

行业标准《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》标准编制说明

一、工作简况

1. 任务来源

根据《国家药监局综合司关于印发2020年医疗器械行业标准制修订计划项目的通知》（药监综械注〔2020〕48号），人工智能医疗器械标准化技术归口单位（以下简称为归口单位）和中国食品药品检定研究院（以下简称为中检院）承担《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》标准（以下简称为数据集标准）的制定任务，项目编号为：A2020005-T-ZJY。

2. 协作单位

本标准的制定工作由人工智能医疗器械标准化技术归口单位提出，起草组包括产学研医检各个领域的15个成员单位，具体由中国食品药品检定研究院、上海长征医院、解放军总医院、中国科学院深圳先进技术研究院、华为技术有限公司、飞利浦（中国）投资有限公司、上海西门子医疗器械有限公司、杭州依图医疗技术有限公司、腾讯医疗健康（深圳）有限公司、上海联影智能医疗科技有限公司、慧影医疗科技（北京）有限公司、广州柏视医疗科技有限公司、美中互利医疗有限公司、北京安德医智科技有限公司、东南大学共同完成。

3. 工作过程：

本标准起草小组具体工作纪要如下：

- 2020年4月10日，中检院组织召开起草组网络会议，为期半天。会议宣布起草组的正式成立，向各参与单位介绍了医疗器械标准制修订管理规定、一般流程和时限，详细讲解了本标准草案的初始版本，对各家的任务分工进行了梳理，明确了标准起草的流程和时间节点。
- 2020年4月17日，中检院组织召开人工智能医疗器械标准化技术归口单位标准工作网络会议，归口单位专家、观察员、秘书处工作人员和标准起草相关人员参加，为期半天。本次会议启动了本年度的标准制定工作，向归口单位专家组介绍了本行业标准草案的起草情况并进行研讨。
- 2020年4月30日，中检院组织召开起草组网络会议，为期半天。会议讨论了各家对标准草案提出的修改意见，对草案的最新版本进行了审议。经讨论，起草组对两个草案形成十六条修改意见，进入细化修改阶段。
- 2020年6月16日，中检院组织召开起草组网络讨论会，会议对修改稿进行了进一步讨论，会后对标准进行了进一步完善和修改，形成标准的征求意见稿。

二、标准编制原则和确定标准主要内容的依据

1. 标准制定的意义、原则

数据集是人工智能医疗器械研发、测试、监管的重要资源，对人工智能医疗器械产品质量有关键影响。当前，用于人工智能医疗器械的数据集开发十分活跃，但质量控制较为薄弱，对数据集质量特性缺乏系统性的描述和要求，对数据集质量的验证缺乏客观定量的指标与方法。

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。

行业内不同厂商、医院、科研机构都在从事数据集的建设，在数据集的描述、评价与质控等方面也存在普遍的差异。

基于上述现状，本标准描述了人工智能医疗器械所使用的数据集的质量通用要求。该基础标准的实施将形成统一的质量评价规范，促进人工智能医疗器械领域数据集的整体建设与质控能力，并且带动人工智能医疗器械各细分领域数据集的规范化。

本标准按照GB/T1.1-2020的起草规则编写。

2. 本标准性能指标制定依据，对于有争议指标的处理及验证情况

本标准是人工智能医疗器械数据集质量相关的基础标准，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数据库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。

与数据库质量要求有关的规范性引用文件和参考文献包括YY/T 0287：2017《医疗器械 质量管理体系 用于法规的要求》、YY/T 0316-2016《医疗器械风险管理对医疗器械的应用》、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》。

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 Co-operation 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、 主要试验（或验证）的分析、综述报告、技术经济论证、预期的经济效果

本标准将选取4个数据集开展技术验证，包括单模态训练数据集、封闭式性能独立测试集、多模态医学影像数据集和企业自用数据集，技术验证由相关医院、企业和检测机构完成。

验证内容包括：

1. 数据集文档检查：根据本标准第4章的要求，检查数据集文档内容，验证各条款的适用性。
2. 数据质量特性评价：针对本标准第5章提出的各项质量测度，采用计数检验、计量检验、过程验证、测试用例等方法，对具体数据集的质量进行评价，验证方法的可行性与科学性。
3. 数据集风险分析：针对数据集偏倚等问题，对具体数据集开展风险分析，形成研究资料，验证偏倚控制的可行性。

各验证单位上述内容基础上，对执行标准的成本与收益进行分析，形成综述报告。本标准预期提升人工智能医疗器械数据集开发与评价活动的规范性，促进后续各细化领域数据集的标准化，为人工智能医疗器械产业发展提供质量保障。

四、 采用国际标准和国外先进标准的程度，以及与国际、国外同类标准水平的对比情况，或与测试的国外样品、样机的有关数据对比情况

本标准与国际标准 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》的相关理念、方法描述无冲突。

本标准的制定参考了国内标准 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》、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》、GB/T 11457-2006《信息技术 软件工程术语》、YY/T 0287:2017《医疗器械 质量管理体系 用于法规的要求》、YY/T 0316-2016《医疗器械风险管理对医疗器械的应用》、WS/T 305-2009《卫生信息数据集元数据规范》等涉及数据质量的标准，无冲突。

五、 与有关现行法律、法规和强制性标准的关系

本标准与有关现行法律、法规和强制性标准均无冲突。

六、 重大分歧意见的处理经过和依据

未出现重大意见分歧。

七、 行业标准作为强制性行业标准或推荐性行业标准的建议

人工智能医疗器械发展变化迅速，涉及的数据模态、参数配置、数据采集条件、访问方式、标注流程等存在差异，牵头建设数据集的机构多种多样。本标准主要提出数据集质量的通用要求，建议成为推荐性行业标准。

八、 贯彻行业标准的要求和措施建议（包括组织措施、技术措施、过渡办法等内容）

建议本标准的过渡期设为1年，本归口单位拟在标准发布后实施前进行宣贯，帮助产业开展应用示范。

九、 废止现行有关标准的建议

无。

十、 其它应予说明的事项

无其他说明。

人工智能医疗器械标准化技术归口单位

2020年6月23日