

附件 2

推荐性国家标准项目申报书

项 目 名 称 : 人工智能医疗器械 质量要求和评价 第 2 部分：数据集通用要求

技术归口单位（或技术委员会） : 国家药品监督管理局

提 出 日 期 : 2025-3-5

一、基本信息

中文名称	人工智能医疗器械 质量要求和评价 第 2 部分：数据集通用要求		
英文名称	Artificial Intelligence Medical Devices — Quality Requirements and evaluation — Part 2: general requirement for datasets		
标准性质	☒ 推荐性国家标准 ☐ 指导性技术文件		
制定/修订	☒ 制定 ☐ 修订	被修订标准号	/
是否采标	☐ 是 ☒ 否	采标类型	/
采标号	/	采标中文名称	/
项目周期	☒ 12 个月 ☐ 16 个月 ☐ 18 个月		
上报单位	国家药品监督管理局		
技术归口单位 (或技术委员会)	国家药品监督管理局		
主管部门	国家药品监督管理局		

二、论证评估报告

（一）制修订推荐性国家标准的必要性、可行性

1. 必要性

数据集是人工智能医疗器械研发、测试、临床试验、日常在用质控、上市后监督等各个环节必备的重要资源，在人工智能医疗器械生产质量管理过程中扮演着原材料或者关键元器件的角色，是产品质控的重要组成部分。

我国已步入数字经济时代，医学健康数据产业开始崛起，但不同机构提供的数据质量参差不齐，缺乏统一的评价规范，对数据集供应链带来风险，关系到人工智能医疗器械产业的安全。为确保人工智能医疗器械健康发展，数据集必须接受严格的监管。

人工智能医疗器械使用的数据集质量涉及交叉学科，相关医疗器械标准尚处于空白。信息技术标准（如 GB/T 36073—2018、GB/T 36344—2018）提出了数据管理能力成熟度评估、数据质量评价的理念，但缺乏对医学场景和人工智能医疗器械特殊风险的考虑，难以提供对人工智能医疗器械使用的数据集质量特性的系统描述。卫生信息标准（如 WS/T 305—2009、WS 539—2017）提出了元数据规范、元数据属性要求等概念，同样缺乏数据集质量特性和质量评价内容，与人工智能医疗器械质量未建立衔接。

在此背景下，人工智能医疗器械数据集质量评价作为支撑

监管的必要举措，急需推动国家标准制修订，在源头上加强数据质量评价公共服务体系，促进数据集供应链健康发展，更好地满足人工智能医疗器械产业和监管需求。

2. 可行性论证

本标准的当前技术内容来自国家药监局 2022 年发布的 YY/T 1833.2-2022 《人工智能医疗器械 质量要求和评价 第 2 部分：数据集通用要求》。该标准首次提出了数据集说明文档要求、数据集质量特性、评价方法三大要素，形成了数据集质量评价的通用框架。在 YY/T 1833.2 的基础上，国家药监局在 2024 年发布了 YY/T 1949-2024 《人工智能医疗器械 数据集专用要求：糖尿病视网膜病变眼底彩照》，形成了通用标准带动专用标准的局面。值得注意的是，该专用标准的牵头单位由国内医院及国家卫健委相关单位，说明国家卫生健康部门认同 YY/T 1833.2 的思路。

YY/T 1833.2-2022 在 2023 年 7 月正式实施，在 2024 年开展了实施评价，相关各方的反馈均为满意。目前，国内医疗器械检验机构已纷纷具备检测能力和检测经验，说明本标准具有应用落地的基础。

3. 拟解决的问题

本标准用于对人工智能医疗器械数据集进行质量评价，包括训练、调优、测试、临床试验、日常质控、上市后监督等各个关键环节使用的数据集，旨在为人工智能医疗器械监

管提供客观证据，加强数据集制造责任方的合规性，确保数据集供应链的高质量发展。本标准的具体评价对象包括数据集说明文档、数据集整体及数据单元的质量、数据集开发过程记录、风险管理文档等，基本覆盖了数据集全生命周期。

（二）主要技术要求

本标准是人工智能医疗器械数据集质量评价的基础标准，在制定过程中，综合参考了医疗器械、信息技术、卫生数据管理、临床数据管理等相关领域的的数据质量要求与评价方式，根据人工智能医疗器械技术特点编制而成。

本标准的范围是：本文件规定了人工智能医疗器械全生命周期使用的数据集的通用质量要求和评价方法本文件适用于人工智能医疗器械研发、生产、测试、质控等环节使用的数据集的开发与评价。

本标准第四章提出了数据集描述文档的通用要求，包括数据采集、数据预处理、数据标注等环节。本标准第五章提出了数据集质量特性的含义及表征，用于指导数据集制造责任方在数据集描述文档中声称数据集质量特性。本标准第六章提出了数据集质量评价的方法和流程。

本标准的内容参考了 GB/T 42061《医疗器械 质量管理体系 用于法规的要求》、GB/T 42062《医疗器械风险管理对医疗器械的应用》、原国家食品药品监督管理总局《关于发布临床试验数据管理工作技术指南的通告（2016 年第 112

号)》、《深度学习辅助决策医疗器械软件审评要点》、GB/T 36344-2018《信息技术 数据质量评价指标》、GB/T 34960.5: 2018《信息技术服务 治理 第5部分: 数据治理规范》、GB/T 35295: 2017《信息技术 大数据 术语》、GB/T 25000.12: 2017《系统与软件工程 系统与软件质量要求和评价(SQuaRE) 第12部分: 数据质量模型》等。GB/T 11457-2006《信息技术 软件工程术语》、WS/T 305-2009《卫生信息数据集元数据规范》, 国际标准 ISO 19157-2013《Geographic Information-Data quality》、ISO 3951-1:2013《Sampling procedures for inspection by variables — Part 1: Specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection for a single quality characteristic and a single AQL》等标准化文件、监管技术文件。

其中, 数据集的质量评价方法部分主要参考和对照了经济合作与发展组织(OECD, Organization for Economic Cooperation and Development)发布的推荐性指南《OECD Recommendation of the Council on Health Data Governance, OECD/LEGAL/0433》、欧盟发布的数据质量评价手册《Handbook on Data Quality Assessment Methods and Tools》、世界卫生组织数据质量审核工具《Data Quality Review (DQR) Toolkit》等文件内容与实施细则。抽样检验

方法主要参考文献包括 GB/T 2828.1-2012《计数抽样检验程序 第1部分：按接收质量限（AQL）检索的逐批检验抽样计划》、GB T 18391.1-2009《信息技术 元数据注册系统（MDR） 第1部分：框架》、GB/T 8054-2008《计量标准型一次抽样检验程序及表》、GB/T 10111-2008《随机数的产生及其在产品质量抽样检验中的应用程序》等。

本标准制定的过程中，未涉及有争议指标。

（三）国内外标准情况、与国际标准一致性程度情况

目前暂无针对人工智能医疗器械数据集质量评价的相关国际标准、国外标准。本项目不涉及国际标准转化，不涉及版权问题。

本标准的技术内容来自国内现行有效的 YY/T 1833.2—2022《人工智能医疗器械 质量要求和评价 第2部分：数据集通用要求》，将对其技术内容进行修订，与最新的监管要求和行业现状相适应。

（四）与相关强制性标准、法律法规配套情况

与有关法律、法规无冲突，协调一致。

（五）标准所涉及的产品、过程或者服务目录

本标准涉及人工智能医疗器械，例如采用人工智能技术的肺结节辅助检测软件、糖尿病视网膜病变辅助诊断软件、冠状动脉 CT 影像处理软件、X 线骨折辅助检测软件等，国内已有 100 多个产品获得 3 类医疗器械注册证，还有两百多个

产品获得 2 类医疗器械注册证。

（六）可能涉及的相关知识产权情况

本标准项目为自主制定项目，未识别到相关的专利情况。

（七）征求国务院有关部门或关联 TC 意见的情况

标准化对象主要为医疗器械，由国家药品监督管理局负责实施监督。

（八）经费预算

本标准项目主要依托自筹经费，预算估计如下：

1. 出版费 1.0 万元
2. 会议费 4.0 万元
3. 专家咨询费及人员劳务费 2.5 万元
4. 标准验证费 1.0 万元

总计：8.5 万元

（九）项目进度安排

本项目预计在 12 个月内完成，工作进度规划如下：

1. 项目预研：75 天
2. 调研起草：75 天
3. 征求意见：60 天
4. 修订送稿：30 天
5. 技术审查：30 天
6. 整理与报批：30 天

（十）需要申报的其他事项

无。