来需要确定 $E(U|R,\ Z,\ X)$。我们可以使用我们的负对照结局 W。

假设

$$E(W|R,U,X) = exp(\beta_{W0} + \beta_{Wx}X + \beta_{Wu}U).$$

注意

$$\begin{aligned} E(W|R,Z,X) &= E[E(W|R,U,Z,X)|R,Z,X] \\ &= E[E(W|R,U,X)|R,Z,X] \\ &= E[exp(\beta_{W0} + \beta_{Wx}X + \beta_{Wu}U)|R,Z,X] \\ &= exp(\beta_{W0} + \beta_{Wx}X)E[exp(\beta_{Wu}U)|R,Z,X] \end{aligned}$$

再次，利用假设 $U = E[U|R,Z,X] + \epsilon$ 其中 $\epsilon \perp\!\!\!\perp (R,\ Z,\ X)$ 且 $E(\epsilon) = 0$，我们有

$$E[W|R,Z,X] = exp(\tilde{\beta}_{W0} + \beta_{Wx}X + \beta_{Wu}E(U|R,Z,X))$$

˜其中 $\beta_{W0} = \beta_{W0} + log(E[exp(\beta_{Wu}\epsilon)])$。

因此， $E(U|R,\ Z,\ X)$可以表示为 X和 $log(E[W|R,Z,X])$的线性组合。

这意味着，

$$f(t|R,Z,X) \approx \tilde{\lambda}_0(t)exp[\beta^*_{Tx}X + \beta^*_{Tw}log(E(W|R,Z,X))]$$

因此

$$\lambda(t|R,Z,X) \approx \tilde{\lambda}_0(t)exp[\tilde{\beta}_{Tx}X + \tilde{\beta}_{Tw}log(E(W|R,Z,X))] \tag{S1}$$

˜ ˜ ˜其中 $\lambda_0(t)$、 β_{Tx} 和 β_{Tw} 是未知参数。 □

S3 违反假设的敏感性分析 1 *

假设潜在的 U 是连续的。然而，为了在只有二元 Z 和 W时满足完备性假设，我们必须假设存在一个二元 U_b ，它是 U的粗化版本。为评估我们的识别对这一粗化假设有多敏感，我们推导出以下界

$$\frac{1}{\Gamma_1^2}E[h(W,X)|S=0] \le E[Y(0)|S=0] \le \Gamma_1^2 E[h(W,X)|S=0],$$

其中我们假设 $\frac{1}{\Gamma_1} \le QR(P(R=0|u_b,\ x,\ z)\ P(R=0|u,\ z,\ x) \le \Gamma_1$对几乎处处的 $u,\ z,\ x$成立。

所需假设:

- 关于负对照的假设1：$(Z,\ R) \perp\!\!\!\perp (Y(0)\ W) \mid \boldsymbol{X},\ U$
- 假设 3^* （完备性）。(a) 对于外部数据集中具有正密度的任意 $\boldsymbol{x}$ 以及任意平方可积函数 g, $E[g(U_b) \mid Z,\ R=1,\ \boldsymbol{X}=\boldsymbol{x}] = 0$ 几乎必然蕴含 $g(U_b) = 0$ 几乎必然。
- h 满足 $E[h(W,\ X)|z,\ x,\ R=1] = E[Y(0)|x,\ z,\ R=1]$ 对几乎处处的 $x,\ z$

注意，在假设 3^* 下， h 也满足

$$E[h(W,X)|u_b,z,x,R=1] = E[Y(0)|u_b,z,x,R=1]$$

那么，

$$
\begin{aligned}
E[Y(0)|u_b,z,x,R=0] &= \sum_u E[Y(0)|u,u_b,z,x,R=0]P(u|u_b,z,x,R=0) \\
&= \sum_u E[Y(0)|u,z,x,R=0]P(u|u_b,z,x,R=0) \\
&= \sum_u E[Y(0)|u,z,x,R=1]P(u|u_b,z,x,R=0) \\
&= \sum_u E[Y(0)|u,z,x,R=1]P(u|u_b,z,x,R=1)\frac{P(u|u_b,z,x,R=0)}{P(u|u_b,z,x,R=1)}
\end{aligned}
$$

令 $r(u,u_b,z,x) = \frac{P(u|u_b,z,x,R=0)}{P(u|u_b,z,x,R=1)}$。

$$
\begin{aligned}
r(u,u_b,z,x) &= \frac{P(u,u_b,z,x,R=0)/P(u_b,z,x,R=0)}{P(u,u_b,z,x,R=1)/P(u_b,z,x,R=1)} \\
&= \frac{P(R=0|u,z,x)/P(R=0|u_b,z,x)}{P(R=1|u,z,x)/P(R=1|u_b,z,x)} \\
&= OR(P(R=0|u,z,x),P(R=0|u_b,z,x)),
\end{aligned}
$$

根据假设，其值介于 $\frac{1}{\Gamma_1}$ 和 Γ_1之间。

因此，

$$
\begin{aligned}
E[Y(0)|u_b,z,x,R=0] &\le \Gamma_1 \sum_u E[Y(0)|u,z,x,R=1]P(u|u_b,z,x,R=1) \\
&= \Gamma_1 E[Y(0)|u_b,z,x,R=1]
\end{aligned}
$$

类似地，我们有

$$
\begin{aligned}
E[Y(0)|u_b,z,x,R=0] &\ge \Gamma_1 \sum_u E[Y(0)|u,z,x,R=1]P(u|u_b,z,x,R=1) \\
&= \frac{1}{\Gamma_1} E[Y(0)|u_b,z,x,R=1]
\end{aligned}
$$

因此，

$$
\frac{1}{\Gamma_1} E[Y(0)|u_b,z,x,R=1] \le E[Y(0)|u_b,z,x,R=0] \le \Gamma_1 E[Y(0)|u_b,z,x,R=1] \tag{S2}
$$

接下来，我们推导 $E[Y(0)|S=0]$ 的一个界。

$$
E[Y(0)|R=0] = \sum_{u_b,z,x} E[Y(0)|u_b,z,x,R=0]P(u_b,z,x|R=0)
$$

$$
\begin{aligned}
&\leq \Gamma_1 \sum_{u_b,z,x} E[Y(0)|u_b,z,x,R=1]P(u_b,z,x|R=0) \\
&= \Gamma_1 \sum_{u_b,z,x} E[h(W,X)|u_b,z,x,R=1]P(u_b,z,x|R=0) \\
&= \Gamma_1 \sum_{u_b,z,x} \sum_{u} E[h(W,X)|u,u_b,z,x,R=1]P(u|u_b,z,x,R=1)P(u_b,z,x|R=0) \\
&= \Gamma_1 \sum_{u_b,z,x} \sum_{u} E[h(W,X)|u,z,x,R=1]P(u|u_b,z,x,R=1)P(u_b,z,x|R=0) \\
&= \Gamma_1 \sum_{u_b,z,x} \sum_{u} E[h(W,X)|u,z,x,R=0]P(u|u_b,z,x,R=1)P(u_b,z,x|R=0) \\
&= \Gamma_1 \sum_{u_b,z,x} \sum_{u} E[h(W,X)|u,z,x,R=0]P(u|u_b,z,x,R=0)P(u_b,z,x|R=0)\frac{P(u|u_b,z,x,R=1)}{P(u|u_b,z,x,R=0)} \\
&= \Gamma_1 \sum_{u_b,z,x} \sum_{u} E[h(W,X)|u,z,x,R=0]P(u,z,x|R=0)\frac{P(u|u_b,z,x,R=1)}{P(u|u_b,z,x,R=0)}
\end{aligned}
$$

注意，$\frac{P(u|u_b,z,x,R=1)}{P(u|u_b,z,x,R=0)} = r_u^1(u_b,z,x)$
也受 $\frac{1}{\Gamma_1}$ 和 Γ_1 的界定。因此，

$$
\begin{aligned}
E[Y(0)|R=0] &\leq \Gamma_1^2 \sum_{u_b,z,x} \sum_{u} E[h(W,X)|u,z,x,R=0]P(u,z,x|R=0) \\
&= \Gamma_1^2 E[h(W,X)|R=0]
\end{aligned}
$$

同样，我们有

$$
\begin{aligned}
E[Y(0)|R=0] &\geq \Gamma_1^2 \sum_{u_b,z,x} \sum_{u} E[h(W,X)|u,z,x,R=0]P(u,z,x,R=0) \\
&= \frac{1}{\Gamma_1^2} E[h(W,X)|R=0]
\end{aligned}
$$

因此，

$$
\frac{1}{\Gamma_1^2} E[h(W,X)|R=0] \leq E[Y(0)|R=0] \leq \Gamma_1^2 E[h(W,X)|R=0]
$$

现在，我们基于利用处理桥函数q进行的识别，推导出一个类似的结果。具体而言，

$$
\frac{1}{\Gamma_2^2}\frac{1}{P(R=0)} E[I(R=1)q(Z,X)Y] \leq E[Y(0)|S=0] \leq \Gamma_2^2 \frac{1}{P(R=0)} E[I(R=1)q(Z,X)Y],
$$

其中我们假设 $\frac{1}{\Gamma_2} \leq OR(P(R=0|u_b,\ w,\ x)\ P(R=0|u,\ w,\ x) \leq \Gamma_2$ 且

$$\frac{1}{\Gamma_2} \leq \frac{E[Y(0)|u_b, w, x, R=1]}{E[Y(0)|u, u_b, w, x, R=1]} \leq \Gamma_2$$

对几乎处处的 u, w, x成立。

所需假设:

- 关于负对照的假设1：$(Z, R) \perp\!\!\!\perp (Y(0)\ W)\ |\ \boldsymbol{X}, U$

- 假设 3^* (完备性)。 (a) 对于在外部数据集中具有正密度的任意 $\boldsymbol{x}$ ， 以及任意平方可积函数 g, $E[g(U_b)\ |\ W, R=1, \boldsymbol{X}=\boldsymbol{x}]=0$ 几乎必然意味着 $g(U_b)=0$ 几乎必然。

- q 满足 $\frac{P}{(R=0|w,x)\ P(R=1|w,x)} = E[q(Z, X)|w, x, R=1]]$ 对几乎处处的 x, w

根据假设， q 满足

$$\frac{P(R=0|W, X)}{P(R=1|W, X)} = E[q(Z, X)|W, X, R=1]] \tag{S3}$$

这等价于

$$P(R=0|W, X) = E[I(R=1)q(Z, X)|W, X]$$

在假设 $3*$ 下， q 也满足

$$P(R=0|U_b, W, X) = E[I(R=1)q(Z, X)|U_b, W, X]$$

原因如下。左端 S3

$$\begin{aligned}
\frac{P(R=0|W, X)}{P(R=1|W, X)} &= 1 - \frac{1}{P(R=1|W, X)} \\
&= 1 - \sum_{u_b} \frac{1}{P(R=1|u_b, W, X)} P(u_b|W, X, R=1) \\
&= \sum_{u_b} \frac{P(R=0|u_b, W, X)}{P(R=1|u_b, W, X)} P(u_b|W, X, R=1) \\
&= E[\frac{P(R=0|u_b, W, X)}{P(R=1|u_b, W, X)}|W, X, R=1]
\end{aligned}$$

右端 S3 是

$$E[q(Z,X)|W,X,R=1]=E[E[q(Z,X)|U_b,W,X,R=1]|W,X,R=1]$$

基于 U_b 的完备性假设，我们得出

$$\frac{P(R=0|U_b,W,X)}{P(R=1|U_b,W,X)}=E[q(Z,X)|U_b,W,X,R=1]$$

这意味着所需结果。

通过与证明（S2）中类似的推导，我们有

$$\frac{1}{\Gamma_2}E[Y(0)|u_b,w,x,R=1]\leq E[Y(0)|u_b,w,x,R=0]\leq\Gamma_2E[Y(0)|u_b,w,x,R=1]. \quad \text{(S4)}$$

然后，

$$\begin{aligned}
E[Y(0)|R=0]&=\frac{1}{P(R=0)}\sum_{u_b,w,x}E[Y(0)|u_b,w,x,R=0]P(u_b,w,x,R=0)\\
&\leq\frac{\Gamma_2}{P(R=0)}\sum_{u_b,w,x}E[Y(0)|u_b,w,x,R=1]P(R=0|u_b,w,x)P(u_b,w,x)\\
&=\frac{\Gamma_2}{P(R=0)}\sum_{u_b,w,x}E[Y(0)|u_b,w,x,R=1]E[I(R=1)q(Z,X)|u_b,w,x]P(u_b,w,x)\\
&=\frac{\Gamma_2}{P(R=0)}\sum_{u_b,w,x}E[Y(0)|u_b,w,x,R=1]E[q(Z,X)|u_b,w,x,R=1]P(R=1|u_b,w,x)P(u_b,w,x)\\
&=\frac{\Gamma_2}{P(R=0)}\sum_{u_b,w,x}E[Y(0)|u_b,w,x,R=1]\Big\{\sum_u E[Y(0)q(Z,X)|u,u_b,w,x,R=1]\\
&P(u|u_b,w,x,R=1)\Big\}P(R=1|u_b,w,x)P(u_b,w,x)\\
&=\frac{\Gamma_2}{P(R=0)}\sum_{u_b,w,x}\Big\{\sum_u\frac{E[Y(0)|u_b,w,x,R=1]}{E[Y(0)|u,u_b,w,x,R=1]}E[Y(0)q(Z,X)|u,u_b,w,x,R=1]\\
&P(u|u_b,w,x,R=1)\Big\}P(R=1|u_b,w,x)P(u_b,w,x)
\end{aligned}$$

根据以下假设：对于任意 u，u_b，w，x，

$$\frac{1}{\Gamma_2}\leq\frac{E[Y(0)|u_b,w,x,R=1]}{E[Y(0)|u,u_b,w,x,R=1]}\leq\Gamma_2$$

然后，

$$E[Y(0)|R=0]=\frac{\Gamma_2}{P(R=0)}\sum_{u_b,w,x}\Big\{\sum_u\frac{E[Y(0)|u_b,w,x,R=1]}{E[Y(0)|u,u_b,w,x,R=1]}E[Y(0)q(Z,X)|u,u_b,w,x,R=1]$$

$$
\begin{aligned}
&P(u|u_b,w,x,R=1)\Big\}P(R=1|u_b,w,x)P(u_b,w,x)\\
&\leq \frac{\Gamma_2}{P(R=0)}\sum_{u_b,w,x}\Big\{\sum_u E[Y(0)q(Z,X)|u,u_b,w,x,R=1]\\
&P(u|u_b,w,x,R=1)\Big\}P(R=1|u_b,w,x)P(u_b,w,x)\\
&=\frac{\Gamma_2^2}{P(R=0)}\sum_{u_b,w,x}E[Y(0)q(Z,X)|u_b,w,x,R=1]P(R=1|u_b,w,x)P(u_b,w,x)\\
&=\frac{\Gamma_2^2}{P(R=0)}\sum_{u_b,w,x}E[I(R=1)Y(0)q(Z,X)|u_b,w,x]P(u_b,w,x)\\
&=\frac{\Gamma_2^2}{P(R=0)}E[I(R=1)Yq(Z,X)]
\end{aligned}
$$

类似地，我们可以证明

$$\frac{1}{\Gamma_2^2}\frac{1}{P(R=0)}E[I(R=1)q(Z,X)Y]\leq E[Y(0)|S=0].$$

在上述推导中，我们强调以下两个假设:

1. 令 $r(u,\ u_b,\ w,\ x) = \frac{P(u|u_b,\ w,\ x,\ R=0)P(u|u_b,}{w,\ x,\ R=1)}$。

$$
\begin{aligned}
r(u,u_b,w,x)&=\frac{P(u,u_b,w,x,R=0)/P(u_b,w,x,R=0)}{P(u,u_b,w,x,R=1)/P(u_b,w,x,R=1)}\\
&=\frac{P(R=0|u,w,x)/P(R=0|u_b,w,x)}{P(R=1|u,w,x)/P(R=1|u_b,w,x)}\\
&=OR(P(R=0|u,w,x),P(R=0|u_b,w,x)),
\end{aligned}
$$

假设其取值介于 $\frac{1}{\Gamma_2}$ 和 Γ_2之间。 —

2. 令 $\tilde{r}(u,\ u_b,\ w,\ x) = \frac{E[q(\ \ Z\ X\ |u_b\ w\ x\ R=1]\ ,\ ),\ ,\ ,\ (,\ ),\ ,\ ,\ ,\ ,}{E[q\ Z\ X\ |u\ u_b\ w\ x\ R=1]}$。我们假设

$$\frac{1}{\Gamma_2}\leq\tilde{r}(u,u_b,w,x)\leq\Gamma_2.$$

注意，根据完备性假设 3*,

$$E[q(Z,X)|u_b,w,x,R=1]=\frac{P(R=0|u_b,w,x)}{P(R=1|u_b,w,x)}$$

如果还满足

$$E[q(Z,X)|u,w,x,R=1]=\frac{P(R=0|u,w,x)}{P(R=1|u,w,x)}$$

那么，上述两个假设可以统一。

最后，对于双重稳健识别，令

$$\psi = \frac{1}{P(R=0)} E[I(R=1)q(Z,X)(Y-h(W,X))] + \frac{1}{P(R=0)} E[I(R=0)h(W,X)].$$

我们知道：

(1) 如果 $h(W, X)$ 满足 $E[Y|Z,X,R=1] = E[h(W, X)|Z, X, R=1]$，则

$$\psi = \frac{1}{P(R=0)} E[I(R=0)h(W,X)]$$

(2) 如果 $q(Z, X)$ 满足 $P(R=0|W, X) = E[I(R=1)q(Z, X)|W, X]$，则

$$\psi = \frac{1}{P(R=0)} E[I(R=1)q(Z,X)Y]$$

在第一种情况下，我们已经证明

$$\frac{1}{\Gamma_1^2}\psi \leq E[Y(0)|R=0] \leq \Gamma_1^2 \psi$$

在第二种情况下，我们已经证明

$$\frac{1}{\Gamma_2^2}\psi \leq E[Y(0)|R=0] \leq \Gamma_2^2 \psi$$

因此，我们有

$$\frac{1}{\min\{\Gamma_1,\Gamma_2\}^2}\psi \leq E[Y(0)|R=0] \leq \min\{\Gamma_1,\Gamma_2\}^2 \psi$$

我们开展了一项简单的模拟研究，以考察违反假设 1*所产生的后果。数据生成方式如下：

- $X \sim \text{Binom}(n, p = 0.3)$
- $U \sim \text{Truncnorm}(n, \text{mean} = 0.5, \text{sd} = 1)$
- $U_b = \text{I}(U > a)$
- $R \sim \text{Binom}(n, p = 1/(1 + \exp(-1.6 + 5U + 0.1X)))$
- $W \sim \text{Binom}(n, p = \exp(-2 + 2U - 0.2X))$

- $Z \sim$ 二项分布(n, $p = 1/(1 + \exp(1.5U_b + 0.5X))$)

对于 $R = 1$ 和 $R = 0$, 平均生存时间 T^* 分别为600个月和4000个月。真实生存时间的风险比通过exp($0.8X + 2U$)和exp($0.5X$ –$1.8U$)进行模拟。对于 $R = 1$ 和 $R = 0$, 平均删失时间 C 分别为30个月和50个月。删失时间的风险比通过exp($0.2Z + 0.2X$)和exp($0.2Z$ +$0.3X$)进行模拟。观测时间为 $T =$ min(T^*, C)。我们考虑的样本量为 $n = 6500$, 并改变 $a = 0.5$, 0。9（即 U_b是 U的良好或弱代理变量）。下表总结了结果。

表 S5：不同 a 条件下违反假设1*时逆概率删失加权估计的比较。模拟总次数为 1000。

	$a = 0.5$		$a = 0.9$	
	Bias	SE	Bias	SE
结局桥接函数	0.013	0.236	0.024	0.411
倾向评分桥接函数	0.010	0.234	0.021	0.411
双重稳健	0.013	0.236	0.024	0.411

当 $a = 0.5$ 时, U_b 与 U 之间以及 U_b 与 Z 之间的相关性更强，因为 U_b 的确定性较低。因此，当 R 的真实数据生成过程依赖于连续的 U （违反假设 1*）时，随着代理变量 Z增强，我们观察到较小的偏差和标准误。

S4 当 W 或 Z 不是 U的代理变量时

在本节中，我们旨在探讨这样一种情形： Z 可能不是 U 的代理变量，以及这会如何影响定理1中给出的识别结果。具体而言，我们不再假设 Z 满足假设 3(a)和4，而只假设 Z 满足

$$Z \perp\!\!\!\perp (U, Y(0), W) | \boldsymbol{X}, R = 1.$$

如果 U 不会直接影响 Z，则该假设成立。

在该假设下，于总体层面，不再存在能够求解以下观测积分方程的平方可积函数 q

$$\frac{P(R=0|W,X)}{P(R=1|W,X)} = E[q(Z,X)|W,X,R=1]. \tag{S5}$$

因为右端等于 $E[q(Z,X)|X,R=1]$，且该式不随 W变化。因此，定理1中的识别公式(4)无效。事实上，定理1中的识别公式(3)很可能同样无效，因为若不具备假设 3(a)，求解以下观测结果桥方程的结果桥函数 h 可能无法满足潜在结果桥方程 $E[Y(0)|U,X,R=1] = E[h(W,X)|U,X,R=1]$, ，而这是控制未测量混杂的关键结果。

$$E[Y(0)|Z,X,R=1] = E[h(W,X)|Z,X,R=1]. \tag{S6}$$

这表明，当 Z 不是 U的代理变量时，$E[Y(0)|R=0]$ 甚至可能不可识别 e.

然而在实践中，我们可能并不清楚 Z 是否是 U足够好的代理变量，以使假设 3(a) 和 4 成立；但仍可能继续使用定理1中的识别公式 (4) 进行分析。也就是说，我们仍可通过求解观测积分方程 S5 来估计 q，再将其代入识别公式 (4)，得到$E[Y(0)|R=0]$的估计值。请注意，尽管S5 中的观测积分方程在总体层面没有解，但在有限样本中仍可能有解，例如，当W、 Z 和 X 均为分类变量，且 q 为饱和模型时。基于识别公式 (4) 得出的结果会收敛到什么？我们假设所有变量均为分类变量，并进行如下有限样本分析。

$$\begin{aligned}
E_n[I(R=1)\hat{q}(Z,X)Y]/\hat{P}(R=0) &= E_n[E_n[I(R=1)Y\hat{q}(Z,X)|W,X]]/\hat{P}(R=0) \\
&= E_n[E_n[Y\hat{q}(Z,X)|W,X,R=1]\hat{P}(R=1|W,X)]/\hat{P}(R=0) \\
&= E_n[E_n[Y\hat{q}(Z,X)|W,X,R=1]\hat{P}(R=1|W,X)]/\hat{P}(R=0) \\
&\approx E_n[E_n[Y|W,X,R=1]E_n[q(Z,X)|W,X,R=1]\hat{P}(R=1|W,X)]/\hat{P}(R=0) \\
&= E_n[E_n[Y|W,X,R=1]E_n[q(Z,X)|W,X,R=1]\hat{P}(R=1|W,X)]/\hat{P}(R=0) \\
&= E_n[E_n[Y|W,X,R=1]\hat{P}(R=0|W,X)]/\hat{P}(R=0)
\end{aligned}$$

$$= E_n[I(R=0)E_n[Y|W,X,R=1]]/\hat{P}(R=0)$$

其中，第四行的近似是因为在给定 X 的条件下， Z 与 Y 的独立性在有限样本中可能并不完全成立；第五个等号则是因为我们假设 q 在有限样本中精确满足观测积分方程。这一非严格的分析表明，$\hat{E}_n[I(R=1)\hat{q}(Z,\ X)Y]/P(R=0)$ 将收敛到 $E[I(R=0)E[Y|W,X,R=1]]/P(R=0)$。容易证明，除非 $E[Y(0)|W、\ X、\ R=1]=E[Y(0)|W,\ X,\ R=0]$ 几乎处处成立，否则 $E[I(R=0)E[Y|W,X,R=1]]/P(R=0) \neq E[Y(0)|R=0]$。

在有限样本中，根据上述推导，我们认为，平均而言，只要 $\hat{q}$ 能精确求解以下有限样本积分方程， $E_n[I(R=\hat{}\ \hat{}\ 1)\hat{q}(Z,\ X)Y]/P(R=0)$ 就可以近似 $E_n[I(R=0)E_n[Y|W,X,R=1]]/P(R=0)$：

$$\frac{\hat{P}(R=0|W,X)}{\hat{P}(R=1|W,X)} = E_n[q(Z,X)|W,X,R=1].$$

$\hat{E}_n[I(R=0)E_n[Y|W,X,R=1]]/P(R=0)$ 不是 $E[Y(0)|R=0]$ 的一致估计量，其偏差大小取决于 $E[Y(0)|W,\ X,\ R=1]$ 与 $E[Y(0)|W,\ X,\ R=0]$ 的接近程度，换句话说，取决于 W 作为 U的代理变量有多强。此外，由于 q 实际上无法识别，我们预计 $\hat{q}$ 只能捕捉噪声，并表现出较大的不稳定性，这使得 $\hat{}$ 代入估计量 $E_n[I(R=1)\hat{q}(Z,\ X)Y]/P(R=0)$ 不稳定。

为进一步确认这些观察结果，我们进行了模拟研究。数据按以下方式生成：

- $X \sim$ 二项分布$(n,0.5)$
- $U \sim$ 二项分布$(n,0.5)$
- $R \sim$ 二项分布$(n, 1/(1+\exp(-0.3+1.5U+0.5X)))$
- $W \sim$ 二项分布$(n,\exp(-w_u+w_uU-0.2X))$
- $Z \sim$ 二项分布$(n, 1/(1+\exp(-0.5+0.5X)))$

我们采用与主要模拟研究类似的方法模拟真实生存时间，但通过假设不存在删失来简化。平均生存时间 T^* 为

$R = 1$ 和 $R = 0$分别为 600 个月和 4000 个月。真实生存时间的风险比通过 $exp($ $-1 + 0.8X + 2U)$ 和 $exp(-1 + 0.5X + U)$ 模拟。我们通过 $Y = I(T^* > 12)$ 获得观测到的 1 年无 HIV 结局。我们考虑样本量为 1e4、1e5 和 1e6，并将 $w_u = 0.5$、2、8 作为变化取值（即， W分别为 U的弱、中等和强代理变量）。在此设定下，目标参数 $E[Y(0)$ $|R = 0] = 0.959$。下表总结了结果。

表 S6：在不同的 λ 和样本量 n下比较两种估计量。标准差由第97.5分位数减去第2.5分位数，再除以 2×1.96计算得出。5000次模拟。

		$E_n[I(R=1)\hat{q}(Z,\ X)Y]/\ \hat{P}(R=0)$		$E_n[I(R=0)E_n[Y\|W,X,R=1]]/\ \hat{P}(R=0)$	
w_u	N	中位数	SD	中位数	SD
0.5	10^4	0.977	0.249	0.976	0.003
	10^5	0.977	0.235	0.976	0.001
	10^6	0.977	0.284	0.976	0.001
2	10^4	0.972	0.543	0.969	0.004
	10^5	0.967	0.535	0.968	0.002
	10^6	0.970	0.556	0.968	0.001
8	10^4	0.964	0.627	0.963	0.005
	10^5	0.961	0.630	0.963	0.002
	10^6	0.962	0.726	0.963	0.002

ˆ 观察结果：（1）估计量 $E_n[I(R = 0)E_n[Y|W, X, R = 1]]/P(R = 0)$ 的偏差随 w_u 增大而减小，这与我们的分析一致。（2）估计量 $E_n[I(R = 1)\hat{q}(Z,\ X)Y]/P(R =0)$ 的估计值平均而言接近 $E_n[I(R = 0)E_n[Y|W, X, R = 1]]$ 得出的结果，但由于 $\hat{q}$ 不稳定，前者的方差要大得多。

接下来，我们进行类似的分析，假设 W 不是 U的代理变量。在此分析中，我们假设 W 满足

$$W \perp\!\!\!\perp (U, R, Z)|X.$$

当这一假设成立时，在总体层面，不再存在平方可积函数 h 能够求解观测到的积分方程

$$E[Y|Z,X,R=1]=E[h(W,X)|Z,X,R=1], \tag{S7}$$

因为右侧等于 $E[h(W,X)|X,R=1]$，且不随 Z 变化。因此，定理1中的识别公式(3)无效。若继续使用基于该识别公式的^代入估计量，并在有限样本中利用上述积分方程求解 h，那么我们可以证明

$$\begin{aligned}
E_n[\hat{h}(W,X)|R=0] &= E_n[E_n[\hat{h}(W,X)|X,R=0]|R=0] \\
&\approx E_n[E_n[\hat{h}(W,X)|X,R=1]|R=0] \\
&= E_n[E_n[E_n[\hat{h}(W,X)|Z,X,R=1]|X,R=1]|R=0] \\
&= E_n[E_n[E_n[Y(0)|Z,X,R=1]|X,R=1]|R=0] \\
&= E_n[E_n[Y(0)|X,R=1]|R=0]
\end{aligned}$$

除非 $E[Y(1)|X,R=1]=E[Y(1)|X,R=0]$，否则其不等于 $E[Y(0)|R=0]$。与上文的评论类似，这一推导实际上无效，因为在总体层面不存在能使 S7 成立的 h。然而，这一分析表明，在有限^样本中，如果存在能在有限样本中精确求解 *S7* 的 h，那么平均而言，我们预计^$E_n[h(W,X)|R=0]$ 等于 $E_n[E_n[Y|X,R=1]|R=0]$，且 $E_n[E_n[Y|X,R=1]|R=0]$ 对 $E[Y(0)|R=0]$ 存在偏差，其偏差大小取决于 $E[Y(1)|X,R=1]$ 与 $E[Y(1)|X,R=0]$ 的接近程度。

为确认这一点，我们进行一项模拟研究。数据生成如下：

- $X\sim$ 二项分布$(n,\ 0.5)$
- $U\sim$ 二项分布$(n,\ 0.5)$
- $R\sim$ 二项分布$(n,\ 1/(1+\exp(-0.3+1.5U+0.5X)))$
- $W\sim$ 二项分布$(n,\exp(-0.5-0.5X))$

- $Z \sim$ 二项分布$(n,\ 1/(1+\exp(-z_u/2+z_uU-0.2X)))$

结局的模拟方式与表 S6类似。我们考虑样本量为 1e4、1e5、1e6，并改变 $z_u = 0.5$、2、8 (即，Z分别是 U的弱、中等和强代理变量)。在此设定下，目标参数为 $E[Y(0)|R=0] = 0.959$。下表总结了结果。

表 S7：不同 λ 和样本量 n下两种估计量的比较。标准差通过将第 97.5 分位数减去第 2.5 分位数后除以 2×1.96计算得出。5000 次模拟。

		$E_n[I(R=1)\hat{q}(Z,X)Y]/\hat{P}(R=0)$		$E_n[I(R=0)E_n[Y\|W,X,R=1]]/\hat{P}(R=0)$	
w_u	N	中位数	SD	中位数	SD
0.5	10^4	0.977	0.249	0.976	0.003
	10^5	0.977	0.235	0.976	0.001
	10^6	0.977	0.284	0.976	0.001
2	10^4	0.972	0.543	0.969	0.004
	10^5	0.967	0.535	0.968	0.002
	10^6	0.970	0.556	0.968	0.001
8	10^4	0.964	0.627	0.963	0.005
	10^5	0.961	0.630	0.963	0.002
	10^6	0.962	0.726	0.963	0.002

观察结果：(1) 估计量 $E_n[E_n[Y(0)|X,\ R=1]|R=0]$ 的偏差不会随 z_u 改变，这与我们的分析一致。(2) 估计量 $E_n[h(W,\ X)|R=0]$提供的估计值平均而言接近 $E_n[E_n[Y(0)|X,\ R=1]|R=0]$的估计值，但前者的方差要大得多，这是由于 h 不稳定。

这些分析凸显了在实践中评估代理变量关联强度的重要性。值得注意的是，与工具变量分析不同——当工具变量较弱时，估计量往往不稳定——近端因果推断在代理变量较弱时仍可能保持稳定，尤其是在小样本中。然而，若不评估

代理变量的强度，应用近端方法的使用者可能会被看似精确、实则有偏、且随样本量增大而愈发不稳定的估计值所误导。

S5 当 h 和 q 精确求解观测积分方程时

ˆ当 W、Z、X均为分类变量（如我们的应用中所示）时，h和ˆq可能会精确求解观测积分方程，尤其是当 h 和 q 由某些饱和模型估计时。也就是说，我们有

$$\frac{\hat{P}(R=0|W,X)}{\hat{P}(R=1|W,X)} = E_n[q(Z,X)|W,X,R=1]$$

and

$$E_n[Y|Z,X,R=1] = E_n[\hat{h}(W,X)|Z,X,R=1].$$

在这种情况下，我们观察到，使用定理1中公式（3）和（4）的识别公式得到的代入估计量是一致的。以下推导可以证明这一点。

$$\begin{aligned}
E[I(R=1)Yq(Z,X)] &= E[q(Z,X)E[I(R=1)Y|Z,X]] \\
&= E[q(Z,X)E[Y|Z,X,R=1]P(R=1|Z,X)] \\
&= E[q(Z,X,)P(R=1|Z,X)E[h(W,X)|Z,X,R=1]] \\
&= E[q(Z,X)E[I(R=1)h(W,X)|Z,X]] \\
&= E[I(R=1)h(W,X)q(Z,X)] \\
&= E[h(W,X)E[I(R=1)q(Z,X)|W,X]] \\
&= E[h(W,X)P(R=1|W,X)E[q(Z,X)|W,X,R=1]] \\
&= E[h(W,X)P(R=0|W,X)] \\
&= E[h(W,X)I(R=0)]
\end{aligned}$$

通过将总体期望 E 替换为样本期望，该推导在有限样本中仍然成立。

S6 二元 U 参数两阶段回归

我们从以下模型开始

$$\lambda(t|R,Z,X,U) = \lambda_0(t)exp(\beta_{Tx}X + \beta_{Tu}U)$$

其中 $U \in \{0,1\}$ 为二元变量。

当事件率较低时，我们有 $S(t|R,\ U,\ X) \approx 1$。

Then, the density function $f(t|R,U,X) \approx \lambda(t|R,U,X)$

注意，

$$\begin{aligned} f(t|R,Z,X) &= E[f(t|R,U,X,Z)|R,Z,X] \\ &= E[f(t|U,X)|R,Z,X] \\ &\approx E[\lambda(t|U,X)|R,Z,X] \\ &= \lambda_0(t)exp(\beta_{Tx}X)E[exp(\beta_{Tu}U)|R,Z,X] \end{aligned}$$

由于 U 是二元变量，我们知道

$$\begin{aligned} E[exp(\beta_{T_u}U)|R,Z,X] &= exp(\beta_{T_u})P(U=1|R,Z,X) + P(U=0|R,Z,X) \\ &= (exp(\beta_{T_u}) - 1)P(U=1|R,Z,X) + 1 \end{aligned}$$

它重新 主要用于识别 $P(U=1|R,Z,X)$。假设二元 W 服从对数线性模型 ：

$$P(W=1|R,Z,U,X) = exp(\beta_{W0} + \beta_{Wx}X + \beta_{Wu}U)$$

然后，

$$\begin{aligned} P(W=1|R,Z,X) &= P(W=1|R,Z,X,U=1)P(U=1|R,Z,X) + P(W=1|R,Z,X,U=0)P(U=0|R,Z,X) \\ &= exp(\beta_{W0} + \beta_{Wx}X + \beta_{Wu})P(U=1|R,Z,X) + exp(\beta_{W0} + \beta_{Wx}X)(1 - P(U=1|R,Z,X)) \\ &= exp(\beta_{W0} + \beta_{Wx}X)\Big[(exp(\beta_{Wu}) - 1)P(U=1|R,Z,X) + 1\Big] \end{aligned}$$

这意味着

$$P(U=1|R,Z,X) = \left\{\frac{P(W=1|R,Z,X)}{exp(\beta_{W0} + \beta_{Wx}X)} - 1\right\}\frac{1}{(exp(\beta_{Wu}) - 1)}$$

然后，我们有

$$
\begin{aligned}
E[exp(\beta_{T_u}U)|R,Z,X] &= (exp(\beta_{T_u}) - 1)P(U=1|R,Z,X) + 1 \\
&= \left\{ \frac{P(W=1|R,Z,X)}{exp(\beta_{W0}+\beta_{Wx}X)} - 1 \right\} \frac{exp(\beta_{T_u})-1}{exp(\beta_{Wu})-1} + 1 \\
&= \frac{P(W=1|R,Z,X)}{exp(\beta_{W0}+\beta_{Wx}X)} \frac{exp(\beta_{T_u})-1}{exp(\beta_{Wu})-1} - \frac{exp(\beta_{T_u})-1}{exp(\beta_{Wu})-1} + 1 \\
&= \frac{P(W=1|R,Z,X)}{exp(\beta_{W0}+\beta_{Wx}X)} \frac{exp(\beta_{T_u})-1}{exp(\beta_{Wu})-1} + \frac{exp(\beta_{Wu})-exp(\beta_{T_u})}{exp(\beta_{Wu})-1}
\end{aligned}
$$

然后，

$$
\begin{aligned}
f(t|R,Z,X) &\approx \lambda_0(t)exp(\beta_{Tx}X)E[exp(\beta_{Tu}U)|R,Z,X] \\
&= \lambda_0(t)exp(\beta_{Tx}X)\left(\frac{P(W=1|R,Z,X)}{exp(\beta_{W0}+\beta_{Wx}X)} \frac{exp(\beta_{T_u})-1}{exp(\beta_{Wu})-1} + \frac{exp(\beta_{Wu})-exp(\beta_{T_u})}{exp(\beta_{Wu})-1}\right) \\
&= \lambda_0(t)exp(\beta_{Tx}X)\left(\alpha \frac{P(W=1|R,Z,X)}{exp(\beta_{W0}+\beta_{Wx}X)} + 1 - \alpha\right)
\end{aligned}
$$

其中 $\alpha = \frac{exp(\beta_{T_u} - 1)}{exp\,\beta_{Wu} - 1}$ ()。

S7 模拟

我们通过一项模拟研究考察半参数逆概率删失加权方法和参数两阶段方法的有限样本性能。

我们首先根据以下分布模拟两个二元协变量： $X_1 \sim$ Binom(n, p =0.7) 和 $X_2 \sim$ Binom($n, p = 0.5$)。我们模拟一个连续的 U ，其中 U 服从截断正态分布， $E(U) = 0.5$, $Var(U) = 1$，且 $U \in [0,1]$。然后，我们生成一个二元 $U_b = I(U > 0.5)$。二元负对照结局 W 服从对数线性模型，其中 $P(W = 1) = exp(\beta_{W0} + \beta_W U + 0.2X_1 + 0.5X_2)$。我们模拟二元 R 和 Z ，其中 $P(R = 1) = 1/(1 + exp(-1.2 + 2U_b + 0.3X_1 + 0.2X_2))$, 且$P(Z = 1) = 1/(1 + exp(\beta_{Z0} + \beta_Z U_b + 0.5X_1 - 0.2X_2))$。我们模拟真实生存时间和删失时间， $T^*(0)$、 $T^*(1)$、 $C(0)$ 和 $C(1)$；对于分别具有 $R = 1$ 和 $R = 0$ 的数据，(平均生存时间，平均删失时间) 分别为 (2000, 30) 天和 (6000, 50) 天。真实生存时间的风险比通过 $exp(-1 + 0.8X_1 + 0.3X_2 - 2U)$ 模拟。

对于具有 $R=1$ 和 $R=0$的数据，则为 $exp(-1+0.5X_1-0.2X_2-1.8U)$。类似地，对于分别具有 $R=1$ 和 $R=0$ 的数据，真实删失时间的风险比分别通过 $exp(-2.5+0.2Z+0.2X_1-0.1X_2)$ 和$exp(-2.5+0.2Z+0.3X_1+0.1X_2)$ 模拟。观测时间为 $T=R\cdot\min(T^*(1)\ C(1))+(1-R)\cdot\min(T^*(0)\ C(0))$，事件指示变量为 $\Delta=I(T^*\leq C)$。我们的估计目标是反事实的1年无人类免疫缺陷病毒生存概率 $P(T^*(0)>365\mid R=0)$（等价于发生率的补数）。请注意，这一数据生成过程满足第 2.2.1 节和第 2.2.2 节中的所有假设。随后，我们可以用以下析因设计概述模拟研究。

因子 1（样本量）：(1) $n=6500$；(2) $n=10000$；

因子 2（负对照变量作为未测量混杂因素代理变量的强度 W，Z；{(β_{W0}, β_W), (β_{Z0}, β_Z)}：(1) {medium, medium}：{$(-4,\ 3)$，$(-1.2,\ 1)$}；(2) {medium, high}：{$(-4,\ 3)$，$(-1.2,\ 2)$}；(3) {high, medium}：{$(-4.7,\ 4)$，$(-1.2,\ 1)$}；(4) {high, high}：{$(-4.7,\ 4)$，$(-1.2,\ 2)$}。

因子 3（安慰剂组下反事实HIV发病率的估计方法）：(1) 调整（X，U）的预言估计量；(2) 调整 X 的朴素方法；(3) 调整（X，Z，W）的朴素方法；(4) 使用结局桥接函数、倾向评分桥接函数和双重稳健估计量的半参数逆概率删失加权；(5) 基于回归的两阶段方法。

在所有数据生成过程中，我们有 $P(R=1)\approx 0.33$，$P(W=1)\approx 0.15$，$P(Z=1)\approx 0.50$，这些与我们的真实数据示例相似。所有设置下，1 年内真实的反事实 HIV 累积发生率为 0.035 。与真实数据分析一致，所有方法的估计和置信区间均使用 `gmm` 包在 R 中获得。我们还按照第 3.2 节所述，考察了我们的统计推断方法的功效和 I 类错误率。

结果见表 S8 和表 S9 ，其中包含 3000 次模拟运行。我们总结了：（i）在$\log(-\log(\cdot))$尺度上，截至 1 年的反事实 HIV 累积发病率估计值的均值；（ii）截至 1 年的转换后反事实 HIV累积发病率；（iii）反事实 HIV累积发病率估计值的标准差

在 $\log(-\log(\cdot))$ 尺度上截至 1 年的反事实 HIV累积发病率；（iv）标准误的中位数；（v）95% 置信区间的经验覆盖率；（vi）在备择假设下估计 $P(T^*(1) > 365 \mid R = 0) - P(T^*(0) > 365 \mid R = 0)$ 时的经验功效；（vii）在零假设下估计 $P(T^*(1) > 365 \mid R = 0) - P(T^*(0) > 365 \mid R = 0)$ 时的经验 I 类错误率。逆概率删失加权估计可能超出 $[0, 100]$ 估计累积发生率的合理范围，这被视为表 3中 IPCW 方法的一项局限性。IPCW 方法报告的结果来自处于 $[0, 100]$ 范围内的模拟运行，我们在最后一列报告了被排除的无效估计值所占的比例。

观察：

1. 所有逆概率删失加权估计和两阶段估计值均无偏，而朴素回归得到的估计值存在偏差。在朴素回归中额外校正 Z, W 可减小偏差。

2. 与逆概率删失加权估计相比，基于回归的两阶段估计值的标准差和标准误平均更小。

3. 随着样本量增加，我们观察到逆概率删失加权方法和两阶段回归方法的标准差均更小、功效均更高。

4. 在所有设定下，逆概率删失加权和两阶段回归的经验覆盖率均近似达到名义 95%，且第一类错误控制在 0.05。

5. 使用作为未测量混杂因素更强代理变量的负对照变量 W, Z 时，所有逆概率删失加权估计和两阶段回归估计值的标准差平均更小。

表 S8：因子 1 的模拟结果： $n = 6500$。

因子 2	方法	设定	估计值	估计发生率	SD	中位数标准误	覆盖率	功效	第一类错误	Prop_ex
中等 W，中等 Z	预言机	调整 X、U	1.212	0.035	0.034	0.034	0.944	1	0.061	-
	朴素	调整 X	1.084	0.052	0.025	0.025	0.002	1	0.938	-
	朴素	调整 X、Z、W	1.101	0.049	0.026	0.026	0.019	1	0.894	-
	IPCW	结局桥接函数	1.207	0.035	0.208	0.205	0.954	0.515	0.047	0.119
	IPCW	倾向评分桥接函数	1.206	0.035	0.206	0.205	0.953	0.518	0.048	0.119
	IPCW	双重稳健	1.207	0.035	0.208	0.205	0.954	0.516	0.047	0.119
	两阶段	-	1.204	0.036	0.128	0.121	0.973	0.492	0.003	-
中等 W，高 Z	预言机	调整 X、U	1.212	0.035	0.033	0.033	0.942	1	0.061	-
	朴素	调整 X	1.086	0.052	0.025	0.025	0.003	1	0.938	-
	朴素	调整 X、Z、W	1.112	0.048	0.027	0.027	0.054	1	0.844	-
	IPCW	结局桥接函数	1.229	0.033	0.143	0.118	0.958	0.726	0.043	0.013
	IPCW	倾向评分桥接函数	1.229	0.033	0.142	0.117	0.958	0.726	0.045	0.012
	IPCW	双重稳健	1.229	0.033	0.143	0.118	0.958	0.726	0.045	0.013
	两阶段	-	1.187	0.038	0.061	0.059	0.940	0.887	0.013	-
高 W、中等 Z	预言机	调整 X、U	1.212	0.035	0.034	0.034	0.944	1	0.061	-
	朴素	调整 X	1.084	0.052	0.025	0.025	0.002	1	0.938	-
	朴素	调整 X、Z、W	1.109	0.048	0.026	0.026	0.036	1	0.842	-
	IPCW	结局桥接函数	1.214	0.034	0.213	0.204	0.952	0.509	0.048	0.106
	IPCW	倾向评分桥接函数	1.216	0.034	0.215	0.203	0.952	0.508	0.050	0.104
	IPCW	双重稳健	1.214	0.034	0.213	0.204	0.952	0.508	0.051	0.106
	两阶段	-	1.185	0.038	0.060	0.122	0.969	0.473	0.002	-
高 W、高 Z	预言机	调整 X、U	1.212	0.035	0.033	0.033	0.942	1	0.061	-
	朴素	调整 X	1.086	0.052	0.025	0.025	0.003	1	0.938	-
	朴素	调整 X、Z、W	1.119	0.047	0.027	0.027	0.087	1	0.789	-
	IPCW	结局桥接函数	1.227	0.033	0.136	0.117	0.962	0.742	0.037	0.010
	IPCW	倾向评分桥接函数	1.228	0.033	0.137	0.117	0.962	0.741	0.041	0.010
	IPCW	双重稳健	1.227	0.033	0.136	0.117	0.962	0.741	0.041	0.010
	两阶段	-	1.196	0.037	0.101	0.060	0.942	0.870	0.012	-

逆概率删失加权 $=$ 逆概率删失加权； $X =$ 基线协变量； $Z =$ 负对照暴露； $W =$NCO。

表 S9：因子 1 的模拟结果： $n = 10000$。

因子 2	方法	设定	估计值	估计发生率	SD	中位数标准误	覆盖率	功效	第一类错误	Prop_ex
中等 W，中等 Z	预言机	调整 X、U	1.212	0.035	0.027	0.027	0.948	1	0.040	-
	朴素	调整 X	1.084	0.052	0.020	0.020	0.000	1	0.995	-
	朴素	调整 X、Z、W	1.101	0.049	0.021	0.021	0.002	1.000	0.981	-
	IPCW	结局桥接函数	1.221	0.034	0.182	0.167	0.949	0.579	0.052	0.068
	IPCW	倾向评分桥接函数	1.221	0.034	0.193	0.166	0.949	0.580	0.054	0.067
	IPCW	双重稳健	1.221	0.034	0.182	0.167	0.948	0.578	0.054	0.068
	两阶段	-	1.185	0.038	0.050	0.096	0.963	0.625	0.003	-
中等 W，高 Z	预言机	调整 X、U	1.212	0.035	0.027	0.027	0.951	1	0.048	-
	朴素	调整 X	1.084	0.052	0.020	0.020	0.000	1	0.998	-
	朴素	调整 X、Z、W	1.113	0.048	0.021	0.022	0.007	1	0.956	-
	IPCW	结局桥接函数	1.221	0.034	0.107	0.095	0.958	0.847	0.037	0.002
	IPCW	倾向评分桥接函数	1.220	0.034	0.107	0.095	0.958	0.848	0.040	0.002
	IPCW	双重稳健	1.221	0.034	0.107	0.095	0.958	0.847	0.040	0.002
	两阶段	-	1.199	0.036	0.099	0.048	0.936	0.938	0.017	-
高 W、中等 Z	预言机	调整 X、U	1.212	0.035	0.027	0.027	0.948	1	0.040	-
	朴素	调整 X	1.084	0.052	0.020	0.020	0.000	1	0.995	-
	朴素	调整 X、Z、W	1.109	0.048	0.021	0.021	0.003	1.000	0.960	-
	IPCW	结局桥接函数	1.227	0.033	0.187	0.168	0.951	0.570	0.045	0.056
	IPCW	倾向评分桥接函数	1.227	0.033	0.187	0.168	0.951	0.571	0.051	0.056
	IPCW	双重稳健	1.227	0.033	0.187	0.168	0.951	0.570	0.051	0.056
	两阶段	-	1.199	0.036	0.099	0.098	0.961	0.604	0.004	-
高 W、高 Z	预言机	调整 X、U	1.212	0.035	0.027	0.027	0.951	1	0.048	-
	朴素	调整 X	1.084	0.052	0.020	0.020	0.000	1	0.998	-
	朴素	调整 X、Z、W	1.119	0.047	0.021	0.022	0.020	1	0.935	-
	IPCW	结局桥接函数	1.220	0.034	0.104	0.094	0.962	0.852	0.036	0.002
	IPCW	倾向评分桥接函数	1.220	0.034	0.105	0.094	0.962	0.854	0.040	0.001
	IPCW	双重稳健	1.220	0.034	0.104	0.094	0.962	0.852	0.040	0.002
	两阶段	-	1.184	0.038	0.049	0.047	0.938	0.913	0.019	-

逆概率删失加权 $=$ 逆概率删失加权； $X =$ 基线协变量； $Z =$ 负对照暴露； $W =$NCO。