

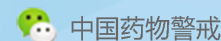


国家药品监督管理局药品评价中心

Center for Drug Reevaluation, NMPA

国家药品不良反应监测中心

National Center for ADR Monitoring, China



请输入要查询的内容



关于发布《真实世界数据支持药品安全性主动监测的一般原则》的通知

发布日期：2025-08-19

为规范和促进真实世界证据在药品安全性评价中的应用，根据国家药品监督管理局工作部署，评价中心组织制定了《真实世界数据支持药品安全性主动监测的一般原则》（见附件）。按照《国家药监局综合司关于印发药品技术指导原则发布程序的通知》（药监综药管〔2020〕9号）要求，经国家药品监督管理局审查同意，现予发布，自发布之日起施行。

附件：[真实世界数据支持药品安全性主动监测的一般原则](#)



国家市场监督管理总局



国家药品监督管理局

国家局直属单位



网站地图 | 联系方式

地址：中国，北京市经济技术开发区广德大街22号院一区4号楼 邮编：100076

本站由国家药品监督管理局药品评价中心主办 版权所有 未经许可禁止转载或建立镜像

Copyright 2004-2023 国家药品监督管理局药品评价中心

All rights reserved. 京ICP备16043083号-3 京公网安备11010502038932 号

附件

真实世界数据支持药品安全性主动监测的一般原则

为促进真实世界数据在药品安全性评价中的应用，指导药品上市许可持有人（以下简称“持有人”）规范开展药品安全性主动监测研究，制定本指导原则。

本指导原则中“药品安全性主动监测”（以下简称“主动监测”）是指采取持续有组织的方式，主动收集一个或多个来源的数据，进而利用数据发现未知风险信号，评价新发现的风险信号，描述药品的安全特征、评估药品风险的过程。

本指导原则仅代表当前的观点和认识，随着药品监管科学研究的不断深入，本指导原则中的相关要求及内容也将不断完善。

一、主要目的

主动监测常用于研究以下药品安全性问题：（1）是否存在新的风险信号；（2）发现的风险信号是否构成新的安全风险；及/或（3）新的安全风险的特征及影响因素等。主要包括：

信号发现（信号检测）：利用主动监测数据，发现或识别药品与某个或某类不良事件之间的新的潜在因果关系，或两者之间已知关联的新变化（如不良反应的频率、持续时间、严重程度等改变）。

信号评价：利用主动监测数据及其他可获得的数据资源对风险信号进行评价，综合判断信号是否已构成新的药品安全风险。

风险评估：利用主动监测获取的数据及其他相关信息对“新的药品安全风险”开展综合评估，描述其性质、频率和严重程度，了解风险的特征和影响因素等。

二、数据来源及类型

基于上市药品安全性评价目的、数据质量及可及性等因素，主动监测可收集不同来源、类型的真实世界数据。根据数据形成的原因，可分为研究特定安全性问题形成的数据、常规诊疗过程形成的数据、其他研究/管理目的形成的数据。

（一）研究特定安全性问题形成的数据

基于预先设定的研究目的，通过制定统一的变量定义标准和收集模式，获取与特定药品安全性问题相关的数据，获取方法包括但不限于哨点监测、药物/处方事件监测、登记等。

1. 哨点监测

指在固定地点（哨点）通过查阅医疗记录或访谈医生、患者等形式，从现场获取更为完整、全面的药品不良反应/事件数据的方法。哨点监测可以获取自发报告方式难以得到的数据，如某些特殊疾病患者亚群的数据。此外，在特定医疗机构还可以收集药物滥用等信息。

2. 药物/处方事件监测

采用观察性队列研究方法，不干涉医师的诊疗行为，根据处方数据确认每个患者，向处方医生或患者发送随访调查表，并获取患者用药与结局信息，如人口统计学信息、适应症、用药持续时间、剂量、临床事件、停药原因等。其优点一是可在一定程度降低选择性报告的影响，二是可识别其他方法难以识别的药品不良反应，包括潜伏期较长的不良反应。其局限性是医师或患者的应答率低、失访率较高、收集的信息种类分散，可能导致重要信号不明确。

3. 登记

采用有组织的系统和观察性研究方法，收集使用某种特定药品（药品登记）或患有特定疾病（疾病登记）人群的特定结局发生过程的数据。登记数据在评估罕见疾病人群、罕见病治疗药物或特殊人群风险信号时具有一定优势。同时，通过与已有外部数据（处方数据库、国家癌症注册数据、死亡登记等）链接，可丰富登记数据库的暴露、协变量、效应修饰因子、结局信息。

（二）常规诊疗过程形成的数据

医院电子病历系统数据（**Electronic Medical Record, EMR**）、医保数据、区域健康医疗数据等均为常规诊疗过程形成的数据，可用于支持主动监测。此类数据包括但不限于以下类型：

1. 医院电子病历系统数据

患者在医疗机构接受诊疗服务过程中所产生的记录，包括患者人口学特征、诊断、医嘱、实验室检验、影像学检查等数据，通常以结构化及/或非结构化的形式呈现。

2. 医保数据

各级医保部门建立的医保险赔数据库，涵盖参保人群基本信息、门诊和住院疾病诊断、处方用药、诊疗项目、结算等结构化字段的数据。

3. 区域健康医疗数据

整合利用某个区域内多来源真实世界数据，包括医疗机构、医保及其他来源的公共卫生数据，有助于获取该区域更多患者、更加全面的药品不良事件信息，可弥补单一来源数据人群代表性相对不足等问题。

4. 居民电子健康档案

居民接受医疗卫生服务过程中产生的电子化记录，是以居民个人健康为核心、贯穿整个生命过程、涵盖各种健康相关因素的系统化数字文档记录，包括个人基本信息、健康体检记录、医院诊疗记录、重点人群健康管理及其他卫生服务记录等。

（三）其他研究/管理目的形成的数据

因其他研究/管理目的形成的真实世界数据，如包含药品安全性评价所需的信息，也可用于主动监测，实现数据的分析利用。此类数据包括但不限于以下类型：

1. 死亡登记数据库

政府或相关机构用于记录公民死亡信息的数据库，详细记录了与死亡事件相关的信息，包括死亡时间、地点、死因、年龄、性别等。

2. 公共卫生监测数据库

由政府部门或公共卫生机构建立关于人群健康状况和

公共卫生事件的数据库系统，如慢性非传染病监测、免疫接种不良事件监测等，通过连续、系统收集疾病或其他公共卫生事件信息，了解其分布和流行趋势。

3. 其他数据

因其他目的开展研究形成的数据，例如因其他安全性问题而开展的药品上市后研究、主动监测等形成的数据；以及可穿戴设备、移动 APP/设备等途径收集的健康数据等，若包含本次安全性评价所需的信息，也可作为主动监测的数据来源。

三、适用范围

实施主动监测收集的真实世界数据及产出的真实世界证据，可用于发现未知风险信号，评价新发现的风险信号，描述药品的安全特征、评估药品风险等，以及用于支持修改药品说明书、识别高风险人群、指导临床合理用药等多种决策。真实世界数据支持主动监测的主要适用范围如下：

（一）发现可疑风险信号

通过对各种途径收集得到的医院电子病历系统数据、医保数据等真实世界数据，借助统计模型、人工智能等大数据挖掘技术开展信号检测，可及时发现新的药品安全风险。需重点关注的可疑信号主要涉及：

1. 新发现的非预期、严重不良事件或已知不良事件的严重程度明显增加或频率显著升高；
2. 新的药物相互作用（包括：药物-药物、药物-食物、药物-医疗器械等）；
3. 涉及新的潜在高风险人群等。

（二）评价新发现的风险信号、评估药品风险

针对新发现的风险信号，获取药品上市后临床使用形成的真实世界数据，可用于评价新发现的风险信号、评估药品风险，具体包括：

1. 明确药品与不良事件的因果关系，识别风险信号是否由目标药品导致。

2. 量化潜在的或已识别的药品风险。如计算潜在的或已识别的药品不良事件发生率，描述其严重程度、持续时间、人群分布等特征。若为已确认的风险，还可与对照人群进行比较获得相对危险度或率差。

3. 描述药品风险特征、分析影响因素等。分析可能引起药品风险、增加风险发生频率或严重程度的危险因素，包括效应修饰因子。

4. 评估罕见及/或迟发的不良反应风险。通常上市前临床研究纳入的患者例数有限、随访时间较短，往往难以发现罕见及/或迟发的药品不良反应。药品上市后，不断累积形成的大样本真实世界数据具有纳入人群广泛、随访时间长的特点，可用于相关风险的评估。

（三）评估特殊人群用药风险

由于伦理等问题，孕妇、儿童、老年人、肝或肾功能不全等人群一般不作为上市前临床试验的受试者，其用药的安全性常缺乏上市前临床研究数据/证据支持。实施主动监测收集特殊人群在临床诊疗过程中产生的真实世界数据，可用于评估其临床用药风险。

（四）评估药物相互作用风险

在临床实践中，常存在两种或两种以上药物同时或序贯使用的情形。不同药物之间可能存在药效或毒性作用的协同、相加或拮抗，以及药动学的相互作用，从而影响药物治疗效果或增加用药风险。上市前临床试验对联合用药的研究较为有限，难以发现或评价潜在的药物相互作用。基于主动监测收集真实世界数据，可用于对药物与药物的相互作用风险进行评价。同样可用于评价药物与食物、药物与医疗器械等的相互作用风险。

（五）描述药源性疾病流行情况

药物性心、肝或肾等重要脏器损害以及药物导致的严重血液系统损害、严重过敏反应等药源性疾病，影响患者临床用药安全，甚至可能危及生命。实施主动监测可获得更为全面、详尽的不良事件数据，可用于描述其发病率、患病率、地域分布、时间趋势等流行情况，为药品安全监管和临床安全用药提供参考。

（六）对风险控制措施开展后效评估

针对已确认的安全风险，基于主动监测数据对采取药品风险控制措施前后的不良反应发生率、严重程度等进行比较，为评估药品风险控制措施的有效性提供依据。

四、实施流程

针对需研究的药品安全性问题，首先应评估自发报告及其他已有数据及资料（文献报道、上市前研究及已有上市后研究数据及结果）是否可以满足要求；当已有数据及资料不足时，可开展主动监测进一步获取真实世界数据。

主动监测的实施流程一般包括：明确研究目的；制定

研究方案；确定适宜的真实世界数据来源和类型；获取主动监测数据；开展数据治理/数据管理，构建研究数据集；基于研究方案进行统计分析，开展信号检测、信号评价、风险评估，制订风险管理措施等（见下图）。

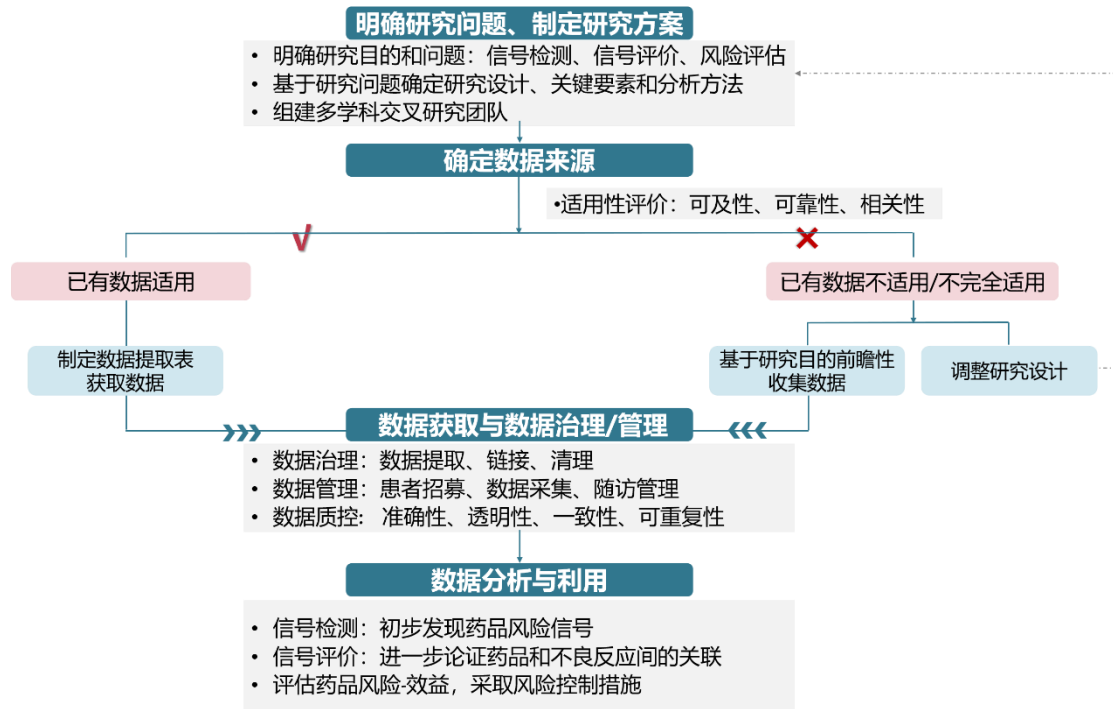


图 药品安全性主动监测的一般流程

（一）明确研究目的

主动监测数据可用于发现未知风险信号，评价新发现的风险信号，描述药品的安全特征、评估药品风险等。基于主动监测数据开展药物流行病学研究，可产出用于支撑监管决策的证据。

（二）制定研究方案

主动监测通常需要组建多学科交叉的研究团队，涵盖如临床医学、药学、药物流行病学及生物统计等领域专家。研究方案应重点明确研究变量定义和统计分析计划。基于

不同的研究问题，所需的研究变量和关键数据要素不同。对于信号检测，通常需要至少确定药物暴露和不良结局相关信息；而信号评价、风险评估可采用传统药物流行病学方法，在研究方案中应明确拟使用的流行病学设计类型、研究关键变量（研究人群、目标药品、对照、结局事件、关键协变量）及其对应的数据要素。

（三）确定数据来源/类型

针对研究方案中所确定的关键数据要素，开展数据适用性评估，确定拟采用的真实世界数据来源/类型（例如医院电子病历系统或医保数据等）及获取方式。

不同的数据源和数据类型的质量和所涵盖的数据元素不同，应就具体研究问题对数据是否适用进行评估，根据评估结果选择合适的数据源、数据类型。

数据的适用性评估主要包括数据可及性、相关性和可靠性等。一般而言，考虑到研究成本和可行性，优先评估已有数据（如医院电子病历系统数据、医保数据或已有的患者登记数据等）对当前研究的适用性。若已有数据缺乏主动监测所需关键变量或质量未达到研究需求，可以考虑进行前瞻性数据收集，或适当调整研究设计。

（四）数据获取与数据治理/管理

针对不同来源和类型的真实世界数据，数据获取与数据治理/管理方式不同。

1. 研究特定安全性问题形成的数据

对于研究特定安全性问题形成的数据，此类数据通常需前瞻性收集，在收集过程中同步实施数据管理以保证数

据质量。数据收集及管理内容一般包括：制作病例报告表明确拟收集的数据内容；制定标准化的数据收集手册如变量定义、测量方式与时间、数据核查流程等；制作数据收集系统，确定合适的录入方式；定期开展数据核查、疑问数据反馈等数据管理等。

2. 常规诊疗过程形成的数据

对于常规诊疗形成的诊疗数据，可以采用人工摘录或电子化等方式，基于明确的数据提取方案获取。数据提取方案应包含拟提取变量及其存储模块。针对提取到的数据，研究者应开展完整性和准确性核查。

常规诊疗过程形成的数据获取可考虑两种方式：一是针对具体的安全性问题从医院电子病历系统数据提取特定药品医嘱处方记录及特定诊断、临床检测异常值等相关数据，用于药品安全性评价。二是不针对具体安全性问题，综合考虑一类或多类安全性问题，将提取到的相关数据治理后构建研究型数据库；后续针对具体安全性问题，构建数据集，二次筛选数据开展研究。

应注意，此类数据的产生并非基于研究问题，通常无法直接进行分析，还需经过数据治理将原数据转化为适用于当前主动监测研究的数据库或数据集。数据治理过程主要包括数据链接、数据标准化（制定变量字典）、数据结构化（文本信息结构化）规则以及数据处理（异常值、极端值、矛盾数据等）。开展数据治理时应注意个体隐私保护和数据安全，并遵循准确性、透明性、一致性和可重复性等原则。

3. 其他研究/管理目的形成的数据

此类数据来源较多，基于不同来源的研究/管理目的形成的数据，其数据获取和治理/管理方式不同。通常死亡登记数据库、公共卫生监测数据库的获取与治理可参照常规划诊疗过程形成的数据；如采用其他研究目的形成的数据，则应根据当前研究目的重新进行信息提取以及必要的数据治理。

此外，如单一数据来源不能满足安全性评价的需求，需从两个或多个数据源获取相关数据，经治理后用于主动监测。此时应注意多源数据间的是否可以有效链接并整合运用。其次，还应针对数据治理过程开展质量评价，包括但不限于是否有清晰明确、合理的数据清理规则及流程、变量字典的合理性、文本结构化结果验证准确性等。数据治理过程的质量评价内容可参考已发布的相关指导原则。

（五）数据分析与利用

基于所构建的研究型数据集，遵循预先制定的统计分析计划，开展统计分析。在数据分析过程中，如果违背了原有计划，应说明原因并报告调整的内容，保障研究的清晰透明和可重复性。

实施主动监测收集的真实世界数据及产出的真实世界证据，可用于发现未知风险信号，评价新发现的风险信号，描述药品的安全特征、评估药品风险等，以及用于支持修改药品说明书、识别高风险人群、指导临床合理用药等多种决策。具体可见本指导原则第三章“真实世界数据支持主动监测的适用范围”。

五、其他考虑

（一）对数据的一般要求

实施主动监测时，应首先明确需获取的真实世界数据要素并评估其适用性。

1. 数据要素

研究所需数据要素的确定，应基于具体的安全性问题，遵循必要和最小化原则。一般而言，可考虑以下方面：可疑药品和其他药品使用信息（剂量、给药途径、疗程等），用药人群和对照人群的人口学信息、不良事件信息、其他治疗、合并疾病、检验检查及生命体征等信息。涉及中药的信号发现或评价，需考虑是否要纳入中药相关信息如：患者病证信息、中药处方构成、剂型、用法用量、中医临床结局指标等。

在保证研究结果有效性和可靠性的情况下，根据研究问题和数据可及性，选择性地收集必要信息，可避免资源的浪费，加快真实世界证据的产生。

2. 数据适用性评估

应考虑不同研究目的对数据的颗粒度、维度及质量等的要求，合理设定数据适用性评估的标准。信号评价、风险评估对数据的颗粒度、维度及质量等的要求通常高于信号检测。数据适用性评估一般可考虑以下因素：

（1）人群代表性

数据库应尽可能覆盖实际临床中的药品使用人群。不同地区、医院级别、诊疗模式、医保政策、数据来源等因素均可能影响数据库涵盖患者人群的特征，从而影响主动

监测结果的代表性。一般而言，医疗机构尤其是单一医疗机构的电子病历系统数据仅覆盖相应医疗机构就诊患者，其代表性通常不及整合多个地区不同级别医疗机构的多中心数据库或区域健康医疗数据。

（2）数据完整性

数据完整性是指评估数据库是否涵盖主动监测所需的关键信息及其缺失比例，如患者人口学信息、药物暴露、不良事件、随访、重要混杂变量等。通常来说，医疗机构的电子病历系统较详尽地记录了患者院内用药的安全性相关信息，但缺乏院外用药信息，且难以获得患者随访数据，在评估迟发型不良事件时存在一定局限性。医保数据库较完整地记录了参保患者的就医行为，同时也涵盖患者院外自购报销药品信息，但通常缺乏用药时间、用药剂量以及患者症状、体征、检验检查结果等详细临床信息。

（3）数据的准确性

真实世界数据特别是常规诊疗过程形成的数据，由于数据的产生并非基于研究目的，数据的准确性受医生、患者、数据库管理者等多方影响，包括原始数据是否准确、编码映射关系是否对应且唯一、数据值域是否合理等。研究者应评估药品暴露、不良事件等关键变量的准确性，例如计算药品或疾病编码的灵敏度、特异度等验证指标。对中药开展主动监测，应特别注意中医方药、病证及临床诊疗术语定义和识别标准的准确性。

（4）是否能获得可比对照人群

当评价不良反应与用药因果关系时，能否从数据库中

获得可比的对照人群也是重要因素。通常考虑使用与待评价药品有相同适应症的同类药品、疾病病程相近的患者作为对照。例如，利用仅涵盖待研究药品用药人群的医院集中监测数据或专药登记数据库，可描述不良反应特征、计算药品不良反应发生率，但因缺乏非用药人群或对照药品使用人群信息，无法计算相对风险，难以探索用药与不良事件的因果关系。

(5) 不良事件数和数据库覆盖时长

当结局为罕见事件时，应评估是否能从数据库中获得足够数量的结局事件，以达到研究结果的预期精度；当评价长期用药或潜伏期较长的药品不良反应时，还应考虑数据库覆盖时长是否足以观察到不良事件的发生，以确保药品与潜在不良反应的关联合理性。

此外，通过链接多源真实世界数据或开展必要前瞻性随访进一步收集数据，可改善数据的适用性。多来源的数据进行相互补充和验证，可提高研究结果的证据强度。

(二) 用于信号检测的考虑

信号检测的核心是从大量用药信息中识别出潜在可疑不良反应风险信号。常见信号检测方法主要有比例失衡法、处方序列对称分析方法（Prescription Sequence Symmetry Analysis, PSSA），序贯统计检验方法（Sequential Probability Ratio Test, SPRT），有监督的机器学习方法和树状扫描统计量（Tree-Based Scan Statistic, TreeScan）等。

比例失衡法可分为频数法、贝叶斯法，常用的频数法包括：比例报告比法（Proportional Reporting Ratio, PRR）、

报告比值比法（Reporting Odds Ratio, ROR）、综合标准法等；常用的贝叶斯方法包括：世界卫生组织乌普萨拉监测中心（WHO-UMC）应用的信息成分法（Information Component, IC），以及美国 FDA 采用的多项伽马泊松分布缩减法（Muti-Item Gamma Poisson Shrinker, MGPS）等。

在信号检测过程中，药品-事件的关联性可能受到混杂因素干扰，可考虑采用回归模型、倾向性评分（Propensity Score, PS）匹配、疾病风险评分（Disease Risk Score, DRS）等方法进行混杂控制。针对罕见不良反应信号，通常考虑获取更大样本量的主动监测数据，以及用人工智能技术和更敏感的分析方法。

（三）用于信号评价/风险评估的考虑

信号评价的目的是确认用药与不良事件发生的因果关系。风险评估的主要目的是识别、定性或定量描述药品安全风险，分析药品安全性特征，以及评估风险控制措施实施效果。

1. 流行病学设计

基于主动监测数据开展信号评价/风险评估，可采用观察性流行病学研究设计，包括队列研究、病例对照研究及其衍生设计类型等。

设计类型的选择取决于具体的研究问题。研究设计阶段应合理选择并定义研究人群、暴露、对照、不良事件与混杂等研究关键变量，从设计阶段尽量控制潜在的偏倚与混杂，包括但不限于：①优选新发病例或新用药者设计，可有效控制选择性偏倚、非死亡时间偏倚；②选择可比对

照，尽量选择与待评价药品相同适应症、病情程度相近的阳性对照，以减少适应症偏倚；③定义研究设计相关时间点，如研究对象基线评估时间窗，药物相关的暴露期、诱导期、洗脱期，随访起止时间等。建议研究者绘制研究时间轴，全面、直观地展示研究设计细节；④考虑并明确定义随访期可能出现的用药状态变化，包括换药、停药、加载等；诱导期、洗脱期，随访起止时间等。建议研究者绘制研究时间轴，全面、直观地展示研究设计细节。

2. 统计分析

研究者应基于研究目的、数据适用性评估结果、研究设计类型等，提前制定统计分析计划（**Statistical Analysis Plan, SAP**），不应根据纳入数据的情况改变统计分析计划。统计分析计划应包括但不限于：样本量计算、分析数据集定义、干预/暴露与对照定义、已知和未知偏倚混杂的识别与控制、结局变量定义、缺失数据处理方法、主分析模型、假设检验水平、检验效能、亚组分析、敏感性分析、定量偏倚分析等内容。

统计分析过程，应重点考虑如何有效地控制和评价偏倚对结果的影响，从而准确评估药品风险信号。常用的统计模型包括但不限于 **Logistic** 回归、**Poisson** 回归、倾向性评分方法、疾病风险评分、工具变量等。不同的统计分析模型适用条件不同，优势和局限性存在差异。应基于拟回答的研究问题和数据特征，选择合适模型处理潜在偏倚。基于真实世界数据的主动监测，由于患者病情变化和依从性等问题，常存在多药联用和用药变化（例如换药、停药、

加载等)等复杂用药模式,特别在中西医结合治疗场景中,中西药联合使用,且用药方案动态调整。此时统计模型选择应考虑如何处理多药联用和用药变化引入的高维时间依赖性混杂。另外,还应注意诸如选择性偏倚、缺失数据、残余混杂等其他观察性研究的常见偏倚。一般而言,除主分析外,应考虑采用定量偏倚分析、敏感性分析等方式评估潜在偏倚对结果的影响,探讨结果的稳健性。对于罕见不良反应的评价,还需特别关注结局事件数量对模型结果估计的预期精度的影响。

3. 其他需考虑的因素

信号检测得到的信号,仅说明用药与不良事件之间的因果关系具有潜在的可能性,尚不能明确风险真实存在。

基于主动监测数据开展信号评价,需结合信号的合理性、报告的可信度、不良事件的严重程度与临床预后等因素,以及其他相关数据及资料。

基于主动监测数据开展风险评估,需综合考虑证据质量,并充分结合现有其他相关证据,评估风险的严重性、发生频率、可预测性、可预防性和可逆性;其他同类药品的可获得性、可替代性、支持关联性的其他证据(生物学合理性、关联一致性、剂量反应关系等)等。

附:词汇表

词汇表

1. 真实世界数据（Real-World Data, RWD）：来自真实医疗环境，反映实际诊疗过程中患者健康状况和医疗服务过程的数据。
2. 真实世界证据（Real-World Evidence, RWE）：利用真实世界数据开展真实世界研究得到的用于临床和医疗决策的结果。
3. 常规诊疗形成的健康医疗数据：基于临床或管理目的收集的健康医疗数据，包括医院电子病历数据、医疗保险数据等。
4. 统计分析计划（Statistical Analysis Plan, SAP）：对临床研究的统计学考虑及拟对数据进行统计分析的清晰描述。
5. 混杂（Confounding）：指某一因素与结局事件和暴露均有关，但不是暴露与结局因果链上的中间变量。若混杂在组间分布不均衡，可能会歪曲（缩小或扩大）暴露与结局间的真实联系。
6. 倾向性评分（Propensity Score）：是指一种针对多个协变量进行调整来处理观察性研究中组间协变量不均衡的降维分析策略。
7. 工具变量（Instrumental Variable）：与内生解释性变量高度相关，但与随机误差项无关，对结局变量的影响只能通过影响处理因素实现的变量，可控制已知和未知的混杂因素。