浙江省药品监督管理与产业发展研究会

关于公开征求《中药真实世界研究数据治理规范》团体标准意见的函

由浙江省药品化妆品审评中心等单位起草的《中药真实世界研究数据治理规范》团体标准已完成征求意见稿的编制，根据《浙江省药品监督管理与产业发展研究会团体标准管理办法（试行）》的相关规定，为保证标准的科学性、严谨性和适用性，现就该项团体标准公开征求意见。

请各有关单位和相关专家审阅标准全文并提出宝贵建议和意见，并于2025年2月22日前以邮件的形式将《团体标准征求意见表》反馈至联系人，感谢您对我们工作的支持！

联系人：胡婷霞　　联系电话：0571-81053259

邮箱：hutingxia@zjcdce.org.cn

附件：《团体标准征求意见表》

浙江省药品监督管理与产业发展研究会

2025年1月22日

附件

团体标准征求意见表

标准名称	《中药真实世界研究数据治理规范》				
专家姓名		所在单位			
通信地址				邮编	
联系电话		电子邮箱		提出意见日期	
标准条款	提出意见及建议内容			备注	

（如填写内容较多，可另加附页）

（注：本表填写后请于 2025 年 2 月 22 日前以电子邮件形式反馈，标准起草单位相关电子邮箱为 hutingxia@zjcdce.org.cn）

ICS 65.020.20
C 05

团　　体　　标　　准

T/CACM ****－20**

中药真实世界研究数据治理规范

Specifications for the Management of Real-World Research Data on Traditional Chinese Medicine

20**-**-**发布　　20**-**-**实施

浙江省药品监督管理与产业发展研究会　发布

目　次

前言 II
1 范围 1
2 规范性引用文件 1
3 术语及定义 2
3.1 真实世界数据（Real-world Data ，RWD） 2
3.2 真实世界证据（Real-world Evidence，RWE） 2
3.3 元数据（MetaData） 2
3.4 通用数据模型（Common Data Model） 2
4 数据治理标准化体系 2
4.1 数据元标准 2
4.2 数据集标准 3
4.3 接口标准 3
4.4 元数据标准 3
4.5 元数据注册管理标准 3
4.6 数据表示标准 3
4.7 数据访问标准 3
5 数据治理 4
5.1 个人信息保护和数据安全性处理 4
5.2 数据提取 4
5.3 数据清洗 5
5.4 数据转化 6
5.5 数据传输和存储 6
5.6 数据质量控制 6
5.7 通用数据模型 7
5.8 数据治理计划书 7
6 治理过程的质量管理 8
6.1 数据质量管理流程 8
6.2 数据质量管理的目标 8
6.3 影响数据质量的相关环节分析 8
6.3.1 数据产生 8
6.3.2 数据集成 9
6.3.3 数据使用 9
7 经治理的中药真实世界数据适用性评价 10
7.1 相关性评价 10
7.2 可靠性评价 10
参考文献 11

前　言

本文件按照 GB/T1.1-2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

本文件由浙江省药品化妆品审评中心提出。

本文件由浙江省药品监督管理与产业发展研究会归口。

本文件起草单位：浙江省药品化妆品审评中心、杭州泰格医药科技股份有限公司、杭州莱迈医疗信息科技有限公司、浙江省中医院、正大青春宝药业有限公司、杭州和泽远志中药科技有限公司。

本规范主要起草人：

中药真实世界研究数据治理规范

1 范围

本规范提出了中药真实世界研究中数据治理的方法、步骤和技术要求，包括：数据标准化体系、数据治理流程、数据评估、数据质控、数据评价。可供从事中药真实世界研究使用。

2 规范性引用文件

下列文件对于本规范的应用是必不可少的。凡是注明日期的引用文件，仅所注明日期的版本适用于本规范。凡是不注明日期的引用文件，其最新版本（包括所有的修改版本）适用于本规范。

ISO/IEC 38505-1：2017 信息技术 - IT 治理 - 数据治理 第 1 部分：ISO/IEC 38500 在数据治理中的应用

GB/T 34960.5-2018 信息技术服务 治理 第 5 部分：数据治理规范

DB50/T 1452-2023 工业数据治理规范

《医疗服务基本数据集》（WS 373-2012）

《电子病历基本数据集》（WS 445-2014）

OHDSI 通用数据模型临床数据通用模型

《中医病证分类与代码》

《中医临床诊疗术语》

《疾病分类与代码》（GB/T 14396-2016）

《用于产生真实世界证据的真实世界数据指导原则（试行）》

《中成药真实世界研究技术指导原则》（T/CATCM 011—2021）

疾病分类代码标准（ICD-9、ICD-10）

中国卫生信息数据元值域代码 WS364.X--2011

WS370-2012 卫生信息基本数据集编制规范

《计算机信息系统安全等级保护相关规范》

3 术语及定义

下列术语和定义适用于本《规范》。

3.1 真实世界数据（Real-world Data ，RWD）

指来自真实医疗环境的患者健康状态相关数据和（或）不同来源的日常医疗健康保健数据。

3.2 真实世界证据（Real-world Evidence，RWE）

指真实世界数据分析得到的干预措施的使用和潜在获益或风险相关的临床证据。

3.3 元数据（MetaData）

元数据是描述数据的数据，它提供了关于数据特征、内容、结构、来源、关系以及数据如何被使用和管理的信息。

3.4 通用数据模型（Common Data Model）

通用数据模型是一种抽象化的数据结构，用于描述特定领域或业务过程中的数据实体、属性以及它们之间的关系。通用数据模型旨在提供一种标准化的方式来组织和表示数据，以便不同系统和应用程序之间能够共享和理解数据。

4 数据治理标准化体系

4.1 数据元标准

明确了系统中最小的数据单元。它是构成业务知识的基础，通过数据元的规范，可以明确项目中的各类对象和每一个对象的特性，并且在此基础上进一步明确了对象的特性的表示。

4.2 数据集标准

是在特定主题下，为了满足业务信息系统规范化建设和业务邻域内部以及领域间数据交换与共享的需求，设计归纳的各个子系统（或者功能模块）所包含的最小数据元素的集合。卫生信息数据集标准规定了内容结构、元数据、数据元属性、数据元索引表达方法。

4.3 接口标准

为了满足获取数据、推送数据的需求。获取数据是上层应用自身需求发起的，为了满足宏观业务的实现。上报数据是下层应用的需求发起的，为了满足微观业务的流转。

4.4 元数据标准

是用于描述其他业务数据的属性数据，根据业务数据涉及的内容包括：数据元、代码、分类、服务、流程等方面，为上层制定业务标准时对于业务知识的表达形成方法和技术的支持。

4.5 元数据注册管理标准

通过提供对元数据的注册和管理服务系统可以把业务知识有效的组织和管理起来，明确规范其他系统对于整个服务中业务知识的访问、检索、提交、更新的技术接口规范。

4.6 数据表示标准

业务数据在不同的系统间进行交换时需要遵循统一的表示规则，可保证多源异构系统间对于业务信息保持统一理解，是实现信息资源交换的重要内容之一。

4.7 数据访问标准

为数据的提供方和使用方明确了各自的行为模式和结果数据格式。数据的使用者无需关心数据提供者提供的数据存储在哪里、采用什么数据库存储等这些基本的、关键的问题。而作为数据的提供方也无需为每一个使用方根据不同的要求开发不同的数据使用的服务。

5 数据治理

5.1 个人信息保护和数据安全性处理

真实世界研究涉及个人信息保护应遵循国家信息安全技术规范、医疗大数据安全管理相关规定，对个人敏感信息应进行去标识化（de-identification）处理，确保根据数据无法进行个人敏感信息匹配还原，通过技术和管理方面的措施，防止个人信息的泄漏、损毁、丢失、篡改。

数据安全性处理应基于中药研究所涉及的各种数据的类型、数量、性质和内容，尤其对于个人敏感信息，建立数据治理各环节的数据加密技术要求、风险评估和应急处置操作规程，并开展安全措施有效性审计。

5.2 数据提取

根据中药研究源数据的存储格式、是否为电子数据、是否包含非结构化数据等因素选择合适的方式进行数据提取，在数据提取时均应遵守以下原则：

1)数据提取的方法应通过验证，以保障提取到的数据符合研究方案的要求。数据提取应确保提取到的原始数据与源数据的准确性，应对提取到的原始数据与源数据进行时间戳管理。

2)使用与源数据系统可互操作或集成的数据提取工具可以减少数据转录中的错误，从而提高数据准确性以及临床研究中数据采集的质量和效率。对于盲法研究，还应评估使用可互操作或集成的数据提取工具带来的揭盲风险。

中药真实世界研究数据的提取范围包括但不限于以下几个方面：

1）患者基本信息

人口学信息：包括年龄、性别、种族、籍贯、职业、教育程度、婚姻状况等，这些因素可能影响疾病的发生、发展及治疗效果。

生活方式：如饮食偏好（素食、肉食、辛辣、油腻等）、吸烟史、饮酒史、运动习惯、睡眠质量等，生活方式与中医的体质及疾病发生密切相关。

体质信息：通过中医体质问卷或医生辨证等方式，判断患者的体质类型，如平和质、气虚质、阳虚质、阴虚质、痰湿质、湿热质、血瘀质、气郁质、特禀质等。

2）中医四诊信息

望诊：观察患者的神色形态、面色、舌象（舌质、舌苔、舌形、舌态等）、脉象、分泌物及排泄物的颜色和性状等。

闻诊：收集患者的声音（言语、呼吸、咳嗽、呃逆等）、气味（口气、体气、排泄物气味等）信息。

问诊：详细询问患者的症状表现，如寒热、汗出、疼痛（部位、性质、程度、发作规律

等）、饮食、二便、情志等情况，以及疾病的发生、发展过程，既往治疗情况等。

切诊：主要是脉象（如浮、沉、迟、数、滑、涩等）和触诊（触摸肌肤、脘腹、四肢等部位，了解有无肿块、压痛、冷热等）信息。

3）疾病相关信息

疾病诊断：包括中医病名、证型诊断，以及西医疾病诊断名称、疾病分期、疾病严重程度评估等。

症状与体征：除了四诊中涉及的症状外，还需详细记录与疾病相关的其他症状和体征，如水肿的程度、关节活动受限情况、视力听力变化等。

实验室检查与影像学检查：采集血常规、生化指标、免疫指标、病理检查结果等实验室数据，以及 X 线、CT、MRI、超声等影像学检查资料，以客观反映疾病的状态和变化。

治疗信息

中医治疗：记录使用的中药方剂（药物组成、剂量、用法用量、疗程等）、针灸治疗（穴位选择、针刺手法、留针时间等）、推拿按摩（手法、部位、频率等）、拔罐、艾灸等各种中医治疗方法的具体操作和应用情况。

西医治疗：如果患者同时接受西医治疗，需记录使用的西药名称、剂量、用法、用药时间、手术治疗情况等。

治疗依从性：了解患者对治疗方案的依从程度，如是否按时服药、是否按照医嘱进行针灸或其他治疗，以及有无自行调整治疗方案等情况。

5）随访信息

症状变化：了解患者在治疗后的症状改善或加重情况，以及新出现的症状等。

疾病转归：观察疾病是否治愈、好转、稳定或恶化，有无复发、并发症发生等情况。

生存质量：通过相关量表评估患者在生理、心理、社会功能等方面的生存质量变化。

终点事件：对于一些慢性疾病或严重疾病，记录患者的终点事件，如死亡、心血管事件、肿瘤复发转移等。

6）其他相关信息

地域与环境：患居住的地理位置、气候条件、生活环境（如城市、农村、山区等）等，这些因素可能对疾病产生影响。

社会心理因素：包括患者的社会支持系统、经济状况、工作压力、心理状态（如焦虑、抑郁、烦躁等情绪）等，社会心理因素与中医的情志致病及疾病的发生发展关系密切。

5.3 数据清洗

对提取的原始数据进行重复或冗余数据的去除、变量值逻辑核查和异常值的处理，以及数据缺失的处理。需要注意，在修正数据时如果无法追溯到主要研究者或源数据负责方签字

确认，数据不应做修改，以保证数据的真实性。

在保证数据完整性的前提下去除重复数据及不相关数据。在不同数据源合并过程中，可能产生重复数据，需要去除。同时由于数据源与通用数据模型映射关系的不准确，可能会采集到与研究目标不相关的数据，从数据集中删除不需要的观测值可以减少不必要的工作。

逻辑核查和异常数据处理。通过逻辑核查可以发现原始数据或者提取数据时产生的错误，例如:出院时间早于入院时间，出生年月按年龄推算不符，实验室检查结果不符合实际，定性判断结果与方案中定义的判断标准不一致等。对异常数据的处理要非常谨慎，避免由此产生的偏倚。对于发现的错误和异常数据应通过进一步核实才能更改数据，数据的更改应保留记录。

对数据缺失进行处理，对于不同研究，数据的缺失程度、缺失原因和变量值的缺失机制不尽相同。如果涉及缺失数据的填补问题，应根据缺失机制的合理假设采用正确的填补方法。

5.4 数据转化

将经过数据清洗后原始数据的数据格式标准、中医医学术语、编码标准、衍生变量计算，按照分析数据库中对应标准进行统一转化为适用中药真实世界证据的过程，包括中药处方标准化、病历标准化、中药编码标准化、中医临床诊断标准化、非药物疗法标准化等。

对于自由文本数据的转化可使用可靠的自然语言处理算法，在保障数据转化准确、可溯源的前提下，提高转化效率。

在进行衍生变量计算时，应明确用于计算的原始数据变量及变量值、计算方法及衍生变量的定义，并进行时间戳管理，以保障数据的准确性和可追溯性。

5.5 数据传输和存储

真实世界数据的传输和存储应当基于可信的网络安全环境，在数据收集、处理、分析至销毁的全生命周期予以控制。在数据传输和存储过程中都应有加密保护。此外，应建立操作设置审批流程、角色权限控制和最小授权的访问控制策略，鼓励建立自动化审计系统，监测记录数据的处理和访问活动。

5.6 数据质量控制

数据质量控制是确保研究数据完整性、准确性和透明性的关键。数据质量控制需要建立完善的中药数据治理质量管理体系和 SOP，建议原则包括:

1）确保源数据的准确性和真实性

对于电子病历作为关键数据源，应有病历质控标准以满足分析要求，来源于门诊的疾病

描述、诊断及其用药信息需要有相关证据链佐证。

2）在数据提取时充分考虑数据完整性问题

评估和确立提取字段，制定相应的核查规则和数据库架构。

3）建立数据录入和结构化的标准指南，确保录入数据与源数据的一致性。

对于录入过程中的任何修改，需要有负责人的确认和签名，并提供修改原因，确保留下完整的稽查轨迹。

4）制定完善的数据质量管理计划

制定系统质控和人工质控计划，确保数据的准确性和完整性。对于关键变量，应进行100%核查和源文件调阅；其它变量可根据实际情况抽样核查，例如，对于人口学信息、数值型变量阈值、编码映射关系等，可按一定比例抽样，核查其准确性与合理性。

5.7 通用数据模型

基于多学科合作模式下对多源异构数据进行快速集中和标准化处理，可将不同标准的源数据转换为统一的结构、格式和术语，以便跨数据库/数据集进行数据整合，建立一套适应中药真实世界研究的数据系统。

由于多源数据的结构和类型的复杂性、样本规模和标准的差异性，在将源数据转换为通用数据模型的整体过程中，需要对源数据进行提取、转换、加载，应确保源数据在语法和语义上与目标分析数据库的结构和术语一致。

理想的通用数据模型应遵循以下原则：

1）需要定义一种数据治理机制，通过该机制可以将源数据标准化为通用结构、格式和术语，从而允许跨多个数据库/数据集进行数据整合。

2）应具有访问源数据的能力，是可动态扩展和持续改进的数据模型，并有版本控制；

3）变量的定义、测量、合并、记录及其相应的验证应保持透明，多个数据库的数据转换应有清晰一致的规则；

4）应有基线概念，安全性和有效性相关的常用变量或概念都应映射到通用数据模型，以适用于不同临床研究问题，并可通过公认或已知的研究结果进行比对。

5.8 数据治理计划书

中药真实世界研究治理计划书应事先制定，与整个项目研究计划同步。计划书中应说明使用中药真实世界数据用于监管决策的目的、使用数据的研究设计，还应对源数据进行说明，包括但不限于：中药源数据/源文件的类型，例如卫生信息系统数据、疾病登记数据、医保数据等；中药的源数据/源文件，适当评价其既往应用情况，说明采用的理由；中药真实世界证据的治理，即由中药数据来源到分析数据库的治理过程，包括数据提取、数据安全性处

理、数据清洗（逻辑核查及异常数据处理、数据完整性处理）、数据导入和结构化（通用数据模型、归一化、自然语言处理、医学编码、衍生变量）、数据传输等；采用的数据模型和数据标准；缺失数据的处理方法；减少或控制使用真实世界数据带来的潜在偏倚所采取的措施；

6 治理过程的质量管理

6.1 数据质量管理流程

数据质量控制的核心是建立数据治理体系，实现数据从产生、流转、加工和服务使用的全生命周期管理，通过数据标准的执行贯彻，完善数据质量管理，建立数据质量监控处理工作机制和技术平台，推进数据资产化管理与数据服务。确保数据的有效性、可访问性、高质量、一致性、可审计和安全性。数据质量评估有六要素：完整性、规范性、一致性、准确性、唯一性、关联性。

6.2 数据质量管理的目标

数据治理过程中的质量管理应达到以下目标:

1）建立数据质量管理组织，承担最终质量审查监督职能，质量问题有人负责;

2）建立完善的数据质量管理规范，质量管理有章可循；

3）建立质量管理平台，进行质量评估、问题改进系统记录，质量评估有据可查:

4）建立质量管理绩效体系，针对质量管理进行目标管理、绩效考核，质量考核有人监督。

6.3 影响数据质量的相关环节分析

为了进行有效的数据质量管理，数据质量管理的不同功能点应分布在流程的恰当环节，基于智能数据治理与服务平台构建数据质量管理系统，并将源系统、相关应用以及相关技术和用户纳入到数据质量中，实现发现、修正、跟踪、评估的全流程闭环管理流程，是实施数据质量管理的最佳选择。

6.3.1 数据产生

1）数据采集环节:医护人员数据录入是数据产生的源头，且难以通过数据清洗得到改善。从数据产生源头，通过增加自动化采集、改善流程设计、增强人员培训、落实质量控制制度，保证数据质量提升是数据质量管理的终极目标。

2）数据管理环节:数据字典管理是业务系统中数据标化的支点，数据字典的使用范、字典更新、版本管理、不同字典间的对照关系维护，任何一个环节不完普都会影啊到数据可用性。

3）数据交换环节:业务系统之间广泛存在数据接口用于共享和交换数据。但是接口设计中的不完善导致系统间传输数据时完整性、一致性校验不足，低质量数据可能通过医疗机构各系统间的接口调用与数据共享影响到整体数据质量。

(4)数据迁移转化环节:新旧业务系统更新时，如果旧系统数据转化、数据迁移工作没有良好的设计和实施，旧系统的数据难以与新系统数据完成融合，必然导致历史数据的损失。

6.3.2 数据集成

1）数据抽取加工环节:构建智能数据治理与服务平台，进行统一数据管理，需要从原始数据系统中抽取数据。该阶段有可能因为抽取操作引入新的数据质量问题，例如:因抽取归档病历时误抽取运行病历导致数据产生错误，或数据字段映射产生错误而导致的数据质量问题。

2）数据汇聚转换环节:在数据汇聚转换过程中，需要将抽取的原始数据根据字典码表进行标准格式转换，转换过程中应避免因转换程序问题导致的错误。

6.3.3 数据使用

在数据应用环节，同一数据可能来自不同的应用场景，导致数据名称不一致，引发数据质量问题。如医保数据分析中，“医保数据”可能来源于“城市医保”

定义好数据质量评估指标，并执行数据质量评估过程后，能够针对数据质量进行详细的检查与分析，并生成质量控制评分报告、问题列表帮助医疗机构进行数据质量改进。质量控制评分报告帮助医疗机构找到关键问题、改进方向，问题列表帮助医疗机构明确问题详情。

针对数据加工汇聚后的结果数据进行质量评价建立质量控制评分报告，帮助数据治理组织的成员从总体角度评价数据质量。数据质量分析首先应包括明确的数据质量评估维度。评估维度是指应从哪些角度、颗粒度以衡量数据质量情况。在面向全院的数据质量管理过程中，可以按照业务维度、严重程度、细致间题类型、时间维度进行查看:

1）评估数据质量的影响情况，应以业务类型作为维度查看数据质量。可将各个数据项目作为汇总的维度展示数据质量评估结果。

2）评估数据质量的严重程度，应以问题严重类型作为维度查看数据质量。根据数据权重和数据违规率主要加和之后得到数据整体的质量等级量表，量表分为不同等级。

7 经治理的中药真实世界数据适用性评价

7.1 相关性评价

旨在评估电子源数据是否与所关注的临床问题密切相关，重点关注关键变量的覆盖度、临床结局定义的准确性、目标人群的代表性和多源异构数据的融合性。

1）关键变量和信息的覆盖度

电子源数据应包含与临床结局相关的重要变量和信息，如药物使用、患者人口学和临床特征、协变量、结局变量、随访时间、潜在安全性信息等。若上述变量存在部分缺失，需充分评估是否能够使用可靠的估计方法进行填补，以及对于因果推断结果可能造成的影响。

2）临床结局定义的准确性

选择具有临床意义的结局并进行准确定义对于临床研究至关重要。临床结局的定义应包括所基于的诊断标准、测量方法及其质量控制（如果有）、测量工具（如量表的使用）、计算方法、测量时点、变量类型、变量类型的转换（如从定量转换为定性）、终点事件评价机制（如终点事件委员会的运行机制）等。当不同数据源对临床结局的定义不一致时，应定义统一的临床结局，并采用可靠的转换方法。

3）目标人群的代表性

在制定纳入和排除标准时，应尽可能地符合真实世界环境下目标人群。

4）多源异构数据的融合性

由于电子源数据的特性，很多情况下属于多来源的异构数据，需要将不同来源数据在个体水平进行数据的链接、融合和同构处理。因此，应通过身份标识符进行个体水平的准确链接，以支持通用数据模型或数据标准对数据源中关键变量进行整合。

7.2 可靠性评价

1) 完整性

是指数据信息的缺失程度，包括变量的缺失和变量值的缺失。对于不同研究，数据的缺失程度、缺失原因和变量值的缺失机制不尽相同，应该予以详尽描述。虽然电子源数据无法避免数据缺失问题，但缺失比例应有一定限度。当特定研究的数据缺失比例明显超过同类研究的比例时，会加大研究结论的不确定性，此时需要慎重考虑该数据能否作为支持产生 RWE 的数据。对缺失原因的详细分析有助于对数据可靠性的综合判断。如果涉及缺失数据的填补问题，应根据缺失机制的合理假设采用正确的填补方法。

2）准确性

是指数据与其描述的客观特征是否一致，包括源数据是否准确、数据值域是否在合理范围、结局变量随时间变化趋势是否合理、编码映射关系是否对应且唯一等。数据的准确性需

要依据较权威的参照进行识别和验证，例如，终点事件是否经独立的终点事件委员会做出判断。

3） 透明性

是指电子源数据的治理方案和治理过程清晰透明，应确保关键暴露变量、协变量和结局变量能够追溯至源数据，并反映数据的提取、清洗、转换和标准化过程。无论采用人工数据处理还是自动化程序处理，数据治理标准化操作程序和验证确认文件要清晰记录和存档，尤其反映数据可信性的问题，如数据缺失度、变量阈值范围、衍生变量计算方法和映射关系等。数据治理方案应事先根据研究目的制定，应确保数据治理过程与治理方案保持一致。数据的透明性还包括数据的可及性、数据库之间的信息共享和对患者隐私的保护方法的透明。

4） 质量控制

是指用以确证数据治理的各个环节符合质量要求而实施的技术和活动。质量控制评价包括但不限于：数据提取、安全处理、清洗、结构化，以及后续的存储、传输、分析和递交等环节是否均有质量控制，以保证所有数据是可靠的，数据处理过程是正确的；是否遵循完整、规范、可靠的数据治理方案和计划，并依托于相应的数据质量核查和系统验证规程，以保障数据治理系统在正常和稳态下运行，确保临床数据的准确性和可靠性。

5） 质量保证

是指预防、探测和纠正研究过程中出现的数据错误或问题的系统性措施。电子源数据的质量保证与监管合规性密切相关，应贯穿于数据治理的每一个环节，考虑的内容包括但不限于：是否建立与临床研究数据有关的研究计划、方案和统计分析计划；是否有相应的标准操作规程；数据收集是否有明确流程和合格人员；是否遵守收集关键数据变量的共同时间框架；用于数据元素捕获的技术方法是否充分，包括各种来源数据的集成、药物使用和实验室检查数据的记录、随访记录等；患者的选择是否将偏倚最小化以体现真正的目标人群；数据输入是否及时、传输是否安全；是否满足监管机构现场核查调阅源数据、源文件等相关要求。

参考文献

[1] 国家药监局药审中心关于发布《基于疾病登记的真实世界数据应用指导原则（试行）》的通告
[EB/OL](2024-11-12).https://www.cde.org.cn/main/news/viewInfoCommon/a8d3a782d45feebfcc700e79f3d86326

[2] 国家中药管理局 国家数据局.关于促进数字中药发展的若干意见的通知
[EB/OL](2024-07-19).https://www.gov.cn/zhengce/zhengceku/202408/content_6968519.htm

[3] 国务院办公厅.印发“十四五”中药发展规划的通知
[EB/OL](2022-03-03).https://www.gov.cn/zhengce/zhengceku/2022-03/29/content_5682255.htm

[4] 国家药品监督管理局.真实世界证据支持药物研发与审评的指导原则(试行)[EB/OL].(2020-01-03).https://www,nmpa.gov.cn/xxgk/ggtg/qtggtg/20200107151901190.html.
[5] FDA.Use of real-world evidence to support regulatory decision-making foimedical devices[EB/OL].(2017-08-31).https://www.fda.gov/media/99447/download.
[6] FDA.Framework for FDA's realworld evidence progarmEB/OL].(2018-12) https://www.fda.gov/media/120060/download.
[7] 赵国桢,闫世艳,郭玉红,等.基于既有医疗数据构建研究型数据库的方法学探讨及实例解读(二):数据治理的方法[J].中国中药信息杂志, 2023, 30(9):17-21.
[8] 石晶金,于广军.跨区域专科联盟健康医疗数据共享的治理框架与对策研究[J].中国医院, 2023, 27(5):43-46.DOI:10.19660/j.issn.1671-0592.2023.05.12.
[9] 陆家发,赵群,唐凯,等.江苏省全民健康信息平台数据治理实践研究[J].中国卫生质量管理, 2023, 30(8):63-67.
[10] 李丽娟,吴强,贾炜,等.基于数据治理的医院绩效考核 PDCA 管理实践[J].中国卫生质量管理, 2023, 30(12):74-77.
[11] Clarke, David., Cocozza, Anna., Appleford, Gabrielle., Thabet, Aya., Ross, Susan.. Data-driven governance and the private sector in mixed health systems. BMJ global health, 2024,.
[12] Shabani, Mahsa.. The Data Governance Act and the EU's move towards facilitating data sharing. Molecular systems biology, 2021, 17(3):e10229.
[13] Brand, Dirk., Singh, Jerome A., Nienaber McKay, Annelize G., Nienaber McKay, Annelize G., Cengiz, Nezerith.. Data sharing governance in sub-Saharan Africa during public health emergencies: Gaps and guidance. South African journal of science, 2022 .