

《专病数据库建设规范》(送审稿) 编制说明

一、项目背景

随着医学科技的快速发展，精准医疗、个性化治疗和大规模临床研究对高质量、结构化的数据需求日益增长，专病数据库作为以疾病为核心、患者为对象的全周期数据管理平台，在打破科室数据孤岛、实现跨系统数据整合与共享、支持临床决策、医疗质量控制和科研分析等方面发挥着重要作用。然而，目前专病数据库的建设在设计、数据采集与整合、数据存储与管理、数据利用、安全管理、质量控制等方面缺乏统一的标准和规范，导致不同医疗机构在数据质量、数据共享、隐私保护方面存在较大的差异。

为了推动专病数据库建设的规范化、标准化，提升临床研究、患者管理和科技成果转化的效率，深圳市卫生健康委员会组织相关医疗机构和科技企业，共同制定《专病数据库建设规范》。本标准旨在通过规范化、系统化的方式，为医疗机构及科研单位提供专病数据库的建设与管理规范，确保不同医疗机构在数据质量、数据共享、隐私保护方面具备一致的技术基础。

二、工作简况

(一) 任务来源

根据深圳市市场监督管理局 2024 年 4 月下达的《深圳市市场监督管理局关于下达 2024 年第二批深圳市地方标准

计划项目任务的通知》，《专病数据库建设规范》正式获得批准立项，本项目由深圳市卫生健康委员会提出并归口，由深圳市妇幼保健院牵头起草。

（二）主要起草过程

2024 年 2 月，深圳市卫生健康委员会提出专病数据库建设规范编制需求。

2024 年 3 月，就《专病数据库建设规范》提交深圳市地方标准制修订计划项目。

2024 年 7 月，《专病数据库建设规范》获深圳市市场监督管理局批准立项。

2024 年 8 月，成立标准起草工作组。

2024 年 9 月，召开工作组启动会议，开展调研、资料收集和分析工作。

2024 年 11 月，完成标准征求意见初稿的编写，并召开第二次工作组会议，对初稿进行讨论和修改。

2024 年 12 月，完成标准征求意见稿第二版修改。

2025 年 1 月，组织专家对标准征求意见稿第二版进行评审，并召开第三次工作组会议，根据专家意见进行修改完善，形成征求意见稿终版。

2025 年 2 月，提交深圳市卫生健康委员会审核，并进行广泛意见征求。

三、地方标准主要内容的依据以及与国内领先、国际先进标准的对标情况

（一）标准主要内容

本标准涵盖专病数据库的设计、数据采集与整合、数据存储与管理、数据利用、安全管理、质量控制、专病数据库评价等关键环节，主要包括：

1、范围：适用于医疗机构、科研单位及相关服务提供方在构建和管理专病数据库时的技术要求与管理规范。

2、规范性引用文件：引用了一系列国家标准和行业标准，确保标准的科学性和权威性。

3、术语和定义：明确了专病库、数据治理、区块链等相关术语和定义。

4、缩略语：明确了文件中的缩略语的含义及英文全称。

5、建设原则和目标：提出了专病数据库建设应遵循的原则和目标。

6、建设内容和技术要求：规范了数据范围、功能模块、设计要求、数据安全和隐私以及专病库的评价标准。

7、建设流程和管理：建设的具体流程和管理要求。

8、专病数据库评价：确定了评价专病数据库的两个方面五个指标。

（二）技术依据

为确保标准的科学性、权威性和实用性，我们充分参考和引用了国内外相关的技术标准和规范。本标准的技术依据主要围绕专病数据库的设计、数据采集与整合、数据存储与管理、数据利用、安全管理、质量控制等关键环节展开，涵盖了一系列国家标准和行业标准。以下是对本标准技术依据的详细阐述：

1、信息安全技术相关标准

GB/T 22239—2019《信息安全技术 网络安全等级保护基本要求》：该标准明确了网络安全等级保护的基本要求和实施指南，为本标准中数据安全与隐私保护部分提供了重要参考。

GB/T 25069—2022《信息安全技术 术语》：该标准对信息安全技术的术语进行了明确定义，并分门别类地编制了术语表，为本标准中统一术语标准提供依据，提高各方对术语的理解和沟通。

GB/T 35273—2020《信息安全技术 个人信息安全规范》：该标准规定了个人信息的收集、存储、使用、共享、转让、公开披露等安全要求，为本标准中患者个人信息保护提供了依据。

GB/T 39725—2020《信息安全技术 健康医疗数据安全指南》：该标准针对健康医疗数据的安全管理提出了具体要求，包括数据的采集、存储、使用、共享、传输等环节，为本标准中专病数据库的安全管理提供了重要指导。

GB/T 39786—2021《信息安全技术 信息系统密码应用基本要求》：该标准规定了信息系统密码应用的基本要求，包括密码算法的选择、密钥管理、密码技术应用等方面，为本标准中数据加密技术的应用提供了依据。

GB/T 37964—2019《信息安全技术 个人信息去标识化指南》：该标准描述了个人信息去标识化的目标和原则，提出了去标识化过程和管理措施，为本标准中的数据提供具体

的个人信息去标识化指导，以及组织开展个人信息安全监督管理、评估等工作的依据。

2、健康医疗数据相关标准

WS/T 305—2023《卫生健康信息数据集元数据标准》：该标准规定了卫生健康信息数据集的元数据描述和管理要求，为本标准中数据集的设计和元数据管理提供了依据。

WS/T 500—2016《电子病历共享文档规范》：该标准规定了电子病历共享文档的结构和内容要求，为本标准中数据采集与整合部分提供了参考。

WS/T 790—2021《区域卫生信息平台交互标准》：该标准规定了区域卫生信息平台的数据交换和共享要求，为本标准中数据共享与协作部分提供了指导。

3、数据标准化与互操作性相关标准

GB/T 42384—2023《健康信息学 数据交换标准 HL7 临床文档架构（版本 2）》：该标准规定了临床文档的数据交换格式和要求，为本标准中数据标准化和互操作性提供了依据。

CDISC（Clinical Data Interchange Standards Consortium）相关标准：CDISC 是一个致力于开发临床数据交换标准的国际组织，其发布的一系列标准（如 SDTM、ADaM 等）为本标准中临床数据的标准化和互操作性提供了重要参考。

4、其他相关标准

ICD-10 (International Classification of Diseases, Tenth Revision)：该标准是国际疾病分类的第十版，为本标准中疾病诊断编码提供了依据。

SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms)：该标准是一个系统化的医学术语集，为本标准中医学术语的标准化提供了参考。

其他国家和行业标准：在制定本标准的过程中，还参考了其他国家和行业标准，如数据治理、数据质量控制、多模态数据整合等方面的标准，以确保标准的全面性和适用性。

(三) 标准创新点与差异化分析

《专病数据库建设规范》在充分吸收国内外医疗信息化及健康数据标准（如 GB/T 22239、GB/T 35273、CDISC、WS/T 305 等）的基础上，立足于疾病研究和临床实践的实际需求，提出了多项具有前瞻性、系统性和实操性的创新内容。与现有医疗数据管理、电子病历、区域卫生信息平台等标准相比，本标准在专病数据库的架构设计、数据质量控制、安全保障、临床适配性与实际应用导向等方面形成了显著差异化，主要体现在以下五个方面：

1、构建了专病数据库“模块化+分层”双重架构设计体系，实现系统灵活部署与功能扩展

本标准创新性地提出了 14 个核心功能模块与 5 层参考系统架构（应用层、服务层、数据层、安全层、基础层）相结合的设计框架，支持从数据采集、治理、质控、共享到安全合规的完整生命周期管理。这一架构较传统的电子病历系

统或区域平台更具专病研究导向性、可组合性和横向集成能力，填补了专病数据平台在标准化建设方面的空白。

2. 首次将“区块链数据存证”机制系统纳入医疗数据标准，用于保障数据完整性与可信性

本标准将区块链技术作为数据完整性保护和多中心协作场景下数据可信存证的核心机制，提出包括“区块链存证模块”“区块链服务支撑体系”等要求，实现数据“全流程上链、可验证、可审计”。这一设计在国内医疗数据标准中尚属首次，显著增强了科研数据的可追溯性与可信度，特别适用于真实世界研究、药械注册和跨机构协同科研。

3. 构建专病数据库“评价指标体系”，明确系统建设质量与应用价值的评估方法

标准提出了由“相关性、可及性、稳健性、可靠性、数据质量”五个维度组成的评价体系，并明确评价流程和指标内涵，强调数据库是否能够“满足临床研究需求”“映射真实世界场景”。该体系不仅可用于项目验收，还可作为持续改进与质量管理的工具，区别于传统仅以功能合规为主的验收判定方式，更加突出医学科研与临床价值导向。

4. 强调“术语标准化+数据字典+指标集”协同机制，打通数据结构化与语义标准的闭环

标准在术语使用方面提出“国际标准优先+术语映射”策略，配合完整的数据字典和由临床专家主导的指标集制定机制，确保专病数据库中数据的结构统一和语义一致。这种从术语—数据元—指标集全链条标准化机制，有效解决了多

源异构数据难以共享和分析的问题，提升了数据在多中心环境下的互操作性与可用性。

5. 明确提出“多模态数据整合与智能分析支持”要求，拓展数据库的深度应用能力

本标准将影像、文本、基因、病理等非结构化与结构化数据视为专病数据库的重要组成部分，提出多模态数据管理模块、知识图谱模块、NLP 支持、AI 模型训练等应用场景，并在附录 B 中进一步明确机器学习、自然语言处理、数据可视化等分析技术要求。这一设计体现了数据库“从数据仓库向智能平台”演进的趋势，显著拓展了标准的适用深度和未来发展空间。

四、各章条款的说明主要条款的说明以及主要技术指标、参数、试验验证的论述

1、第 1 章到第 4 章：第 1 章明确了文件的适用范围，涵盖医疗机构、科研单位等开展专病数据库建设与管理工作的关键环节。第 2 章列出了规范性引用文件，为后续建设规范提供依据。第 3 章详细解释了专病库、数据治理等重要术语，统一概念认知。第 4 章则对 ABAC、API 等缩略语进行定义，方便文件理解。这些章节相互配合，搭建起了专病数据库建设规范的基本框架，为后续具体规范内容的展开奠定基础。

2、第 5 章 建设原则和建设目标：阐述了专病数据库建设应遵循的原则和目标。建设原则包含规范性、系统性、安全与可信、质量、易用高效、可扩展性七项原则。建设目标

旨在通过标准化、系统化采集、存储、治理和分析特定疾病相关数据，提升临床研究、患者管理和医疗决策的效率。

3、第6章 建设内容与技术要求：包括数据范围、模块化系统架构、参考技术框架、设计要求、数据要求、数据安全与隐私保护要求、专病数据库评价：

3.1 数据范围：规定了专病数据库应涵盖的多维度数据内容。

3.2 模块化系统架构：规定了包含数据采集模块、数据存储模块、数据治理模块等14个专病数据库架构的核心功能模块，及其用途和要求。

3.3 参考技术架构：规定了专病数据库包含的应用层、服务层、数据层、安全层和基础层5个层次，几个层次内部的关键模块。可以作为专病数据库架构的参考。

3.4 设计要求：详细规定了专病数据库设计的关键要素，包括设计要求上，强调业务需求驱动、模块化架构，保障系统实用性和扩展性。模块化系统架构要求采用模块化设计，支持各功能模块独立可扩展。数据服务操作要求规范数据审核、导出、稽查等操作，保障数据安全性与准确。

3.5 数据要求：从多维度对专病数据库中的数据进行了规定。术语标准化通过采用国际标准、管理术语库和实现术语映射，保证医学术语一致性。元数据基本要求确保数据集元数据规范统一，便于数据处理。数据字典要求以术语标准化和元数据管理为基础，对每一个数据元进行标准化描述，规范字段名称、定义、数据类型、单位、值域、代码系统及

适用范围等内容，支撑专病数据库的数据采集、治理、共享与分析。专病指标集要求以数据字典为基础建立，由临床专家和数据专家共同制定并定期更新。数据结构与模型要求采用标准化结构和通用模型，具备扩展性，支持数据集成共享。数据质量控制标准化建立质控规则和流程，实现持续改进。数据共享与互操作要求通过标准化接口，保障数据在不同系统和机构间安全高效共享。

3.6 数据安全与隐私保护要求：围绕数据全生命周期保障安全隐私。数据分类分级要求依据数据的属性、敏感性及潜在影响，对所采集、存储和处理的数据资源开展分类分级管理。数据加密采用国家密码标准，结合权限管理保护数据安全。数据访问控制基于角色或属性分配权限，采用多种认证机制防止未授权访问。数据去标识化与隐私保护针对不同机构提供技术选型建议，保护患者隐私。数据传输与存储安全通过加密传输、存储，采用分布式存储备份保障数据安全。数据完整性保护利用区块链技术确保数据真实完整。系统日志与审计记录操作行为，保证所有数据操作的可追溯性和合规性。

3.7 数据合规要求：规定了专病数据库的建设和使用应严格遵守国家有关法律法规和伦理，确保数据利用的合法性与合规性，包括数据合法合规、遵循伦理要求、数据留存与销毁机制、合规审计与监督。

4、第7章 建设流程和管理：系统阐述建设流程和管理要点。建设流程分为建设准备阶段、系统部署与配置阶段、

测试与验收阶段、数据运维与生命周期管理阶段，以及数据库实施和迁移。各环节紧密相连，保障项目有序推进。

5、第 8 章 专病数据库评价：规定了专病数据库评价体系。评价标准从满足临床研究需求和映射真实世界场景两方面展开，涵盖相关性、可及性、稳健性、可靠性和数据质量五个维度。评价流程明确了从确定评价主体与对象到反馈整改的各环节。此外，还涉及与其他标准规范的衔接、动态更新机制等内容，确保评价工作科学权威，推动专病数据库持续改进。

6、附录 A 参考实施流程：提供了专病数据库实施的标准流程。该流程分为需求调研与规划、系统设计与架构搭建、数据采集与治理等六个主要步骤。每个步骤都有明确的任务和目标，如需求调研阶段确定建设目标和数据集范围，系统设计阶段明确技术方案并搭建环境。各步骤环环相扣，从规划到上线运维，为建设者提供了清晰的实施路径，有助于确保系统有效建设、部署和稳定运行。

7、附录 B 数据利用策略与方法：聚焦专病数据库的数据利用。基本原则强调合法合规、支持科研临床、高效多维分析和保障安全隐私。数据应用场景广泛，支持单中心和多中心研究、药物器械研发等。数据分析与挖掘支持多种方法，提升数据价值。数据共享与协作遵循标准规范，保障数据安全。合规与伦理审查确保数据使用合法合规，保护患者权益。

五、是否涉及专利等知识产权问题

本文件未涉及专利等知识产权问题。

六、重大意见分歧的处理依据和结果

无。

七、实施地方标准的措施建议

在专病数据库建设规范的实施过程中，为了确保标准的有效落地，建议采取以下实施标准的措施：

1、明确实施目标与责任分工 清晰定义专病数据库建设的具体目标，如提高数据质量、促进临床研究、支持临床决策等。明确各参与方的职责，包括医疗机构、科研单位、信息技术部门等，确保各方协同合作。

2、制定详细的实施计划 制定详细的时间表，明确各阶段的任务和时间节点。合理调配人力、物力和财力资源，确保项目的顺利进行。

3、加强人员培训与技术支持 对参与专病数据库建设的人员进行专业培训，包括数据采集、数据整合、数据分析等方面的技能。提供必要的技术支持和维护服务，确保系统的稳定运行和数据的安全。

4、推动数据共享与协作研究 建立数据共享机制，促进不同医疗机构和研究机构之间的数据交换和共享。鼓励跨学科的协作研究，利用专病数据库开展大规模、多中心的临床研究。

5、加强安全与隐私保护 制定完善的安全策略，包括数据加密、访问控制、备份恢复等措施。严格遵守相关法律法规和伦理规范，确保患者个人信息的安全和隐私保护。

6、定期评估与持续改进 定期对专病数据库建设项目的进展和成效进行评估。收集用户反馈，针对存在的问题进行改进和优化，不断提升系统的性能和用户体验。

7、促进标准的应用与推广 加大对专病数据库建设规范的宣传力度，提高医疗机构和科研单位对标准的认知度和接受度。

8、采标示范项目 开展标准应用示范项目，展示专病数据库建设规范的实施效果，带动更多医疗机构和科研单位参与标准的应用与推广。

9、建立反馈与沟通机制 设立反馈渠道，鼓励用户提出意见和建议，及时收集并处理用户反馈。定期与医疗机构、科研单位等相关方进行沟通，了解他们的需求和期望，不断优化和完善标准。

八、其他

暂无。