



【专家介绍】王伟炳,教授,博士生导师。复旦大学公共卫生学院流行病学教研室主任。兼任上海市重大传染病和生物安全研究院新发传染病预警预测中心主任,教育部公共安全重点实验室副主任。中华医学会公共卫生学分会现场流行病学组副组长、上海市预防医学会流行病学分会主委,中华预防医学会流行病学分会常委,教育部全国学校结核病防控专家组专家。长期从事流行病学教学与科研工作,培养研究生30余名,作为第一作者或通讯作者发表论文200余篇,其中SCI收录100余篇。主要研究方向为传染病疾病负担的估计、传染病分子流行病学、临床预后及疫苗效果评价。承担一系列国际合作和国家级项目,包括盖茨基金会捐赠项目、世界卫生组织项目、瑞典国家科学院项目、科技部重大专项子项、国家自然科学基金、上海市市级科技重大专项等。作为课题第一完成人获上海市科技进步奖二等奖(2023年)。

观察性研究中倾向性评分匹配的方式和应用

吴晔¹ 田婕^{1,2} 王伟炳^{1,2△}

¹复旦大学公共卫生学院流行病学教研室-公共卫生安全教育部重点实验室 上海 200032;

²复旦大学上海市重大传染病和生物安全研究院 上海 200032)

【摘要】 观察性研究是获取真实世界证据的重要途径,但由于缺乏随机化分组,不同组间的特征差异使结果易受混杂偏倚的影响。传统的分层分析和多因素回归在协变量较多时控制混杂效果有限。倾向性评分匹配通过估计个体接受干预的倾向性评分进行匹配,将倾向值相近的个体在处理组和对照组间进行配对,从而平衡协变量分布、降低混杂干扰。近年来,倾向性评分匹配已在慢性病管理与药物效果评价、卫生政策和健康服务评估、疫苗效果人群差异研究和远程医疗评价等领域广泛应用。本文简述倾向性评分在观察性研究中的常见匹配方式和应用场景。

【关键词】 倾向性评分匹配法(PSM); 观察性研究; 混杂因素

【中图分类号】 R311 **【文献标志码】** A **doi:**10.3969/j.issn.1672-8467.2025.06.020

Propensity score matching in observational studies: methods and applications

WU Ye¹, TIAN Jie^{1,2}, WANG Wei-bing^{1,2△}

¹Department of Epidemiology, School of Public Health, Fudan University-Key Laboratory of Public Health Safety, Ministry of Education, Shanghai 200032, China; ²Shanghai Institute of Infectious Disease and Biosecurity, Fudan University, Shanghai 200032, China)

【Abstract】 Observational studies are important approaches for obtaining real-world evidence. However, due to the lack of randomized allocation, differences in baseline characteristics between groups often introduce confounding bias, which may distort the results. Traditional stratified analyses and multivariable regression models have limited ability to control confounding when multiple covariates are involved.

上海市市级科技重大专项(ZD2021CY001);上海市加强公共卫生体系建设三年行动计划(2023—2025年)重点学科项目(GWVI-11.1-03)

[△]Corresponding author E-mail: wwb@fudan.edu.cn

网络首发时间:2025-10-31 14:48:18 网络首发地址:<https://link.cnki.net/urlid/31.1885.R.20251030.1428.002>

Propensity score matching estimates the probability of receiving an intervention based on observed covariates, then matches individuals with similar propensity scores between treatment and control groups, thereby balancing covariate distributions and reducing confounding. In recent years, propensity score matching has been widely applied in various fields, including chronic disease management and drug effectiveness evaluation, public health policy and health service assessment, vaccine effectiveness studies focusing on population disparities, and evaluations of telemedicine interventions. This review summarizes common matching methods and application scenarios of propensity score matching in observational studies.

【Key words】 propensity score matching (PSM); observational studies; confounding factors

* This work was supported by the Shanghai Municipal Science and Technology Major Project (ZD2021CY001) and the Key Disciplines Project of Shanghai Municipal Three-year Action Plan (2023–2025) for Strengthening Public Health System Construction (GWVI-11.1-03).

观察性研究在公共卫生领域具有重要地位。相较于高成本的随机对照试验(randomized controlled trial, RCT),观察性研究纳入标准相对宽松,能更好反映真实世界的人群特征,结论外推性更好^[1],特别在伦理或实践条件下难以实施RCT的情境中,观察性研究往往成为获取循证依据的主要手段。然而由于缺乏随机化分组,观察性研究常面临混杂偏倚的影响,不同组间的基线特征(如人口学特征、疾病史、健康行为等)的分布可能存在系统差异,导致对分组因素和结局的分析受潜在混杂因素的干扰^[2]。传统的分层分析和多因素回归虽能在一定程度上控制混杂,但当协变量数量过多或组间差异较大时,其调整能力有限,残余混杂的影响难以完全消除^[3]。倾向性评分匹配(propensity score matching, PSM)是一种近年来广泛应用于观察性研究的统计方法,其主要思想是通过倾向性评分在不同组间进行匹配,使两组的可观测协变量分布达到近似平衡,减少混杂偏倚,使处理效应的估计更接近无偏^[4]。本文将对目前观察性研究中PSM的常用方式与具体应用进行阐述。

基本概念 倾向性评分(propensity score)由Rosenbaum和Rubin在1983年提出^[5]。对于每一研究个体,倾向性评分 $e(x_i)$ 的定义为该个体 i 在观测协变量为 x 时分组到处理组($Z_i=1$)的条件概率,即 $e(x_i) = Pr(Z_i=1|X_i=x_i)$,若分组变量 Z_i 与给定的观测协变量 x_i 相互独立,则对于每一个观测协变量 x_i ,其分配到某一组的概率可表示为:

$$Pr(z_1, \dots, z_n | x_1, \dots, x_n) = \prod_{i=1}^n [e(x_i)^{z_i}][1 - e(x_i)]^{1-z_i}$$

倾向性评分将多维协变量信息转换成表示个

体接受处理的倾向强弱的概率函数,不同组间倾向性评分相近的个体可视为协变量基本均衡。因此,将不同组间倾向性评分相近的个体进行匹配,可以平衡处理组和对照组的多个协变量,使效应估计更接近无偏。

PSM的应用基于若干前提假设^[5]。其关键假设是:(1)条件独立假设。在控制了所有已观测的协变量 X 后,处理分配 T 与潜在结果 Y 独立。即所有重要的混杂因素都必须已被测量并包含在模型 X 中。(2)共同支撑或重叠性。处理组与对照组的倾向性评分分布应具有足够的重叠区域,使每个处理个体都能在对照组中找到可比对象。若倾向性评分的分布重叠不足,大量样本因无匹配对象被排除,会降低效应估计的外部有效性。

PSM的实现流程 PSM的实现包括4个核心步骤:(1)数据预处理与协变量选择;(2)倾向性评分模型估计;(3)匹配策略与执行;(4)匹配质量评估与结果检验。第二步与第三步的模型选择与匹配设计是PSM实现的关键步骤,可直接影响匹配结果的平衡性与分析的有效性。

协变量选择与倾向评分模型估计 PSM协变量的选择一般遵循“必要充分原则”,尽可能纳入所有影响分组和结局的已观测混杂因素,常见协变量包括人口学特征、社会经济水平、疾病严重程度等。对于变量较多的研究,可采用逐步筛选法、LASSO法或专家共识法确定最终协变量集。

倾向性评分模型估计旨在估计个体接受处理的概率估计值。最常用方法为逻辑回归模型:

$$\text{logit}(P(T=1|X)) = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n$$

假定处理概率与协变量呈线性对数关系,结果

易于解释。近年来,研究者也常采用广义加性模型、随机森林等机器学习方法,提高非线性和高维数据下估计的精度与稳健性^[6]。模型构建后,可通过ROC曲线、AUC值等检验拟合优度,并比较处理组与对照组倾向值分布的重叠情况。

匹配策略选择 目前国内外研究常见的PSM策略主要分为贪婪匹配(greedy matching)和最优匹配(optimal matching)两大类^[7]。贪婪匹配按顺序逐个匹配处理组个体,不保证全局最佳但计算快速、实现简单,包括最近邻匹配(nearest neighbor matching)、卡钳匹配(caliper matching)等;最优匹配则通过整体优化算法,使匹配对的距离总和或不平衡度达到全局最优,计算量较大但匹配平衡性更佳,如最优完全匹配(optimal full matching)。另外,部分方法如马氏距离匹配法(Mahalanobis metric matching)和核匹配(kernel matching)并不完全归属于这两类,而是基于不同的距离度量或加权原则构建,常作为贪婪匹配和最优匹配的补充,用于提高匹配精度与样本利用率。

最近邻匹配 最近邻匹配法是目前最常用的PSM方法。对于每个处理组个体,该方法从对照组中选择倾向性评分最接近者作为匹配对,可分为不放回(每个对照只能用一次)或有放回(允许一个对照个体匹配多个处理个体)两种形式。该方法简单直观,便于实现和结果解释。缺点是当处理组与对照组倾向值差异较大时,可能造成“不良匹配”引入剩余混杂偏倚^[8]。此外,不放回匹配的结果可能受处理组匹配顺序影响,有放回匹配虽缓解了此问题,但某些高相似度的对照易被重复使用,增加后续分析的复杂性。

卡钳匹配法 卡钳匹配又称半径匹配(radius matching),指在最邻近匹配的基础上,设定一个最大允许的倾向性评分差异,即卡钳宽度(caliper)或允许半径(radius),常用推荐阈值为倾向性评分标准差的0.2~0.25倍。仅当处理组个体与对照组个体的倾向性评分差在该阈值以内才允许进行匹配,否则视为无法匹配而舍弃^[9]。卡钳匹配能有效避免倾向性评分相差过大,降低残余混杂。当处理组和对照组倾向性评分存在长尾分布或重叠不多时,卡钳匹配可提高匹配的精度,但可能导致样本损失,影响结果代表性。

马氏距离匹配法 马氏距离由印度统计学家P.C.Mahalanobis提出,考虑各协变量之间的相关

性,是一种衡量各观测点之间标准化协方差距离的统计量^[10]。通过计算处理组与对照组的马氏距离,选择马氏距离最接近者进行匹配,常与卡钳匹配法结合使用,即先限定卡钳范围,再在范围内选择马氏距离最小者。该方法的优点是兼顾了倾向性评分之外的协变量平衡,缺点在于计算复杂,结果不易解释,且需假定协变量的分布情况,在多分类变量的场景应用受限。

核匹配 核匹配是一种基于加权的非参数匹配方法,通过核函数计算所有对照个体的权重,倾向性评分距离越小的权重越大,距离越大则权重越小,形成一个加权平均的虚拟对照组合值^[11]。核匹配能充分利用所有样本信息,避免单个匹配产生的随机误差,常用于大样本研究。局限在于纳入远距离样本可能引入残余偏倚,此外核匹配须考虑核函数和带宽,参数不当会影响结果的平衡性。

最优完全匹配 最优完全匹配的基本思想是整体最优化,将所有样本分配到一系列匹配集中,每个匹配集内包含至少一名处理个体和一名对照个体,并通过最优化过程使各匹配组的全局距离总和最小化^[12]。该方法的优点是充分利用所有样本,在全局上达到最佳平衡,适用于样本量较大且处理组与对照组倾向性评分分布重叠良好的情况。缺点是算法复杂,计算量大且结果分析相对复杂,需要对不同匹配集中的个体权重进行计算。

匹配质量评估与软件实现 匹配完成后还需要对协变量进行平衡性诊断,常用指标如标准化平均差(standardized mean differences, SMDs),可反映处理组与对照组各协变量差异程度,匹配后各变量SMD<0.1通常视为平衡良好;倾向性分布重叠率和核密度图可用于直观比较两组倾向性分布的重叠程度,判断匹配样本在倾向空间的可比性。

PSM目前已广泛集成于主流统计软件中。R语言中常用包如MatchIt、optmatch、matching,可实现从倾向性评分估计、匹配执行到平衡性可视化的全过程;Stata软件可通过psmatch2与teffectspsmatch等命令进行匹配与加权分析;Python中可利用psmpy和CausalML等库进行自动化匹配和可视化。这些工具的应用显著提升了PSM的可操作性,使其在公共卫生、临床流行病学及健康政策研究中进一步推广。倾向性评分匹配的整体流程见图1。

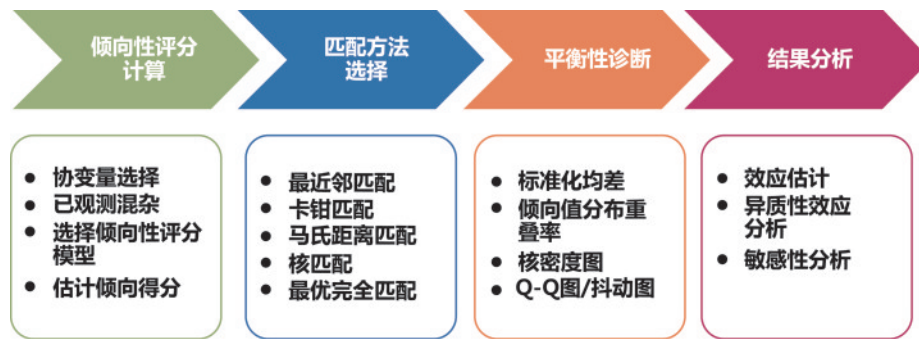


图1 倾向性评分匹配流程图

Fig 1 Flowchart of propensity score matching

PSM在观察性研究中的应用

慢性病管理与药物效果评价 尽管RCT是临床治疗和药物效果评价的金标准,慢性病管理措施和药物治疗方案的效果评价在真实世界中大量依赖观察性研究数据^[13]。由于医师常基于患者病情严重程度、共病状况等因素选择不同治疗管理方案,不同处理组患者的结局可能受基线风险差异的影响^[14]。PSM通过在研究设计阶段计算每位患者接收某治疗的倾向概率并进行匹配,可减轻混杂偏倚并提高观察性研究的内部有效性。

一项奥地利的队列研究评估了糖尿病管理计划的长期有效性,通过最近邻匹配法对年龄、性别、经济水平、糖尿病病程等基线特征进行匹配,结果显示PSM控制混杂因素后,糖尿病管理项目仍与患者生存质量改善和随访期内的医疗成本降低相关^[15]。美国一项研究分析了联合血小板单抗与单独使用抗凝剂的出血风险差异,基于年龄、既往出血史、肝肾功能、消化系统疾病等变量构建倾向性评分并进行匹配,减少了因患者病情差异导致的混杂偏倚,使结果更加稳健^[16]。Myles等^[17]评估了术中单次使用地塞米松对患者慢性术后疼痛的长期风险,采用PSM和逆概率加权,平衡了不同组间患者手术类型、术前状态、用药情况等方面的差异,降低了已知混杂因素对术后疼痛结局的干扰。

公共卫生政策与健康服务评价 卫生政策和健康服务在实施时难以实现随机化,政策常在特定地区或人群中推广,与未实施人群在经济社会和健康状况上存在系统差异,例如评估医疗保险覆盖扩大政策时,参保者往往收入更高、健康状况更好^[18]。PSM可构建可比对照组,平衡已知混杂的干扰,还可结合敏感性分析评估未观测混杂的影响,进一步增强效应的可信度。

美国一项研究分析了孕产妇导乐护理服务对不同州产妇剖腹产率和产后抑郁情况的影响,采用1:1最近邻匹配对导乐护理服务组和对照组进行匹配,以平衡母亲年龄、种族、社会经济状态、医院类型等混杂因素。研究通过PSM减少了由于个体因素(如经济能力)带来的选择偏倚,使组间的结果差异更能反映护理服务本身的真实效果^[19]。Bezie等^[20]在撒哈拉以南的非洲地区使用PSM评估产前保健与医疗机构分娩政策对新生儿死亡率的影响。研究采用倾向评分计算不同基线特征孕妇接受产前保健和医疗机构的可能性,并在对照组中进行卡钳匹配。结果显示,产前护理与医疗设施分娩与新生儿死亡率显著下降相关,政策干预的效应在控制多种基线特征后依然显著。进一步敏感性分析显示,结果对未测混杂的影响不敏感,表明该研究结论具有稳健性。以上两项研究均通过PSM控制了经济水平和医疗资源可及性等混杂因素,使比较组在可测协变量上趋于可比,为公共卫生政策效果评估和决策提供了更可靠的科学依据。

疫苗效果评估与人群差异研究 疫苗效果评价常利用人群队列、医疗数据库和公共卫生监测数据评估疫苗在真实世界中的有效性和潜在偏倚。疫苗观察性研究中,自愿接种疫苗的人往往具有更高的健康意识和健康基础状况,这种“健康志愿者”偏倚可能导致疫苗效果被高估^[21]。PSM可根据人口统计学特征和共病情况等协变量进行匹配,平衡已知混杂因素,从而更可靠地评估疫苗对不同群体健康结果的影响。

美国一项研究采用PSM方法评估类风湿关节炎患者接种重组带状疱疹疫苗的有效性,匹配了社会人口学特征、生活方式、医疗服务利用、共病和用药等协变量,结果显示接种疫苗组带状疱疹发病风

险和全因死亡率显著降低,该保护效果在50~64岁年龄组、女性以及白人和非裔患者中更为突出^[22]。法国学者利用PSM评估65岁以上人群中高剂量四价流感疫苗的效果,将高剂量组和标准剂量组按年龄、性别、接种时间和地区进行匹配,发现高剂量疫苗组流感相关住院率较标准剂量组降低约3.3%,但非流感住院率反而更高,提示可能存在剩余混杂,高剂量疫苗或许应优先接种更体弱多病的老年人^[23]。虽然PSM可能存在剩余混杂或选择偏好,但能有效减少已知混杂的干扰,为评估疫苗对脆弱群体的健康效益和免疫策略制定提供重要参考。

远程医疗与数字健康干预评价 远程医疗与数字健康干预在近年来快速普及,成为提升医疗可及性的重要途径^[24]。与传统临床干预不同,患者是否接受远程医疗服务常受个人偏好、地理条件、电子设备使用能力等因素影响,且地区间应用差异较大,实践中常难以实现随机化分组。PSM为此类研究提供了可行的准实验方法,通过在观察性数据中构建平衡对照组,平衡两组间已知混杂因素,近似模拟随机化过程。一项多国前瞻性队列研究利用

PSM评估远程医疗在术后伤口评估中的应用效果^[25]。远程组与常规随访组按手术类型、年龄、基线感染风险等变量进行匹配。匹配前远程组患者总体更年轻且手术创伤更轻,匹配后各协变量SMD<0.1,提示协变量平衡较好。两组在术后伤口感染检出率上无显著差异,但远程评估显著缩短评估等待时间并降低了往返成本,在不降低安全性的前提下提升了随访效率。德国一项针对慢性心力衰竭患者的队列研究同样采用PSM匹配年龄、性别、心衰分级、既往住院史等变量,结果显示远程干预组的全因死亡率和心衰再入院率显著降低,且远程组患者自我管理依从性和生活质量评分均显著提高^[26]。PSM在真实世界的远程医疗与数字健康干预研究中,有助于减少患者健康意识、疾病严重程度差异、医疗服务可及性等导致的混杂,为评估数字健康干预安全性与有效性提供了更可靠的证据。未来研究可结合PSM与时间序列分析等方法,进一步探索远程医疗对长期健康结局和卫生系统负担的影响。表1总结了观察性研究中PSM的应用场景。

表1 观察性研究中的倾向性评分匹配的常见应用场景概述

Tab 1 Common application scenarios of propensity score matching in observational studies

应用场景	应用背景	常见平衡协变量	倾向性评分匹配作用
慢性病管理与药物效果评价	慢病需长期管理,药物安全性与疗效常依赖真实世界数据;RCT受限于伦理与成本	健康意识差异、疾病严重程度、共病情况、既往医疗记录等	匹配多种基线风险因素,减少混杂,提高因果推断效率
公共卫生政策与健康服务评价	政策服务在特定地区推广,难以实现RCT,需评估其人群健康效应	经济水平、医疗服务可及性、年龄性别等人口学特征	构建特征相似的对照组,平衡社会经济与医疗资源差异
疫苗效果评估与人群差异研究	疫苗接种率与效果在人群间差异大,易受健康志愿者偏倚影响	社会经济差异、基础健康状况、医疗可及性差异	匹配不同人群接种与未接种者,评估对脆弱群体保护效应
远程医疗与数字健康干预评价	远程医疗快速普及,需要评估其对医疗利用与健康结局的影响	地理、经济、基础健康差异及患者选择倾向	匹配远程使用者与未使用者,平衡患者基线特征差异

结语 作为一种控制混杂偏倚的准实验设计方法,PSM在观察性研究中发挥了重要作用。近年来,随着因果推断框架的发展,PSM逐渐被应用于模拟RCT,以更系统地设计和分析观察性数据,从而在非随机环境下逼近RCT的推断效力。通过平衡处理组与对照组间协变量分布,PSM可有效减少已观测混杂因素的干扰,提高因果效应估计的可信度与解释性。然而,PSM仍存在一定局限,比如无法控制残余混杂,严格匹配导致样本量下降、外推性受限。基于PSM的结果仍属于观察性研究的结论,其因果推断的强度依赖于“无未测混杂”假设的合理性。目前,PSM已广泛应用于慢性病管理与药物干预、公共卫生政策与健康服务评估、疫苗效果

和远程医疗评价等领域,为循证医学和公共卫生决策提供更多参考。未来可进一步结合因果推断和机器学习等方法,提高对未测混杂的稳健性,优化倾向评分模型构建和匹配效率。

作者贡献声明 吴晔 论文撰写。田婕 论文修订。王伟炳 研究指导,论文修订,经费支持。

利益冲突声明 所有作者均声明不存在利益冲突。

参 考 文 献

[1] 黄丽红,王永吉,王素珍,等.倾向性评分方法及其规范化

- 应用的统计学共识 CSCO 生物统计学专家委员会 RWS 方法学组[J]. 中国卫生统计, 2020, 37(6): 952-958.
- [2] WAN F. Propensity score matching: should we use it in designing observational studies? [J]. *BMC Med Res Methodol*, 2025, 25(1): 25.
- [3] 焦明旭, 张晓, 刘迪, 等. 倾向性评分匹配在非随机对照研究中的应用[J]. 中国卫生统计, 2016, 33(2): 350-352.
- [4] GARRIDO MM, KELLEY AS, PARIS J, *et al.* Methods for constructing and assessing propensity scores [J]. *Health Serv Res*, 2014, 49(5): 1701-1720.
- [5] ROSENBAUM PR, RUBIN DB. The central role of the propensity score in observational studies for causal effects [J]. *Biometrika*, 1983, 70(1): 41-55.
- [6] MEDAGLIO D, STEPHENS-SHIELDS AJ, LEONARD CE. Research and scholarly methods: propensity scores [J]. *J Am Coll Clin Pharm*, 2022, 5(4): 467-475.
- [7] AUSTIN PC. A comparison of 12 algorithms for matching on the propensity score [J]. *Stat Med*, 2014, 33(6): 1057-1069.
- [8] STUART EA. Matching methods for causal inference: a review and a look forward [J]. *Stat Sci*, 2010, 25(1): 1-21.
- [9] LUNT M. Selecting an appropriate caliper can be essential for achieving good balance with propensity score matching [J]. *Am J Epidemiol*, 2014, 179(2): 226-235.
- [10] RUBIN DB. Bias reduction using Mahalanobis-metric matching [J]. *Biometrics*, 1980: 293-298.
- [11] MAO H, LI L. Flexible regression approach to propensity score analysis and its relationship with matching and weighting [J]. *Stat Med*, 2020, 39(15): 2017-2034.
- [12] AUSTIN PC, STUART EA. Optimal full matching for survival outcomes: a method that merits more widespread use [J]. *Stat Med*, 2015, 34(30): 3949-3967.
- [13] JACKSON R, JOHNSON P, BERHANE S, *et al.* Estimating treatment effects using parametric models as counter-factual evidence [J]. *BMC Med Res Methodol*, 2025, 25(1): 91.
- [14] FORTIN SP, JOHNSTON SS, SCHUEMIE MJ. Applied comparison of large-scale propensity score matching and cardinality matching for causal inference in observational research [J]. *BMC Med Res Methodol*, 2021, 21(1): 109.
- [15] RIEDL R, ROBAUSCH M, BERGHOLD A. Disease management program in patients with type 2 diabetes mellitus, long-term results of the early and established program cohort: A population-based retrospective cohort study [J]. *PLoS One*, 2022, 17(12): e0279090.
- [16] FARROKH S, NALLEBALLE K, ONTEDDU S, *et al.* Bleeding risk with combining antiplatelets and anticoagulants for secondary stroke prevention: a propensity score-matched analysis [J]. *J Am Heart Assoc*, 2025, 14(16): e042767.
- [17] MYLES PS, CORCORAN TB, CHAN MT, *et al.* Intraoperative dexamethasone and chronic postsurgical pain: a propensity score-matched analysis of a large trial [J]. *Brit J Anaesth*, 2024, 133(1): 103-110.
- [18] XIAN W, XU X, LI J, *et al.* Health care inequality under different medical insurance schemes in a socioeconomically underdeveloped region of China: a propensity score matching analysis [J]. *BMC Public Health*, 2019, 19(1): 1373.
- [19] FALCONI AM, BROMFIELD SG, TANG T, *et al.* Doula care across the maternity care continuum and impact on maternal health: evaluation of doula programs across three states using propensity score matching [J]. *EClinicalMedicine*, 2022, 50: 101531.
- [20] BEZIE MM, FENTE BM, ASMARE ZA, *et al.* Impact of health facility delivery and antenatal care on neonatal mortality in Sub-Saharan Africa: a propensity score matching analysis [J]. *BMC Pregnancy Childbirth*, 2025, 25(1): 440.
- [21] REMSCHMIDT C, WICHMANN O, HARDER T. Frequency and impact of confounding by indication and healthy vaccinee bias in observational studies assessing influenza vaccine effectiveness: a systematic review [J]. *BMC Infect Dis*, 2015, 15(1): 429.
- [22] LIN YL, WANG SI, WEI JC. Effectiveness of recombinant zoster vaccine in reducing herpes zoster incidence and all-cause mortality among patients with rheumatoid arthritis: a retrospective cohort study of 21, 046 individuals from TriNetX US Collaborative Network [J]. *EClinicalMedicine*, 2025, 85: 103319.
- [23] BRICOUT H, LEVANT MC, ASSI N, *et al.* The relative effectiveness of a high-dose quadrivalent influenza vaccine versus standard-dose quadrivalent influenza vaccines in older adults in France: a retrospective cohort study during the 2021—2022 influenza season [J]. *Clin Microbiol Infect*, 2024, 30(12): 1592-1598.
- [24] ZHANG Y, LAL LS, CHANDRA S, *et al.* Primary care telehealth in a dynamic healthcare environment from digital divide to healthcare outcomes [J]. *NPJ Digit Med*, 2025, 8(1): 211.
- [25] GLASBEY J, TROUT I, ADEYEYE V, *et al.* Use of telemedicine for postdischarge assessment of the surgical wound: international cohort study, and systematic review with meta-analysis [J]. *Ann Surg*, 2022, 277(6): e1331-e1347.
- [26] KNOLL K, ROSNER S, GROSS S, *et al.* Combined telemonitoring and telecoaching for heart failure improves outcome [J]. *NPJ Digit Med*, 2023, 6(1): 193.

(收稿日期: 2025-10-24; 编辑: 张秀峰)