

《药物流行病学研究方法学指南（第2版）》及其系列解读（8）：统计分析策略的制定与注意事项



王 晴¹, 冯欣妍¹, 王丹丹¹, 顾泽南¹, 毛广运¹, 王 怡^{1, 2}

1. 温州医科大学公共卫生学院流行病与卫生统计学系（浙江温州 325035）
2. 再生调控与眼脑健康浙江省实验室（瓯江实验室）（浙江温州 325000）

【摘要】制定规范标准的药物流行病学研究统计分析策略，对确保药物流行病学研究价值和推动学科高质量发展具有重要意义。本文基于《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”），对其中统计分析相关内容进行解读。文章第一部分阐述了药物流行病学研究统计分析策略的重要性及具体包含内容，对变量定义、样本量估计等研究概述进行说明，详细叙述了数据预处理及统计分析方法，简要概括了偏倚的控制和敏感性分析。第二部分关注统计分析策略的注意事项，对亚组分析、期中分析和统计分析报告的撰写进行阐明。

【关键词】药物流行病学；方法学；指南；统计分析策略；偏倚控制

【中图分类号】R 181.3+5 **【文献标识码】**A

Guide on Methodological Standards in Pharmacoepidemiology (2nd edition) and their series interpretation (8): development of statistical analysis strategies and key considerations

WANG Qing¹, FENG Xinyan¹, WANG Dandan¹, Gu Zenan¹, MAO Guangyun¹, WANG Yi^{1,2}

1. Department of Epidemiology and Biostatistics, School of Public Health, Wenzhou Medical University, Wenzhou 325035, Zhejiang Province, China

2. Zhejiang Lab for Regenerative Medicine, Vision and Brain Health (Oujiang Laboratory), Wenzhou 325000, Zhejiang Province, China

Corresponding author: WANG Yi, Email: wang.yi@wmu.edu.cn

【Abstract】The establishment of standardized statistical analysis strategies is of great significance to ensuring scientific value and promoting high-quality development in pharmacoepidemiology research. Based on the *Guide on Methodological Standards in Pharmacoepidemiology (2nd edition)*, this article interprets the statistical analysis content within it section of the Guidelines. The first part outlines the importance and specific content of statistical analysis strategies in pharmacoepidemiology studies, introduces the variable definitions and sample size estimation and elaborates on data preprocessing and statistical methods. It also briefly discusses bias control and sensitivity analysis. The second part focuses on considerations for developing statistical analysis strategies, with further clarification on subgroup analysis, interim analysis, and the preparation of the statistical analysis reports.

DOI: 10.12173/j.issn.1005-0698.202508028

基金项目：国家自然科学基金面上项目（82371744）；温州市重大科技创新攻关项目（ZY2023019）

通信作者：王怡，博士，教授，博士研究生导师，Email: wang.yi@wmu.edu.cn

<https://ywlxbx.whuzhmedj.com/>

【Keywords】Pharmacoepidemiology; Methodology; Guildelines; Statistical analysis strategies; Bias control

统计分析作为药物流行病学研究的关键环节，其规范性与结果可靠性直接影响对药物有效性与安全性的评估。不规范的统计分析可能带来诸多隐患，如混杂控制不足引入偏倚、错误应用方法扭曲效应估计、结果报告不全损害可重复性等。随着方法学研究的深化与国际国内指南的不断完善，统计分析过程的规范化日益受到重视。欧盟药物流行病学和药物警戒网络中心（the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance, ENCePP）发布的《药物流行病学研究方法学标准指导手册》^[1]、国际药物流行病学学会（International Society for Pharmacoepidemiology, ISPE）发布的《药物流行病学规范指南》^[2]、欧洲药品管理局（European Medicines Agency, EMA）发布的《药物警戒质量管理规范》模块 VIII《上市后药品安全性研究》^[3]以及我国发布的《药物流行病学研究方法学指南（第1版）》^[4]等国内外权威指南，均对统计分析计划（statistical analysis plan, SAP）、分析流程及敏感性分析提出了明确要求。其中 ENCePP 与 EMA 的指南还对 SAP 的制定时机和具体内容提供了详细要求。

《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”）^[5]在借鉴国际先进规范的基础上，进一步明确了药物流行病学研究中统计分析策略的制定标准与实施路径。本文重点解读指南第2版中涉及统计分析的内容，为规范开展药物流行病学研究提供方法学依据。

1 统计分析策略的制定

统计分析策略是涵盖研究概述、变量定义、数据处理、分析方法、结果表达与偏倚控制的系统性框架，其规范化程度直接关系到研究结论的可靠性与外推性。统计分析的内容及内涵见表1。指南建议在研究方案确定后立即制定 SAP，以切实提高研究的透明性与可重复性。

1.1 假设检验

研究假设应基于理论提出，是对变量间关系的可检验陈述，有助于避免无预设假设下的“数据疏浚”（data dredging）及其导致的 I 型错误增加。假设分为零假设（ H_0 ）和备择假设（ H_1 ），前者假设无差异或无关，后者为研究者希望证实的假设，并反映检验的单双侧。假设检验基于小概率反证法，通过检验 H_0 间接推断 H_1 是否成立，

表1 统计分析的内容及内涵
Table1. Components and descriptions of statistical analysis

统计分析要素	内涵
假设检验	提出明确、具体的可检验假设，假设需与研究问题和目标紧密关联，基于演绎推理，详细说明结局指标测量方法，重视阴性结果的价值
变量定义	明确定义所有研究变量（包括时依性变量），区分基于历史数据的临床变量转换与统计模型拟合转换。对于统计模型拟合转换，说明拟合检验方法和转换条件
研究设计和样本量	观察性研究中样本量非首要考虑因素，但需提供预期的研究对象总数及分层分析中各层的最小样本数作为参考
期中分析（如适用）	需预先规划期中分析的标准、条件及缺陷处理措施（包括停止规则），并适当调整统计方法
数据源、研究人群和分析集	描述数据源和链接方法、纳入/排除标准、退出/随访机制、研究开始/索引日期、基线受试者特征、潜在混杂变量及完整分析集。明确受试者分组逻辑和亚组定义（如适用）
分析方法	描述用于解决每个研究目标的效应测量和统计方法，包括： • 数据处理，缺失值、异常值，多源数据整合； • 统计技术，点估计/区间估计、复合变量/派生变量计算规则、统计模型的选择、治疗交互作用分析、亚组分析； • 偏倚控制，选择偏倚、信息偏倚、时间相关性偏倚、混杂偏倚四类偏倚。其中混杂偏倚的识别和控制是关键环节； • 软件工具，明确使用的统计分析程序和统计软件包
统计原则	明确显著性水平（单/双侧检验）、置信区间；针对重复检验需调整显著性水平，任何计划进行的显著性水平调整应予以说明；注意统计软件默认值（如 $\alpha=0.05$ ）的适用性
敏感性分析	用于探究不同假设或特定观测值对结果的影响，支持主要分析结论。探索性分析需明确标注性质
决策标准	若研究结果用于决策，需预先定义影响决策的关键统计数据及其阈值（如不同结局对应的统计量临界值）

是统计推断的核心。

1.2 定义研究变量及变量转换

暴露 / 干预可定义为二分类变量，或者进一步细化如持续时间、时间窗（如当前与既往暴露）或剂量（如当前用量、累积剂量）。结局需区分主要、次要和探索性结局，并明确评估时间、方法及事件判定标准，提供编码列表。若结局变量需由原始数据（如实验室或影像学资料）转换而来，应详述转换理由与方法，并通过敏感性分析验证。协变量需说明识别方法及评估时间点，其时间变异性对结果准确性有重要影响，可基于不同时间段或全部历史数据评估。协变量常用于研究对象筛选、匹配、倾向性评分（propensity scoring, PS）构建、亚组分析及混杂控制。变量转换（如对数转换、Box-Cox 正态性转换等）常用于满足模型前提条件或提升模型拟合效果。

1.3 估计样本量

样本量估计是保证统计功效和优化资源的关键。需根据结局指标的变量类型（定量 / 定性）选择相应公式或软件（如 PASS 软件）。样本量估计的核心参数包括：I 型错误 α （常取值 0.05 或 0.025）、II 型错误 β （检验效能 $=1-\beta$ ，一般 $\geq 80\%$ ）、效应量（如均值差、率差、相对危险度、比值比、风险比等），以及其他因素如组间分配比例、非劣效界值、脱落率（常增加 10%~20% 样本量）等。不同研究设计对应的样本量估计参数详见表 2。多因素分析中广泛采用 EPV（events per variable）原则，即每个预测变量至少需要 10 个事件^[6]；更推荐使用 EPP（events per candidate predictor parameter）以应对多分类、交互项等复杂参数结构。预测模型研究可借助如 pmsampsize 工具进行样本量估算^[7]。

表2 按研究设计类型匹配的样本量估计参数
Table 2. Sample size estimation parameters based on study design

研究设计类型	关键参数
随机对照试验	干预组和非干预组的发病率 (p_1, p_2) 或 δ, σ
队列研究	暴露组和非暴露组的发病率 (p_1, p_2), RR 或 δ, σ
病例对照研究	病例组与对照组的暴露率 (p_1, p_2), OR 或 δ, σ , 匹配比
横断面研究	预期现患率 (p), 容许误差 (d)
生存分析	实验组和对照组的风险率 (λ_1 和 λ_2), HR, 随访时间

注：RR. 相对危险度（relative risk）；OR. 比值比（odds ratio）；HR. 风险比（hazard ratio）。

1.4 确定分析集

分析集指经纳入排除标准筛选，并经基线评估后最终用于分析的受试者集合。干预性研究常按依从性及数据完整性分为完整分析集、符合方案集和安全性分析集^[8]；观察性研究通常定义一个主分析集，并通过变更暴露定义等进行敏感性分析。亚组分析用于探索暴露效应修饰，需依据分析目标（主要、次要或探索性）选择对应分析集，并通过流程图清晰展示人群筛选过程。

1.5 数据预处理

数据预处理旨在将原始数据转化为适合统计分析的高质量、结构化数据集，包括清洗、集成、规约和变换，多源数据还需进行整合。

1.5.1 数据清洗

数据清洗是删除原始数据集中的无关数据、重复数据，平滑噪声数据，处理缺失值与异常值。缺失值按缺失机制分为完全随机缺失（missing completely at random, MCAR）、随机缺失（missing at random, MAR）和非随机缺失（missing not at random, MNAR）。处理方法包括：删除法[如完全案例分析（complete case analysis, CCA）^[9]、插补法[单一插补法，如均值插补法、回归插补法、末次观测值结转法、最差观测值结转法^[10-11]和哑变量法^[12]等；多重插补法（multiple imputations, MI）^[11]；基于模型插补法，如最大似然估计法、最大期望值法、其他机器学习模型插补法^[13]、逆概率加权法^[14]。这些处理方法中，MI 常用全条件定义法，适用于 MCAR 和 MAR 数据^[15]。对于 MCAR 数据，样本量大且缺失率低可使用 CCA，或者进行单一插补或多重插补；对于 MAR 数据常使用 MI，当预测缺失值的辅助信息不足时使用最大似然估计法或逆概率加权法；MNAR 需借助选择模型、模式混合模型和共享参数模型等^[16]。缺失值处理后需要进行敏感性分析评估处理方式的稳健性^[17]。异常值的识别可通过可视化（如箱型图和散点图）、Z-score 标准化、 3δ 原则以及基于模型的方法（如孤立森林、局部离群因子等），并可进行删除、替换或保留处理。

1.5.2 数据集成与数据规约

数据集成需统一实体识别、解决冗余与冲突。数据规约包括维度规约和数值规约。维度规约旨在减少数据集中的特征数量（列数），主要通过特征选择（如过滤法、包裹法、嵌入法）和特征

提取（如主成分分析）实现。数值规约旨在减少数据集中的实例数量（即行数），主要通过参数方法（如回归模型）和非参数方法（如抽样或聚类）实现。

1.5.3 数据变换

数据变换包括数据规范化与数据离散化。数据规范化也称数据标准化，用于消除指标之间的量纲和量级差异的影响，方法包括最小-最大缩放、Z-score 标准化、对数变换和小数定标。数据离散化是将连续数据分段处理使其变为离散化区间，方法包括等频法、等宽法、聚类法等。

1.5.4 多源数据整合

为提高研究结果的普适性与准确性，药物流行病学研究日益采用多源数据。然而，数据源间的异质性要求对数据进行统一整合与管理。所有基于多源数据的研究均需遵循共同的研究方案和 SAP。常用的整合策略包括以下四类：①本地分析：各中心独立分析本地数据，无需传输原始数据或使用通用数据模型（common data model, CDM），仅共享分析结果（如患者层面数据集、聚合数据或估计值）；②数据共享：将原始数据集中传输至中央平台进行统一分析，适用于结构相似的数据库，但隐私风险较高且受法规严格限制；③研究特定数据模型：将原始数据转换为研究特定数据模型格式，各中心采用统一程序分析后共享结果；④CDM：提前将数据转换为 CDM 存储，研究时按统一方案与 SAP 进行分析和结果共享^[18]。

数据异质性是整合过程中的核心挑战，常源于医疗体系、人群特征及数据质量等方面的差异。可通过制定共同研究方案与 SAP，共享数据管理程序，采用 CDM 统一变量定义、结构与格式（如字段命名、术语体系、编码标准等）以提升标准化与互操作性^[1, 19]。数据整合后需实施全流程质量管理，涵盖数据收集、存储与传输的标准化记录，采用“适配研究目标”（fit-for-purpose）原则进行评估，并利用概率匹配技术（如基于间接标识符的记录交叉验证）增强数据一致性。针对跨国数据整合中的隐私与法规壁垒（如伦理审查制度、数据管理要求差异），应采用匿名化与去标识化技术、遵循“最小必要原则”限制数据使用范围，并借助基于 CDM 的联邦分析等分布式技术^[20]解决。建议研究参照《药物流行病学研究

中数据库选择与使用的良好指南》^[21]进行系统性核查，以保证操作规范性与结果可靠性。

1.6 统计分析方法的选择

1.6.1 描述性分析

描述性分析通过统计指标或图表呈现数据的分布及特征。定量变量应报告集中趋势、离散趋势和分布形态，通常先进行正态性检验：符合正态分布时采用均数和标准差，非正态分布采用中位数与四分位数间距，可借助直方图或箱式图进行可视化。定性变量（包括分类与等级变量）以频数、率、构成比或相对比描述，常用直条图、饼图或列联表展示。统计表是呈现分析结果的重要形式，应主宾分明、条理清晰、格式统一，并按照变量重要性排序以突出关键信息。

1.6.2 关联性分析

在描述性分析基础上，需进行假设检验以探索变量间关联。定量变量满足正态性和方差齐性时可采用 *t* 检验或方差分析，否则建议使用秩和检验等非参数方法。分类变量常用 χ^2 检验或 Fisher 确切概率法，等级变量适用 Wilcoxon 秩和检验、Wilcoxon 符号秩检验或 Kruskal-Wallis 秩和检验。方法选择需结合数据特征、设计类型及样本量，大样本非正态分布数据有时也可使用参数检验，小样本则更推荐非参数方法^[22]。

多因素分析常用于识别暴露 / 干预与结局的独立关联性，常用模型包括多元线性回归、Logistic 回归、Cox 比例风险回归和 Poisson 回归^[22]。建模时需检验自变量共线性及模型拟合优度，变量类型与模型选择关系参见图 1。

多元线性回归适用于连续因变量与多个自变量的线性关系分析，需满足线性、独立性、正态性和方差齐性等条件。自变量筛选可选用前进法、后退法、逐步回归法或全局择优法。分类自变量需进行哑变量转换。Logistic 回归适用于二分类、有序或多分类因变量，可通过 OR 量化自变量与结局的关联强度。条件 Logistic 回归适用于配对设计。模型需满足观测对象独立性、因变量服从二项分布、连续自变量与 logit 转换后的因变量呈线性关系，且自变量间无严重共线性。自变量筛选方法分为前进法、后退法和逐步法。Cox 比例风险回归是生存分析的常用方法，适用于时间-事件数据。需先通过 Kaplan-Meier 曲线或寿命表法进行生存率等的描述，再根据是否满

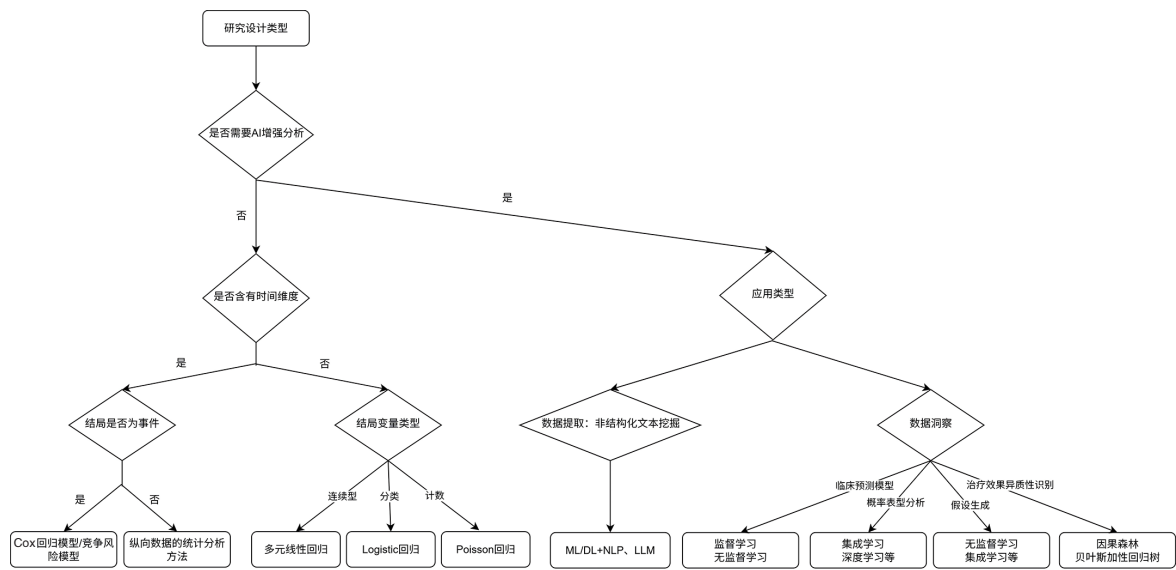


图1 统计分析建模方法选择决策树
Figure1. Decision tree for modeling methods

足比例风险（proportional hazards, PH）假设选择后续方法：满足时可用 log-rank 检验或 Breslow 检验进行组间比较并构建 Cox 模型；不满足时可选用加权 log-rank 检验、时依协变量 Cox 模型、分层 Cox 模型或加速失效时间模型。PH 假设可通过 Kaplan-Meier 曲线、对数-对数生存曲线、Schoenfeld 残差图及 Schoenfeld 残差检验进行判断。若存在多终点事件竞争风险，应选用竞争风险模型。Poisson 回归适用于罕见计数结局，要求事件独立发生且符合 Poisson 分布（可通过数据平稳性、独立增量性和普通性等特征或是 Poisson 检验进行判断）。

1.6.3 纵向数据的统计分析

在药物流行病学研究中常需处理纵向数据，即对同一研究对象在不同时间点或聚类下进行多次测量所获得的数据。此类数据的分析方法主要包括重复测量方差分析、混合效应模型、非线性混合效应模型、广义线性混合模型（generalized linear mixed model, GLMM）、广义估计方程（generalised estimation equations, GEE）和潜变量增长模型（latent growth model, LGM）等^[23]。重复测量方差分析是方差分析的拓展，除考虑组内变异外，还纳入时间点变异及处理与时间的交互作用。该方法要求数据满足独立性、正态性、方差齐性与“球对称”假设，且需保持测量时点与间隔相同的平衡设计，适用于比较不同处理因素和各时间点结局指标总体平均水平的差异。混

合效应模型同时包含固定效应（自变量对因变量的总体影响）和随机效应（个体或组间变异），适用于具有分层、分组或嵌套结构的数据，要求因变量服从正态分布且与自变量呈线性关系，可评估观测值随时间的变化及不同处理或个体水平的差异。非线性混合效应模型适用于变量间存在非线性关系的情形（如药代动力学数据），通过在非线性结构中引入固定效应与随机效应，捕捉不同层次的差异。该模型对变量分布无特定要求。广义线性模型（generalized linear model, GLM）通过连接函数建立因变量与自变量线性预测值之间的关系，适用于因变量服从指数分布族（如正态分布、Poisson 分布、二项分布等）的数据。GLMM 为 GLM 与混合效应模型的结合，在线性预测项中加入随机效应，能够同时处理非正态分布因变量及数据相关结构，适用于嵌套数据、重复测量等具有层次结构的分析。GEE 属于边际模型，通过设定“作业相关矩阵”刻画组内相关结构，适用于非正态因变量和简单相关结构（如交换相关、自回归相关）数据，主要估计总体平均效应。在大样本情况下，即使误设“作业相关矩阵”，仍可获得参数估计的一致性结果。LGM 是结构方程模型的一种，通过截距和斜率两个潜变量刻画个体随时间的变化轨迹，可评估协变量（包括时依协变量）对变化过程的影响，适用于个体内及个体间变异均显著的数据，要求变量为连续型且满足正态性和独立性。纵向数据分析方法的选择

应综合考虑其适用条件、优势与局限性，具体可参考表3进行决策。

1.6.4 人工智能方法的应用

人工智能（artificial intelligence, AI）技术，主要包括机器学习（machine learning, ML）与自然语言处理（natural language processing, NLP），在药物流行病学研究中的应用集中于两类任务：一是从非结构化文本中提取和构建数据，二是开展数据洞察。AI可以借助命名实体识别和关系抽取技术，将电子健康记录、自发报

告和产品文档等非结构化文本转换为可用于分析的结构化信息，实现药物不良事件识别和目标人群（可计算表型）的自动提取^[24-25]。大语言模型（large language models, LLM）作为先进的NLP工具，能够更高效地处理大规模复杂文本，自动完成数据标准化与真实世界证据的整理^[26]。在数据洞察方面，ML常用于混杂控制、临床预测模型构建与概率表型分析。ML方法（如随机森林、神经网络）通过自动化变量筛选与加权，更灵活地模拟非线性及交互效应，从而提高混杂

表3 纵向数据的统计分析方法
Table 3. Statistical analysis methods for longitudinal data

模型	适用条件	优势	局限性
重复测量方差分析	1. 结局变量：连续且满足正态性、独立性、方差齐性； 2. 平衡数据； 3. 满足球对称假设； 4. 小样本量	1. 模型简单，结果易于解释； 2. 适用于总体平均水平差异分析	1. 适用范围窄：对平衡数据和球对称假设要求极为严格，在真实世界研究中很难满足； 2. 无法处理时依协变量和缺失数据
混合效应模型	1. 结局变量：连续且满足正态性； 2. 因变量与自变量呈线性关系； 3. 随机效应服从多元正态分布，残差满足正态性与方差齐性； 4. 大样本量	1. 可以处理非平衡数据和随机缺失数据； 2. 可同时估计固定效应（总体平均）和随机效应（个体变异），适用于分层、分组、嵌套结构的数据； 3. 可纳入时依协变量进行时间发展轨迹差异分析	1. 适用于少数自变量与一个因变量的相对简单的回归结构； 2. 模型设定更复杂，需要明确指定固定和随机效应结构； 3. 需要较大样本量才能获得足够的检验效力
非线性混合效应模型	1. 结局变量：连续型、离散型； 2. 对数据分布无要求； 3. 随机效应服从多元正态分布，残差满足正态性与方差齐性； 4. 大样本量	1. 同混合效应模型，且适用于参数间的非线性关系； 2. 适用于拟合复杂的、理论驱动的非线性增长轨迹	1. 模型极复杂，计算强度大，收敛困难； 2. 需要较强的先验知识来设定正确的非线性函数形式； 3. 对初始参数值敏感
GLMM	1. 结局变量：连续型、离散型，满足指数分布族； 2. 在连接函数转换的尺度上建立线性预测值，随机效应服从正态分布； 3. 需要较大样本量，尤其对于分类结局或复杂随机效应	1. 扩展了混合效应模型的应用范围，使其能处理各种类型的结局变量； 2. 可以处理非平衡数据和随机缺失数据； 3. 适用于分层、分组、嵌套结构的数据和具有相关性的数据； 4. 适用于个体和总体平均水平差异分析	1. 参数解释是在连接函数尺度上，而非原始概率尺度，解释稍复杂； 2. 对于二分类等数据，模型收敛问题可能更常见； 3. 边际效应需要额外计算
GEE	1. 结局变量：连续型、离散型，对数据分布无要求； 2. 无需指定随机效应分布，只需指定一个“工作相关矩阵”； 3. 大样本量	1. 稳健性强：对相关结构假设不敏感，只要样本量足够，工作相关矩阵指定错误，系数估计仍稳健； 2. 可以处理随机缺失数据，能提供边际效应推断； 3. 适用于总体平均水平差异分析； 4. 有效处理纵向数据中结局变量的相关性	1. 无法处理非平衡数据； 2. 不包括随机效应，不提供个体水平的推断或预测； 3. 小样本时，基于经验的标准误差会低估真实的误差，导致I型错误率增加
LGM	1. 结局变量：连续且满足正态性、独立性、方差齐性； 2. 平衡数据； 3. 在结构方程模型框架下，将个体轨迹视为潜变量； 4. 需要大样本量（通常 $n>200$ ），模型才稳定	1. 纳入时依协变量进行时间发展轨迹差异分析，并检验轨迹的异质性； 2. 适用于处理个体内变化和个体间差异均明显的的数据，并找出变化/差异的预测因素； 3. 可以处理测量误差； 4. 可以处理随机缺失数据	1. 模型复杂，需要非常大样本量； 2. 对分布假设敏感，模型识别和收敛可能存在问题

控制能力^[27]。基于机器学习（如回归算法、无监督学习）构建的临床预测模型已广泛应用于药物安全性、有效性评价及信号检测，但目前其预测问题定义、模型报告规范和性能评估标准仍待完善^[28-29]。概率表型分析依托模型估计特定表型的发生概率，是机器学习应用的新兴方向^[30]，例如借助集成学习（如 XGBoost）和深度学习（如贝叶斯神经网络）识别潜在亚群与表型关联。机器学习还可用于因果探索中的假设生成及治疗异质性识别。

尽管 AI 显示巨大潜力，其应用常因报告不透明、流程不清晰而面临可重复性差挑战。为提高结果可靠性并满足监管要求，研究者在报告 AI 模型时需遵循以下原则^[31-33]：①开发流程透明化：明确说明训练集、验证集和测试集的划分方法、样本量及比例，公开随机种子；阐述算法选择与超参数优化策略及最终取值；鼓励公开代码和去标识化数据并提供运行环境说明。②模型验证严格化：应报告模型区分度[如曲线下面积（area under the curve, AUC）、C-index]、校准度（如 Calibration 校准曲线、Brier 得分）、临床效用指标[如决策曲线分析（decision curve analysis, DCA）、临床影响曲线（clinical impact curve, CIC）]^[34]；必须进行外部验证以证明模型跨人群稳健性。③可解释性优先原则：应优先选用本质可解释模型（如 Logistic 回归、Cox 回归）或采用 SHAP（shapely additive explanations）、LIME（local interpretable model-agnostic explanations）等事后解释方法提供变量贡献度^[35]；若使用复杂模型，须提供特征重要性、典型实例解释及临床评审意见^[36]。

1.7 偏倚的控制

偏倚指在研究设计、实施或分析过程中引入的系统性误差，导致暴露与结局之间的关联被扭曲，常见类型包括选择偏倚、信息偏倚、时间相关性偏倚和混杂偏倚。

1.7.1 选择偏倚

选择偏倚源于研究对象筛选或参与机制的缺陷，导致入组人群与未入组人群在某些特征上存在差异，影响样本代表性。常见类型包括入院率偏倚、现患病例-新发病例偏倚、志愿者偏倚、无应答偏倚和失访偏倚等。在研究设计阶段，可采用新用药者设计、随机分组、严格纳入排除标准及多中心/多数据库策略予以控制；在分析阶段，

可将已知选择偏倚变量作为混杂因素加以调整，明确报告失访原因并对缺失数据采用多重插补等方法处理。

1.7.2 信息偏倚

信息偏倚产生于数据收集、测量或分类过程中的误差，涉及暴露、协变量或结局变量的错误记录。其常见来源包括数据标准不一致、变量定义模糊和信息不完整，具体表现为回忆偏倚、检出征候偏倚和错误分类偏倚等。控制方法包括制定详细且统一的数据收集规程、实施质量控制、采用客观暴露指标、使用标准病例定义（如世界卫生组织标准）、推行盲法评估，以及通过多源信息（如病历、处方和实验室结果）交叉验证。

1.7.3 时间相关性偏倚

时间相关性偏倚在药物流行病学中常见却易被忽视，主要包括原发性偏倚、非死亡时间偏倚、时间窗偏倚、延迟时间偏倚、滞后时间偏倚、不可测量的时间偏倚和易感者损耗偏倚^[37-38]。原发性偏倚，又称因果倒置偏倚，指因疾病早期症状或高风险而接受暴露（如药物治疗），误将疾病归因于该暴露；非死亡时间偏倚，又称幸存者治疗选择偏倚，源于队列定义、随访开始和暴露发生时间的不匹配，错误地将未暴露归入暴露时间^[39-40]；时间窗偏倚发生于病例对照研究中，因暴露组与非暴露组观察时间窗长度不同导致^[41]。控制方法因类型而异：原发性偏倚可通过设置暴露前时间窗或限制于症状与治疗无关的病例^[42]；非死亡时间偏倚需统一关键时间点或采用时依 Cox 模型；时间窗偏倚可通过巢式病例对照设计或时间依赖性抽样控制^[41, 43]；滞后与延迟时间偏倚需统一病程或设置诱导期；不可测量时间偏倚应选用暴露信息完整的数据；易感者损耗偏倚可采用新用药者设计^[38]。

1.7.4 混杂偏倚

混杂偏倚指在暴露与结局关联估计中，受一个或多个既与结局相关又与暴露相关的混杂因素影响，导致效应估计失真。适应证混杂是药物流行病学中的典型类型，因药物使用与疾病严重程度相关，而疾病严重程度又影响结局（如药物不良反应），此时疗效差异可能源于疾病严重程度而非药物本身^[44]，可采用阳性对照进行控制。未测量混杂或残余混杂指未被记录或调整的混杂因素所致偏倚；时依性混杂指随时间变化的混杂

因素同时受既往暴露影响、又影响后续暴露和结局^[45]。控制方法包括：设计阶段采用限制、匹配或随机化；分析阶段可采用直接标化、分层分析、多因素回归、PS 和疾病风险评分（disease risk score, DRS）等控制已知混杂^[46]；对未知混杂可使用敏感性分析、工具变量、阴性对照、本底事件率校正法（prior event rate ratio adjustment, PERR）和趋势-趋势设计；时依性混杂常用 G-方法（G-computation, G-estimation）和边际结构模型（marginal structure model, MSM）^[47]。如在某降压药与心梗风险的队列研究中，若将“队列入组至用药前的无事件时间”误归为“用药暴露时间”，则会低估风险，此时可通过时依 Cox 回归将暴露定义为随时间变化的变量，以纠正偏倚。PERR 是一种新兴方法，通过比值的方式比较暴露前两组发生率比（RR_p）和暴露后率比（RR_s）（RR_{perr}=RR_s/RR_p）以校正未测量混杂，也适用于 HR 和 OR 的调整^[48]。

1.8 敏感性分析

敏感性分析旨在评估主要统计方法在面对模型假设偏离或数据缺陷时的稳健性^[49]。该分析应围绕主要研究目标展开，通过变换关键假设或方法（如调整纳入排除标准、更换变量定义标准、调整统计模型、改变缺失数据处理方式等），比较效应估计值的变化，并可使用 E 值等工具评估未测量混杂的影响。所有敏感性分析均需预先在 SAP 中明确说明。尽量避免事后敏感性分析，若确有必要，需详细解释理由。结果建议汇总于统一表格中，包含效应估计值、置信区间及所用假设；当与主要分析结论不一致时，应以预先设定的主要分析结果为准^[50]。

2 统计分析注意事项

2.1 亚组分析

在药物流行病学研究中，尽管通过严格纳入排除标准、设立对照或匹配等方法力求研究对象间的同质性，但真实世界研究中的对象常因遗传、人口、环境、病理生理、合并症与合并用药、地域等因素存在不可避免的异质性。为探索和解释这种异质性，亚组分析通过按不同特征将研究人群分组，以考察暴露或干预效应在各亚组间是否存在差异。根据研究目的，亚组分析通常分为三类：探索性亚组分析，旨在发现可能导致效应

差异的因素，用于生成假设，为后续研究指引方向，无需事先定义；支持性亚组分析，用于评估效应在不同亚组中的一致性，为研究结论的普遍性提供支持证据；确证性亚组分析，用于验证暴露/干预在特定亚组中的效应（如哪个亚组疗效更优），常为药物说明书撰写提供支持，需要预先定义并进行多重性校正^[51]。进行亚组分析时，应注意：不宜设置过多亚组，以免降低结果可信度；预先估计每个亚组的样本量，确保有足够的检验效能；合理选择统计学显著性水平（ α ）并进行多重检验校正以控制 I 型错误膨胀；对于亚组分析结果的解释应基于生物学理论，谨慎对待偶然发现。

2.2 期中分析

期中分析指在临床试验正式完成前，依据预先制订的 SAP，对不同治疗组间的有效性和安全性所进行的分析。其目的在于及时监测试验安全性，尽早确认药物疗效，并在必要时重新估计样本量^[52]。部分投入大、周期长的观察性研究也可采用此方法。研究方案若计划进行期中分析，需明确阐明试验的终止规则、分析时点、具体实施方式以及所采用的 α 、 β 和样本量^[53]。任何对预先计划的修改或偏离均应说明原因并详细记录。为控制多次检验带来的 I 型错误膨胀，需对显著性水平（ α ）进行校正。常用方法包括 Pocock 法、O'Brien-Fleming 法、Peto 法及 Lan-Demets 的 α 消耗函数法^[54-55]。期中分析次数应受到严格控制。

2.3 决策标准

若研究结果将用于临床或监管决策（如药物上市批准、说明书修改），必须在研究开始前的方案中明确规定决策所用的统计量（如 HR、置信区间宽度、P 值）及具体的决策阈值（如 $P < 0.01$ ）。清晰的决策标准有助于确保分析结果的客观性，避免事后解释带来的偏倚，并为监管机构提供明确的审评依据。

2.4 统计分析报告

在研究结束后需根据研究类型选择合适的报告规范撰写统计分析报告，以提高结果的透明度和科学性。常用报告规范包括：随机对照试验的 CONSORT 声明，干预性研究的 TREND 声明及 TIDieR 检查表，观察性研究的 STROBE 声明和 RECORD-PE 规范，系统评价与 Meta 分析的 PRISMA 声明，以及观察性研究 Meta 分析的

MOOSE 声明等。选择时应注意：观察性研究（如队列研究、病例对照研究）宜选用 STROBE；真实世界干预研究可结合 TREND；使用数据库的研究需遵循 RECORD-PE；Meta 分析则应参考 PRISMA 规范。统计分析报告通常包括研究概述、统计分析方法、结果与结论三个部分，其中前两部分需与 SAP 一致。研究概述应明确研究目标、研究人群及样本量、变量（包括复合变量或派生变量）的定义及计算规则。统计分析方法需要说明使用的统计软件（如 SAS、SPSS、Stata、R 语言和 Python 等软件）与统计包、统计描述和模型构建方法。SAS 软件更适用于多中心临床试验和上市后安全性研究的大数据处理，R 语言和 Python 软件在真实世界研究和 AI 应用中优势显著。假设检验须注明单 / 双侧及显著性水平，并关注多重性问题（如多组比较、多个主要终点、亚组分析、期中分析导致的多次检验），可采用 Bonferroni 法、Hochberg 序贯法等调整 α 以控制 I 型错误。结果与结论应使用统计表和图呈现，效应值需报告点估计值、置信区间和 P 值，并对阴性结果慎重解读。

3 结语

药物流行病学研究中的统计分析策略制定是一个系统而复杂的过程，其规范实施对产生真实、可靠的证据至关重要。本文基于指南第 2 版，系统阐述了统计分析计划的制定内容与关键步骤，包括假设构建、变量定义、样本量估算、分析集确定、数据预处理、方法选择、偏倚控制及敏感性分析等方面，并强调了亚组分析、期中分析、决策标准与报告规范在实际应用中的注意事项。规范的统计分析不仅有助于提升研究的科学性与透明度，也是支持医药卫生决策和药物安全性评价的重要基石。未来仍需持续关注方法学进展，并在实践中不断优化与验证统计分析策略。

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

- 1 The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on Methodological Standards in Pharmacoepidemiology (revision 11) [EB/OL]. (2023-07-

- 13) [2025-06-04]. https://encepp.europa.eu/encepp-toolkit/methodological-guide_en.
- 2 International Society of Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practice (GPP)[J]. Pharmacoepidemiol Drug Saf, 2016, 25(1): 2-10. DOI: 10.1002/pds.3891.
- 3 European Medicines Agency. Guideline on Good Pharmacovigilance Practices (GVP) Module VIII-Post-Authorisation Safety Studies (Rev 3) [EB/OL]. (2017-12-10) [2025-06-04]. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf.
- 4 胥洋, 丁呈怡, 詹思延. 中国药物流行病学研究方法学指南[J]. 药物流行病学杂志, 2019, 28(1): 5-9. [Xu Y, Ding CY, Zhan SY. Guide on Methodological Standards in Pharmacoepidemiology[J]. Chinese Journal of Pharmacoepidemiology, 2019, 28(1): 5-9.] DOI: 10.19960/j.cnki.issn1005-0698.2019.01.002.
- 5 颜济南, 吴昀效, 聂晓璐, 等. 《中国药物流行病学研究方法学指南（第 2 版）》的制订 / 修订过程[J]. 药物流行病学杂志, 2025, 34(2): 121-35. [Yan JN, Wu YX, Nie XL, et al. Revision process of the Guide on Methodological Standards in Pharmacoepidemiology in China (2nd edition)[J]. Chinese Journal of Pharmacoepidemiology, 2025, 34(2): 121-135.] DOI: 10.12173/j.issn.1005-0698.202502028.
- 6 Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis[J]. J Clin Epidemiol, 1996, 49(12): 1373-1379. DOI: 10.1016/s0895-4356(96)00236-3.
- 7 Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model[J]. BMJ, 2020, 368: m441. DOI: 10.1136/bmj.m441.
- 8 European Medicines Agency. ICH E9 statistical principles for clinical trials[EB/OL]. (1998-09-01) [2025-06-13]. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf.
- 9 Heymans MW, Twisk JWR. Handling missing data in clinical research[J]. J Clin Epidemiol, 2022, 151: 185-188. DOI: 10.1016/j.jclinepi.2022.08.016.
- 10 Rigopoulos D, Tampouratzi E, Angelakopoulos C, et al. Real-world data on the effectiveness of brodalumab in patients with moderate-to severe plaque psoriasis in the Greek clinical setting (the BrIDGE study)[J]. J Eur Acad Dermatol Venereol, 2024, 38(6): 1121-1130. DOI: 10.1111/jdv.19816.
- 11 朱荣慧, 许金芳, 王睿, 等. 多重填补技术在医学研究缺失值处理中的应用及发展[J]. 中国卫生统计, 2022, 39(2): 293-295, 298. [Zhu RH, Xu JF, Wang R, et al. Application and development of multiple imputation in handling missing values in medical research[J]. Chinese Journal of Health Statistics, 2022, 39(2): 293-295, 298.] DOI: 10.3969/j.issn.1002-3674.202202034.
- 12 Vetter C, Devore EE, Wegryzn LR, et al. Association between rotating night shift work and risk of coronary heart disease among women[J]. JAMA, 2016, 315(16): 1726-1734. DOI: 10.1001/

- jama.2016.4454.
- 13 黎剑锋, 张静怡, 李立康, 等. 临床试验中缺失值的处理方法探讨[J]. 协和医学杂志, 2024, 15(5): 1165–1172. [Li JF, Zhang JY, Li LK, et al. Approach to handling missing data in clinical trials[J]. Medical Journal of Peking Union Medical College Hospital, 2024, 15(5): 1165–1172.] DOI: [10.12290/xhyxzz.2023-0415](https://doi.org/10.12290/xhyxzz.2023-0415).
 - 14 Mongin D, Lauper K, Finckh A, et al. Accounting for missing data caused by drug cessation in observational comparative effectiveness research: a simulation study[J]. Ann Rheum Dis, 2022, 81(5): 729–736. DOI: [10.1136/annrheumdis-2021-221477](https://doi.org/10.1136/annrheumdis-2021-221477).
 - 15 Lee KJ, Tilling KM, Cornish RP, et al. Framework for the treatment and reporting of missing data in observational studies: the treatment and reporting of missing data in observational studies framework[J]. J Clin Epidemiol, 2021, 134: 79–88. DOI: [10.1016/j.jclinepi.2021.01.008](https://doi.org/10.1016/j.jclinepi.2021.01.008).
 - 16 韦连慧, 郭晓晶, 许金芳, 等. 缺失数据处理在药物流行病学研究中的应用及进展[J]. 药物流行病学杂志, 2022, 31(1): 60–64. [Wei LH, Guo XJ, Xu JF, et al. Application and progress of missing data handling methodology in pharmacoepidemiology[J]. Chinese Journal of Pharmacoepidemiology, 2022, 31(1): 60–64.] DOI: [10.19960/j.cnki.issn1005-0698.2022.01.009](https://doi.org/10.19960/j.cnki.issn1005-0698.2022.01.009).
 - 17 Madley-dowd P, Hughes R, Tilling K, et al. The proportion of missing data should not be used to guide decisions on multiple imputation[J]. J Clin Epidemiol, 2019, 110: 63–73. DOI: [10.1016/j.jclinepi.2019.02.016](https://doi.org/10.1016/j.jclinepi.2019.02.016).
 - 18 Gini R, Sturkenboom MCJ, Sultana J, et al. Different strategies to execute multi-database studies for medicines surveillance in real-world setting: a reflection on the European model[J]. Clin Pharmacol Ther, 2020, 108(2): 228–235. DOI: [10.1002/cpt.1833](https://doi.org/10.1002/cpt.1833).
 - 19 Dedman D, Cabecinha M, Williams R, et al. Approaches for combining primary care electronic health record data from multiple sources: a systematic review of observational studies[J]. BMJ Open, 2020, 10(10): e037405. DOI: [10.1136/bmjopen-2020-037405](https://doi.org/10.1136/bmjopen-2020-037405).
 - 20 Gedeberg R, Igl W, Svennblad B, et al. Federated analyses of multiple data sources in drug safety studies[J]. Pharmacoepidemiol Drug Saf, 2023, 32(3): 279–286. DOI: [10.1002/pds.5587](https://doi.org/10.1002/pds.5587).
 - 21 Hall GC, Sauer B, Bourke A, et al. Guidelines for good database selection and use in pharmacoepidemiology research[J]. Pharmacoepidemiol Drug Saf, 2012, 21(1): 1–10. DOI: [10.1002/pds.2229](https://doi.org/10.1002/pds.2229).
 - 22 曾慧娟, 杨之雨, 柳东红, 等. 医学研究中常见统计分析方法的应用[J]. 上海预防医学, 2023, 35(8): 831–839. [Zeng HX, Yang ZY, Liu DH, et al. Selection and application of statistical methods in medical research[J]. Shanghai Journal of Preventive Medicine, 2023, 35(8): 831–839.] DOI: [10.19428/j.cnki.sjpm.2023.22771](https://doi.org/10.19428/j.cnki.sjpm.2023.22771).
 - 23 汤宁, 宋秋月, 易东, 等. 医学纵向数据建模方法及其统计分析策略[J]. 中国卫生统计, 2019, 36(3): 441–444, 447. [Tang N, Song QY, Yi D, et al. Modeling methods and statistical analysis strategies for medical longitudinal data[J]. Chinese Journal of Health Statistics, 2019, 36(3): 441–444, 447.] [https://d.wanfangdata.com.cn/periodical/ChVQZXJpb2RpY2FsQ0hJMjA5NTA2MjJISD3pnd3N0ajIwMTkwMzAzNhoIZjR1cnJ2ang%3D](https://doi.org/https://d.wanfangdata.com.cn/periodical/ChVQZXJpb2RpY2FsQ0hJMjA5NTA2MjJISD3pnd3N0ajIwMTkwMzAzNhoIZjR1cnJ2ang%3D).
 - 24 Jeong E, Su Y, Li L, et al. Discovering severe adverse reactions from pharmacokinetic drug–drug interactions through literature analysis and electronic health record verification[J]. Clin Pharmacol Ther, 2025, 117(4): 1078–1087. DOI: [10.1002/cpt.3500](https://doi.org/10.1002/cpt.3500).
 - 25 Yang Z, Dehme R M, Yli-harja O, et al. Combining deep learning with token selection for patient phenotyping from electronic health records[J]. Sci Rep, 2020, 10(1): 1432. DOI: [10.1038/s41598-020-58178-1](https://doi.org/10.1038/s41598-020-58178-1).
 - 26 Matheny ME, Yang J, Smith JC, et al. Enhancing postmarketing surveillance of medical products with large language models[J]. JAMA Netw Open, 2024, 7(8): e2428276. DOI: [10.1001/jamanetworkopen.2024.28276](https://doi.org/10.1001/jamanetworkopen.2024.28276).
 - 27 Cannas M, Arpino B. A comparison of machine learning algorithms and covariate balance measures for propensity score matching and weighting[J]. Biom J, 2019, 61(4): 1049–1072. DOI: [10.1002/bimj.201800132](https://doi.org/10.1002/bimj.201800132).
 - 28 Yang C, Kors JA, Ioannou S, et al. Trends in the conduct and reporting of clinical prediction model development and validation: a systematic review[J]. J Am Med Inform Assoc, 2022, 29(5): 983–989. DOI: [10.1093/jamia/ocac002](https://doi.org/10.1093/jamia/ocac002).
 - 29 Yamamoto H, Kayanuma G, Nagashima T, et al. Early detection of adverse drug reaction signals by association rule mining using large-scale administrative claims data[J]. Drug Saf, 2023, 46(4): 371–389. DOI: [10.1007/s40264-023-01278-4](https://doi.org/10.1007/s40264-023-01278-4).
 - 30 Pfaff ER, Girvin AT, Bennett TD, et al. Identifying who has long COVID in the USA: a machine learning approach using N3C data[J]. Lancet Digit Health, 2022, 4(7): e532–e541. DOI: [10.1016/s2589-7500\(22\)00048-6](https://doi.org/10.1016/s2589-7500(22)00048-6).
 - 31 Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods[J]. BMJ, 2024, 385: e078378. DOI: [10.1136/bmj-2023-078378](https://doi.org/10.1136/bmj-2023-078378).
 - 32 van Royen FS, Asselbergs FW, Alfonso F, et al. Five critical quality criteria for artificial intelligence-based prediction models[J]. Eur Heart J, 2023, 44(46): 4831–4834. DOI: [10.1093/eurheartj/ehad727](https://doi.org/10.1093/eurheartj/ehad727).
 - 33 Chen L, Shi B, Feng S. Where medical statistics meets artificial intelligence[J]. N Engl J Med, 2023, 389(25): 2403. DOI: [10.1056/NEJMc2312332](https://doi.org/10.1056/NEJMc2312332).
 - 34 Lei J, Zhai J, Zhang Y, et al. Supervised machine learning models for predicting sepsis-associated liver injury in patients with sepsis: development and validation study based on a multicenter cohort study[J]. J Med Internet Res, 2025, 27: e66733. DOI: [10.2196/66733](https://doi.org/10.2196/66733).
 - 35 Ali S, Akhlaq F, Imran AS, et al. The enlightening role of explainable artificial intelligence in medical & healthcare domains: a systematic literature review[J]. Comput Biol Med, 2023, 166: 107555. DOI: [10.1016/j.combiomed.2023.107555](https://doi.org/10.1016/j.combiomed.2023.107555).

- 36 中国卫生信息与健康医疗大数据学会重症医学分会, 北京肿瘤学会重症医学专业委员会. 重症大数据应用中国专家共识 (2022) [J]. 中华医学杂志, 2023, 103(6): 404-424. DOI: [10.3760/cma.j.cn112137-20221008-02098](https://doi.org/10.3760/cma.j.cn112137-20221008-02098).
- 37 邓思危, 赵厚宇, 詹思延. 基于概况性综述的药物流行病学研究中时间相关性偏倚识别流程 [J]. 中华流行病学杂志, 2024, 45(9): 1273-1282. [Deng SW, Zhao HY, Zhan SY, et al. Identification process of time-related bias in pharmacoepidemiologic research based on a scoping review[J]. Chinese Journal of Epidemiology, 2024, 45(9): 1273-1282.] DOI: [10.3760/cma.j.cn112338-20240123-00037](https://doi.org/10.3760/cma.j.cn112338-20240123-00037).
- 38 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).
- 39 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).
- 40 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).
- 41 Suissa S, Dell'aniello S, Vahey S, et al. Time-window bias in case-control studies: statins and lung cancer[J]. Epidemiology, 2011, 22(2): 228-231. DOI: [10.1097/EDE.0b013e3182093a0f](https://doi.org/10.1097/EDE.0b013e3182093a0f).
- 42 Lieslehto J, Tiihonen J, Lähteenvuoto M, et al. Comparative effectiveness of pharmacotherapies for the risk of attempted or completed suicide among persons with borderline personality disorder[J]. JAMA Netw Open, 2023, 6(6): e2317130. DOI: [10.1001/jamanetworkopen.2023.17130](https://doi.org/10.1001/jamanetworkopen.2023.17130).
- 43 许璐, 赵厚宇, 詹思延. 药物流行病学研究中的时间窗偏倚[J]. 药物流行病学杂志, 2019, 28(12): 828-832. [Xu L, Zhao HY, Zhan SY. Time-window bias in pharmacoepidemiology Study[J]. Chinese Journal of Pharmacoepidemiology, 2019, 28(12): 828-832.] DOI: [10.19960/j.cnki.issn1005-0698.2019.12.010](https://doi.org/10.19960/j.cnki.issn1005-0698.2019.12.010).
- 44 Klungel OH, Martens EP, Psaty BM, et al. Methods to assess intended effects of drug treatment in observational studies are reviewed[J]. J Clin Epidemiol, 2004, 57(12): 1223-1231. DOI: [10.1016/j.jclinepi.2004.03.011](https://doi.org/10.1016/j.jclinepi.2004.03.011).
- 45 Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology[J]. Epidemiology, 2000, 11(5): 550-560. DOI: [10.1097/00001648-200009000-00011](https://doi.org/10.1097/00001648-200009000-00011).
- 46 黄丽红, 魏永越, 陈峰. 如何控制观察性疗效比较研究中的混杂因素: (一) 已测量混杂因素的统计学分析方法[J]. 中华流行病学杂志, 2019, 40(10): 1304-1309. [Huang LH, Wei YY, Chen F. Confounder adjustment in observational comparative effectiveness researches: (1) statistical adjustment approaches for measured confounder[J]. Chinese Journal of Epidemiology, 2019, 40(10): 1304-1309.] DOI: [10.3760/cma.j.issn.0254-6450.2019.10.024](https://doi.org/10.3760/cma.j.issn.0254-6450.2019.10.024).
- 47 Sabaté M, Montané E. Pharmacoepidemiology: an overview[J]. J Clin Med, 2023, 12(22): 7033. DOI: [10.3390/jcm12227033](https://doi.org/10.3390/jcm12227033).
- 48 Uddin MJ, Groenwold RH, van Staa TP, et al. Performance of prior event rate ratio adjustment method in pharmacoepidemiology: a simulation study[J]. Pharmacoepidemiol Drug Saf, 2015, 24(5): 468-477. DOI: [10.1002/pds.3724](https://doi.org/10.1002/pds.3724).
- 49 European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials[EB/OL]. (2020-07-30) [2025-06-13]. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guideline-statistical-principles-clinical-trials-step-5_en.pdf.
- 50 Cheng DM, Hogan JW. The sense and sensibility of sensitivity analyses[J]. N Engl J Med, 2024, 391(11): 972-974. DOI: [10.1056/NEJMp2403318](https://doi.org/10.1056/NEJMp2403318).
- 51 国家药品监督管理局药物审评中心. 药物临床试验亚组分析指导原则 [EB/OL]. (2020-12-31) [2025-06-13]. <https://www.cqn.com.cn/ms/att/2021-02/09/f8fcb4fa-a5e3-4a47-9e16-3ce3a7b4f7ca.pdf>.
- 52 Delgado-Herrera L, Anbar D. A model for the interim analysis process: a case study[J]. Control Clin Trials, 2003, 24(1): 51-65. DOI: [10.1016/s0197-2456\(02\)00275-1](https://doi.org/10.1016/s0197-2456(02)00275-1).
- 53 Walter SD, Han H, Guyatt GH, et al. A systematic survey of randomised trials that stopped early for reasons of futility[J]. BMC Med Res Methodol, 2020, 20(1): 10. DOI: [10.1186/s12874-020-0899-1](https://doi.org/10.1186/s12874-020-0899-1).
- 54 郭洋, 蒋志伟, 夏结来, 等. 肿瘤药物临床试验中成组序贯设计的统计学考虑 [J]. 中国临床药理学与治疗学, 2013, 18(2): 175-180. [Guo Y, Jiang ZW, Xia JL, et al. Statistical considerations of group sequential designs in oncology drug clinical trials[J]. Chinese Journal of Clinical Pharmacology and Therapeutics, 2013, 18(2): 175-180.] https://kns.cnki.net/kcms2/article/abstract?v=ZscdH8NaPi8y1GIKMfij9MwC7XO8STdcOmO6ANRJKrEly8dyKhBY8Him__damh7TSLFTLqylwJGTicq_ZxPmxPJGm2MrMcNwD-xCY7T1feKLkBI4ACrxd8gS3uEEBzA_KXu5gQiaUty-p0fCZRPqwG0ZSj8CjGzw2SUnVs5CEHhPhQ30TM7Q==&uniplatform=NZKPT&language=CHS.
- 55 黄丽红. 临床试验期中分析中的两个统计学问题 [D]. 南京: 南京医科大学, 2017.

收稿日期: 2025 年 08 月 06 日 修回日期: 2025 年 08 月 27 日
 本文编辑: 沈静怡 周璐敏