

ICS号(请添加)

中国标准文献分类号(请添加)

团 体 标 准

T/CGCPU 00X-2021

多维数据融合的心血管疾病药物临床研究 数据体系

Data System of cardiovascular drug clinical research based on
multi-dimensional data fusion

(征求意见稿)

在提交反馈意见时, 请将您知道的相关专利连同支持性文件一并附上

2021-xx-xx 发布

2021-xx-xx 实施

中关村玖泰药物临床试验技术创新联盟 发布

前 言

本部分按照 GB/T 1.1-2020 给出的规则起草。

请注意本文件的某些内容可能涉及专利，本文件的发布机构不承担识别这些专利的责任。

本标准由中关村玖泰药物临床试验技术创新联盟归口。

本标准起草单位：北京光大隆泰科技有限责任公司、中关村玖泰药物临床试验技术创新联盟、首都医科大学附属北京安贞医院、宝创瑞海（北京）科技发展有限公司。

本标准主要起草人：×××，×××，×××等。

本标准为首次发布。

引 言

心血管疾病是全球范围内的重大公共卫生问题,加快我国具有自主知识产权的心血管疾病药物研发速率,提高研发质量刻不容缓。大数据时代下,大数据分析驱动新药创制方法和技术策略,利用数据模型评价体系及大数据云计算对创新药物研发各个阶段积累的海量数据进行深入挖掘、整合,全方位揭示药物特性,实现药物临床研究的高效、动态、准确的评价。

多维数据融合的心血管疾病药物临床研究数据体系

1 范围

本标准给出了基础性的心血管疾病药物临床研究数据及数据来源，经系统性的数据收集形成数据库，融合多维数据，应用大数据统计分析方法，完善心血管疾病药物临床研究。

本标准仅针对心血管疾病药物治疗的临床研究数据，不包含医疗器械治疗的临床研究数据。

2 规范性引用文件

下列文件对于本文件的引用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 22239 信息安全技术 网络安全等级保护基本要求

《药物临床试验质量管理规范》（2020年第57号） 国家药品监督管理局 国家卫生健康委员会

3 术语和定义

《药物临床试验质量管理规范》界定的以及下列术语和定义适用于本文件。

3.1

心血管疾病 Cardiovascular disease

循环系统包括心脏、血管和血液循环的神经体液调节装置。循环系统疾病包括心脏和血管病，合称心血管疾病。

3.2

维度 Dimensions

具有某类性质的数据集合。

3.3

多维数据分析 Multidimensional data analysis

多重维度的数据组织起来，经过各种分析操作，使分析者、决策者能从多个角度、多个侧面观察数据，深入了解包含在数据中的信息和内涵。

4 总则

心血管疾病的分类有其特殊性，应包括病因、病理解剖和病理生理的分类。这些疾病因不同的病因、不同的病理解剖以及不同的病理生理过程，导致心血管疾病种类繁多，因此心血管疾病临床研究的数据复杂。本标准给出部分数据维度供使用者参考，依据临床研究数据的种类进行数据收集，数据分析时，应根据需求开展多维数据分析，并不断的反哺到心血管疾病药物临床研究中，不断完善包括方案设计、

试验实施、质量管理在内的临床研究/试验全过程评价体系，为心血管疾病临床研究提供数据理论。

5 多维数据

心血管疾病药物临床研究数据的数据维度，可按照心血管疾病类型、治疗方法、治疗途径、治疗环境、研究类型等等进行划分。本标准给出部分维度供使用者参考。心血管疾病药物临床研究的划分因素繁多，数据维度广泛，要尽可能的收集不同维度数据，借助大数据分析手段，开展数据分析。

心血管疾病的形成因素复杂，根据致病因素分为先天性和后天性两大类：

1. 先天性心血管病（先心病）

2. 后天性心血管病：有以下类型。①动脉粥样硬化；②风湿性心脏病；③原发性高血压；④肺源性心脏病；⑤感染性心脏病；⑥内分泌性病性心脏病；⑦血液病性心脏病；⑧营养代谢型心脏病；⑨心脏神经症；⑩其他：如药物或化学制剂中毒、心脏肿瘤和不明原因的心肌病等等。

曾有研究者基于国际疾病分类的心血管疾病亚型的基因组学研究，将心血管疾病分为：（1）缺血性心脏病，（2）肺源性心脏病和肺循环疾病，（3）脑血管疾病，（4）动脉、小动脉和毛细血管疾病，（5）静脉、淋巴管和淋巴结疾病，不可归类在他处者^[1]。

心血管疾病治疗药物的作用机制、给药途径、给药方法、给药时间、适应症等均有不同。按照适应症可分为抗高血压药物、抗心绞痛药物、扩血管药、抗动脉粥样硬化药物、强心苷药物、降血脂药物、抗心律失常药物等等。近年来使用干细胞治疗心血管疾病的研究也有突破进展。

临床研究的类型按照临床试验分期可分为 I-IV 期临床试验、生物等效性试验、药代动力学试验、药效动力学试验；根据方案设计有平行、交叉、队列、随机、盲法、单臂、双臂、三臂、多中心等临床研究方法。不同的临床研究中心、检测设备、检测方法、研究者等也对临床研究结果产生影响。

6 数据来源

本标准涉及的数据是心血管疾病药物的临床研究数据。临床研究数据一般来源于合规性文件、临床试验方案、病例报告表（CRF）、知情同意书（ICF）、原始记录表、过程性记录表、统计分析报告、总结报告等临床研究文件。可将其分为：诊疗数据、受试者管理数据、试验用药品数据、生物样本数据、安全性数据、统计分析数据，及辅助数据。

诊疗数据：住院信息、门诊信息；生命体征，如体温、脉搏、呼吸、血压；血常规，血生化，尿常规、便常规，出凝血功能；心电图，医学影像；传染病检查；评价量表、问卷调查等。根据《药物临床试验质量管理规范》的要求，以患者为受试者的临床试验，相关的医疗记录应当载入门诊或者住院病历系统。遂诊疗数据一般记录在医院门诊病历或住院病历系统中，医院管理系统一般有 HIS、LIS、PACS 等系统。

受试者管理数据：受试者信息、筛选入选数据、随机分组数据、随访数据、方案依从性数据等。以上均可在医院门诊病历或住院病历系统中采集。随机分组数据一般来源于随机化系统。同时一些更为原始的数据记录在表单中，如筛选入选表、鉴认代码表、随机表、随访登记表、受试者日记卡等。

试验用药品数据：临床试验用药品全过程管理数据，包括临床前研究数据、试验用药品信息（名称、编码、规格、使用方法、使用量、存储条件、检验报告等），试验用药品运输、接收、验收、入库、存

储、出库、领用发放、用药、回收、返还、退回、销毁、留样（如适用）等数据。试验用药品的数据一般记录在研究者手册、试验用药品的说明书、检测报告、温湿度监控记录、接收验收入库登记表、领用返还记录、销毁记录等表单中。部分研究中心已建立了试验药品管理信息化系统，将试验药品全过程的管理数据记录在系统中。

生物样本数据：血液、组织样本信息及对样本的检测分析信息等；根据人类遗传资源管理要求，申报审批的人类遗传资源信息等。生物样本数据一般记录在检验检查单、手术记录、组织样本登记表及交接记录中。人类遗传资源信息按照人类遗传资源管理要求，开展相应的检验检查及记录，形成相应的表单。

安全性数据：试验方案中要求记录的安全性数据，如 AE、SAE、SUSAR 及处理上报信息。对于已上市药品还应收集药品不良反应监测信息。指标异常检验检查报告单，不良事件的记录、处理上报表单以及药品不良反应监测数据来源，如文献收集、医院使用药品不良反应数据收集等，都是安全性数据的来源。

统计分析数据^[2]：临床研究统计分析计划、统计分析方法及结果等。统计分析数据一般来源于病例报告表（CRF），近年来临床试验数据一般使用电子病例报告表（eCRF）记录临床试验数据。统计分析计划书、统计分析报告记录着详细的统计分析方法、过程及结论。由专业的生物学统计分析师依据事先规定的统计分析方法，利用专业统计分析软件，对临床研究数据进行统计分析并出具统计分析报告。

辅助数据：合规性数据来源于实验方案、临床试验批件/默示许可信息、伦理批件、研究合同等记录的数据；监查数据、稽查、检查数据来源于相应的监查/稽查/检查记录表及报告；检测方法及检测设备，检查设备数据来源于仪器使用记录及报告单；研究中心信息、研究者信息来源于分中心目录、研究者授权分工表、研究者资质材料、培训记录等。

7 数据收集与分析

在临床研究数据信息化管理的时代下，利用大数据收集分析系统，系统化收集相应的临床研究数据，建立多维数据融合的心血管疾病药物临床研究数据库。使用者可根据需求，选取特定统计方法与数据维度，融合多维数据，构建系统分析模型，开展多维数据分析，为心血管疾病的临床研究提供数据支持，提升心血管疾病药物临床研究设计水平，建立更加完善得临床研究的方案设计、实施与质量管理体系。

系统应支持数据收集和数据分析功能。通过系统录入收集数据，或能够对接其他数据来源系统，如 HIS、LIS、PACS、eCRF、CTMS 等，抓取多维数据，同时具有数据处理分析功能。

数据系统在基础设施与体系的支撑上，应提供分布式服务、缓存服务、文件存储服务、数据库服务、日志服务、服务治理、业务服务等。系统包含以下模块但不限于：受试者管理模块、研究者管理模块、项目管理模块、试验方案辅助设计模块、生物样本库模块、质控模块、风险预警管理模块、数据分析模块、登录模块、日志模块、账户设置模块等。

8 数据管理与信息安全

数据库系统应当具有完善的权限管理，不同角色之间的权限进行严格隔离，经过培训后才可授权使用。数据库系统应具备完善得稽查轨迹，应支持根据试验用药品编号、受试者随机号/筛选号、项目编号等不同关键词进行一键溯源。

数据库系统应满足相应的基础设施要求，如存放空间应保证环境参数的规范性及环境的安全性；配备应急设施，如双电路/不间断电源(UPS)，或者配有备用发电机组，保障 24h 供电；具备满足系统运行要求的网络环境。

系统应满足 GB/T 22239-2019 网络安全等级保护基本要求中的第三级安全要求，还应满足网络安全、主机安全、应用安全、数据安全和安全管理方面的其他要求。

参考文献

- [1] 郭子宁, 梁志生, 周仪, 等. 基于国际疾病分类的心血管疾病亚型的基因组学研究[EB/OL]. <https://kns.cnki.net/kcms/detail/11.4691.R.20210512.1349.006.html>, 2021-05-12/2021-06-10.
- [2] 药物临床试验的生物统计学指导原则[EB/OL]. 国家食品药品监督管理总局, <https://www.nmpa.gov.cn/yaopin/ypggtg/ypqtgg/20160603161201857.html>, 2016-06-03/2021-06-10.