

## 专论——人工智能医疗器械质量管理体系专栏（上）

**编者按：**人工智能医疗器械（Artificial Intelligence Medical Device, AIMD）已经在辅助诊断、治疗、筛查、图像重建等多个医疗场景得到了良好的应用，并有一些产品在国内外获批上市。作为一种特殊的医用软件产品，AIMD在生产企业的质量管理体系上具有特殊性，例如数据质量的控制、标注过程的规范、变更迭代的验证等都备受业界关注。本专栏特邀AIMD领域相关的质量专家、科研学者以及生产企业开发人员，从质量管理体系的关键控制要素、标准化趋势、代表产品研发过程实践等角度进行讨论和分享，希望为产业界不断提升质量管理水平、不断自我提升和进步提供参考。



栏目主编：刘毅

刘毅，高级工程师，曾任职国家食品药品监督管理局医疗器械司、市场监督司，曾任北京市医疗器械检验所所长，现任复星集团全球合伙人、复星医药副总裁、医疗器械事业部董事长兼CEO，并任第一届人工智能医疗器械标准化技术归口单位专家组成员和第一届医用机器人标准化技术归口单位专家组成员。作为医疗器械质量管理和产业实践专家，曾带领团队制定医疗器械国家/行业标准制定260余项、承担国家等各级研究课题70余项。

# 人工智能医疗器械质量管理标准化趋势分析

王浩，李澍，王晨希，唐桥虹，李佳戈，李静莉

中国食品药品检定研究院医疗器械检定所，北京 102629

**[摘要]** 目的 旨在研究人工智能（Artificial Intelligence, AI）医疗器械质量管理的发展趋势，促进标准规范研究。方法 根据AI医疗器械相关监管、法规的最新动态，梳理质量管理面临的特殊问题；结合产品检测实践情况，分析AI医疗器械质量管理的需求与发展方向。结果 在AI医疗器械的质量管理过程中，可溯源性、数据集管理、质量控制等方面都需进一步增强。结论 本文研究的思路和方法有助于建立适合AI医疗器械的专用标准规范。

**[关键词]** 人工智能医疗器械；质量管理；溯源；质量控制

## Trend Analysis of Quality Management Standardization of Artificial Intelligence Medical Devices

WANG Hao, LI Shu, WANG Chenxi, TANG Qiaohong, LI Jiage, LI Jingli

Institute for Medical Device Control, National Institutes for Food and Drug Control, Beijing 102629, China

**Abstract:** Objective To study the development trend of quality management of artificial intelligence (AI) and promote the research of standards and specifications. Methods The special issues faced by quality management of AI medical devices were summarized according to the latest AI medical devices related regulations. The demand and development direction of quality management of AI medical devices were analyzed in combination with the practice of product testing. Results In the process of quality management of AI medical devices, traceability, data set management, quality control and other aspects need to be further enhanced. Conclusion The ideas and methods of this paper are helpful to establish the special standards for AI medical devices.

**Key words:** artificial intelligence medical devices; quality management; traceability; quality control

[中图分类号] R197.39; TP18

[文献标识码] A

doi: 10.3969/j.issn.1674-1633.2021.03.004

[文章编号] 1674-1633(2021)03-0020-04

## 引言

近年来，人工智能（Artificial Intelligence, AI）医疗

器械发展迅速，上市产品数量不断增加，核心算法的迭代更新频率加快。AI产品的质量管理框架需要创新，以便在满足产品的安全有效的前提下促进产品的迭代更新。质量管理过于严苛，会影响产品的竞争力，限制行业发展；质量管理过于松散，又会降低产品质量，带来临床风险。因此，如何开展AI产品质量管理是监管与标准化的重要研究

收稿日期：2021-01-26

基金项目：国家重点研发计划课题（2019YFC0118801）。

通信作者：李静莉，主任药师，主要研究方向为医疗器械检定和医疗器械标准。

通信作者邮箱：lijli@nifdc.org.cn

对象,对行业发展有重要意义。

目前, AI 医疗器械的质量管理尚未形成统一标准,各国的研究进展不同。我国药品监管部门在通用医疗器械、医疗器械软件质量管理方面已开展大量工作<sup>[1-2]</sup>,在 AI 医疗器械软件的注册审评方面也发布了相关技术文件<sup>[3-4]</sup>。美国食品药品监督管理局 (Food and Drug Administration, FDA) 在今年发布了 AI/ 机器学习医疗器械软件行动计划<sup>[5]</sup>,继续研究产品快速变更的监管路径<sup>[6]</sup>,构建良好机器学习框架<sup>[7]</sup>。欧盟从伦理角度出发,以“可解释”“可信赖”为关键词,构建 AI 产品的监管框架<sup>[8]</sup>。

由于 AI 产业刚刚起步,产品上市后召回事件、企业飞行检查通报等监管数据比较缺乏, AI 企业在质量管理方面的问题缺乏系统的梳理。产业在实践环节还存在一些问题和困惑需要解决。本文对 AI 医疗器械质量管理面临的特殊问题进行了梳理,结合相关的行业与监管情况,对 AI 医疗器械专用的标准规范研究提出了建议,旨在帮助行业提升质量管理能力。

## 1 AI 特殊问题分析

与传统医疗器械相比,当前 AI 医疗器械存在以下特点:

① 内核采用以深度学习为代表的新一代 AI 算法,缺乏可解释性,工作过程比较抽象,难以直观理解;② 产品研发过程比较开放,前沿学术成果的转化速度快,调用的算法框架、工具包等技术资源也处于快速变化中,例如业内常见的 pytorch 深度学习框架从 2016 年问世以来,版本更新了 50 多次<sup>[9]</sup>;③ 产品的验证和确认在方法、指标、测试集等方面差异较大,不同企业之间缺乏可比性。例如,近年来美国已上市的部分计算机辅助诊断/检测类产品的算法测试集规模从几十到几千不等,性能指标、受试者响应曲线的定义和构造方式也有差别<sup>[10]</sup>;④ 常见形态为医疗器械独立软件,开始向嵌入式组件、AI 芯片扩展。软件的部署与配置方式灵活多变,包括本地服务器、云平台、移动计算平台、芯片即系统 (System on a Chip)、软件即服务 (SaaS) 等。在当前 AI 医疗器械特点背景下,现行的医疗器械软件监管与法规,对 AI 医疗器械的质量管理具有宏观的指导意义<sup>[11-12]</sup>,例如强调全生命周期管理、软件生存周期的理念等。但在具体实施中,部分细节尚未明确,有待研究。本文依据医疗器械监管、AI 技术资源发展的历史,从更大的维度发掘线索,建议关注以下特殊问题。

### 1.1 可追溯性

本文对美国 2016 至 2020 年医疗器械的召回事件<sup>[13]</sup>进行了简要分析,汇总了由于软件、算法、数据问题直接导致的召回事件分布如图 1 所示。根据 5 年的总数量,出现质量问题的频率从高到低排名为软件 > 数据 > 算法。

AI 产品与软件、算法、数据都有密切联系,结合其中的具体案例考虑,在以下方面宜加强可追溯性。

(1) 数据的可追溯性。在图 1 所示的医疗器械召回事件中,与数据有关的常见现象包括数据丢失、数据错误、

信息记录混乱等,说明相关产品在数据完整性、数据管理等方面存在缺陷,质量管理需要加强数据的可追溯性。

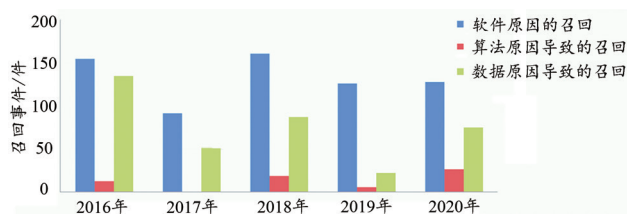


图1 美国医疗器械召回事件统计

数据集是 AI 医疗器械全生命周期使用的重要资源,对产品质量有重要影响。目前, AI 医疗器械行业不断提升对数据集的认识和标准建设<sup>[14-16]</sup>,包含了对数据集本身的版本控制、标志标识的要求,以及数据集开发管理过程的可追溯性要求。在实际的生产研发活动中,需要根据数据集的作用和角色,明确数据采集、标注、存储、使用等各个环节的具体过程记录,对人员、设施、工具、操作规程等要素进行覆盖,对医疗器械产品本身的数据处理、数据管理能力进行溯源。

(2) 算法和软件的可追溯性。在图 1 所示的医疗器械召回事件中,部分现象包括软件及组件的过期/擅自更新/不兼容、算法异常、计算错误等,从软件系统到 GPU 都有所涉及。上述缺陷说明算法、软件的质量管理需要加强。

根据医疗器械软件注册相关指导文件<sup>[11]</sup>,软件生存周期均应开展可追溯性分析。对 AI 医疗器械而言,算法的来源和功能较为丰富,包括企业本身、开源平台、外部供应商等渠道,以及联邦学习、迁移学习等其他方式。在研发与测试阶段,算法架构、参数配置和对应的源代码、软件组件是动态变化和完善的,可追溯的难度较大。虽然软件产品整体有版本控制的要求<sup>[17]</sup>,源代码、软件组件、核心算法的版本控制仍需加强。

### 1.2 风险与缺陷管理

AI 产品的风险与缺陷,由多种因素决定,包括自身设计、外部组件、硬件等。根据公开数据<sup>[18]</sup>, AI 研发相关的技术资源多次出现安全漏洞如图 2 所示。其中, Notebook、TensorFlow、MongoDB、GPU、Python 分别是开发工具、深度学习框架、数据库、核心硬件和编程语言的代表。在此情况下, AI 产品的制造商在设计开发、设计变更、供应商资质审核等方面应广泛提高警惕。

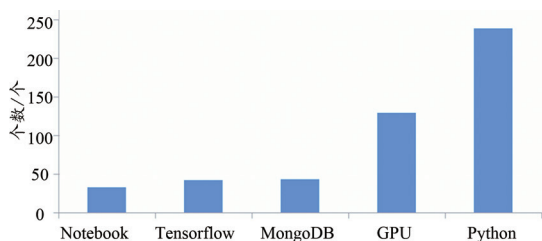


图2 近3年公布的安全缺陷一览表

此外,在临床部署后,产品的性能会随着输入数据的波动而变化。这种波动来自真实世界的数据采集设备、人

员操作、外界干扰、受试人群等各个方面,具有随机性和发散特性。质量管理体系需要对产品的真实世界表现进行监测,需要应对和处理不良事件,明确何种情形下对产品进行召回<sup>[19]</sup>。

### 1.3 更新控制

由于 AI 行业鼓励算法的迭代更新,为保障 AI 医疗器械算法更新后的安全有效,从管理的角度意味着需要加强更新控制,尤其是对算法性能开展验证和确认<sup>[20-21]</sup>。

在测试环节,应加强回归测试的管理。回归测试是使用相同的测试用例或测试集对迭代前后的算法进行测试。如果算法迭代频繁,应保障测试集动态管理的资源需求。另一方面,当测试集发生更新时,当前最新版本的算法应立即开展测试,作为性能的基线,用于和后续版本的算法进行比对。同时,为了充分验证算法更新后的泛化能力,除了开展回归测试外,宜动态引入新的测试数据,扩展算法验证与确认的维度。例如对现有测试集添加噪声和干扰,合成“对抗测试集”,用于比较新旧算法的抗扰能力。

## 2 来自实践的启示

在 AI 医疗器械的性能验证环节,实践中发现一些常见问题,容易影响验证工作的进程如图 3 所示。这些问题涵盖了检品的安装、部署、运行、设计等方面,从侧面反映了 AI 产品质量管理的薄弱之处,建议从以下方面予以加强。

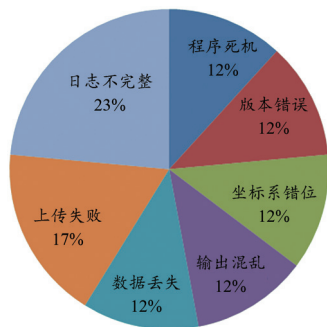


图3 产品检测实践中发现的问题

(1) 样品管理。测试前,样品的代码编译、发布、程序部署等环节需要做好验证记录,确保将正确的版本投入测试,将算法版本与测试结果之间形成关联;测试中,样品宜形成和显示过程记录,如当前测试进度、当前测试样本编号、剩余测试时间等,提高测试过程的可追溯性;测试后,应检查内存、存储空间使用情况,检查算法输出结果的格式、顺序是否符合预期,形成记录。

(2) 数据管理。参照行业标准《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(报批稿)》<sup>[14]</sup>,数据采集、标注、存储、流通、停用等环节均应建立记录;数据标注的最终结论或形成的参考标准应能追溯至原始标注结论、仲裁与审核结果。数据采集、标注人员的资质要求、选拔、培训、活动记录应当齐全。

(3) 评价指标管理。在产品的测试环节,随着测试集的抽样与组合,评价指标可能是动态变化的,但需要与产

品的设计输入、需求分析、风险分析之间保持协同。评价指标的确定和改变需要进行评审并建立记录。对于鲁棒性等评价指标,在风险分析环节应明确描述其对应的真实世界风险。

## 3 讨论与总结

本文根据国外医疗器械上市后监管数据和行业安全数据,对 AI 医疗器械质量管理的特殊性进行了分析,并结合在产品验证实践中发现的问题,对产品质量管理提出建议,主要强调可追溯性、数据管理、质量控制等角度:①可追溯性方面:需要专业的技术规范保证 AI 医疗器械使用的数据集、源代码、软件组件的可追溯性,例如唯一标识编码等,考虑各种技术要素的生成时间、修改时间、使用地点、执行机构、版本等信息,密切跟踪软硬件、组件的版本更新、安全更新等,做好验证、确认和变更记录控制。同时,为了加强算法的可追溯性,宜细化明确算法研发记录的要求,例如参与研发/测试的人员记录、训练记录与过程参数、训练集/调优集的分配记录等;②数据管理方面:训练/调优、封闭测试、临床试验、真实世界监测等各种用途的数据集均应建立专用的管理体系与过程记录,明确管理职责和人员资质,并提供相应的存储与安全设施,保证数据集的版本控制和可重复使用,能与算法的快速迭代相适应。产品、研发工具的数据管理、数据处理能力,同样应纳入考量范围;③在产品质控方面,应加强质量评价方法标准的研究,把真实世界风险评估的关口前移,增强上市前的验证与确认,模拟产品上市后面临的风险;在产品临床部署后,提高日常质控的完备性,设立产品性能监测指标、接口和工具,为用户反馈和产品变更提供可靠的、持续的客观数据。

AI 医疗器械属于新兴事物,监管的历史较短,公开数据较少,制约着质量管理研究。本文参考了近 5 年来国外医疗器械召回事件中与软件、算法、数据相关的案例,丰富了信息收集的渠道,为后续研究提供了有益的素材。同时,本文对 AI 产品研发资源自身的更新、安全漏洞等情况进行了量化分析,侧重 AI 自身特色,为质量管理研究提供了新的线索。

综上所述,AI 医疗器械的质量管理还有很大的进步空间,需要加强算法、数据的可追溯性,提高数据管理能力,提升对产品的全生命周期质控能力,需要相关基础标准、方法标准的协同与支撑。

### 【参考文献】

- [1] 国家药品监督管理局.医疗器械生产质量管理规范附录独立软件[Z].北京:国家药品监督管理局,2019.
- [2] YY/T 0287-2017,医疗器械 质量管理体系用于法规的要求[S].
- [3] 国家药品监督管理局医疗器械技术审评中心.深度学习辅助决策医疗器械软件审评要点[Z].北京:国家药品监督管理局医疗器械技术审评中心,2019.
- [4] 国家药品监督管理局医疗器械技术审评中心.肺炎CT影像辅



助分诊与评估软件审评要点(试行)[Z].北京:国家药品监督管理局医疗器械技术审评中心,2020.

- [5] US Food and Drug Administration. Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan[Z]. Washington, DC.: U.S. Food and Drug Administration, 2021.
- [6] US Food and Drug Administration. Proposed Regulatory framework for modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback[Z]. Washington, DC.: US Food and Drug Administration, 2019.
- [7] Xavier Health. Perspectives and good practices for AI and continuously learning systems in healthcare[Z]. Cincinnati, OH.: Xavier Health, 2018.
- [8] European Commission. Whitepaper on artificial intelligence-a European approach to excellence and trust[Z]. Brussels: European Commission, 2020.
- [9] Pytorch. Tags.[EB/OL].[2021-01-31]. <https://github.com/pytorch/pytorch/tag>.
- [10] Wang L, Wang H, Xia C, et al. Towards standardized premarket evaluation of computer aided diagnosis/detection products: insights from FDA-approved products[J]. *Expert Rev Med Devic*, 2020, 17:899-918.
- [11] 国家药品监督管理局医疗器械技术审评中心.《医疗器械

软件技术审查指导原则(第二版征求意见稿)》[Z].北京:国家药品监督管理局医疗器械技术审评中心,2020.

- [12] YY/T 0664-2020, 医疗器械软件 软件生存周期过程[S].
- [13] US Food and Drug Administration. Medical device recalls [EB/OL].[2021-01-31]. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRES/res.cfm>.
- [14] 中国食品药品检定研究院. 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(报批稿)[S].北京:中国食品药品检定研究院,2020.
- [15] 王权,王浩,孟祥峰,等.人员管理对人工智能医疗器械用数据集质量的影响分析[J].中国医疗设备,2018,33(12):6-9.
- [16] 王浩,孟祥峰,李澍,等.数据集在人工智能医疗器械质控中的角色与要求[J].中国医疗器械杂志,2019,43(1):54-57.
- [17] 彭亮,刘晓燕.软件版本与医疗器械软件监管[J].中国医疗器械杂志,2015,(1):48-50.
- [18] NIST. National Vulnerability Database[EB/OL].[2021-01-31]. <https://nvd.nist.gov/vuln/search/>.
- [19] 彭亮,于一.医疗器械的软件相关召回分析[J].中国医学装备,2016,13(7):97-99.
- [20] 蔡小舒,吕晖,于广军.美国FDA医疗人工智能软件审核指南研究[J].中国数字医学,2019,14(11):34-37.
- [21] 《中国食品药品监管》杂志编辑部.基于人工智能和机器学习技术的医疗器械软件[J].中国食品药品监管,2019,(6):59-61.

本文编辑 崔丽君

#### 上接第10页

between the electrode configuration and histopathology of irreversible electroporation ablations in prostate cancer patients[J]. *World J Urol*, 2016, 34(5):657-664.

- [59] Neal RE, Millar JL, Kavnoudias H, et al. In vivo characterization and numerical simulation of prostate properties for non-thermal irreversible electroporation ablation[J]. *Prostate*, 2014, 74(5).
- [60] Murray KS, Ehdaie B, Musser J, et al. Pilot study to assess safety and clinical outcomes of irreversible electroporation for partial gland ablation in men with prostate cancer[J]. *J Urology*, 2016, 196(3):883-890.
- [61] Van Den Bos W, Jurhill RR, De Bruin DM, et al. Histopathological outcomes after irreversible electroporation for prostate cancer: results of an ablate and resect study[J]. *J Urology*, 2016, 196(2):552-559.
- [62] Bos VD, Jurhill RR, de Bruin DM, et al. Histopathological outcomes after irreversible electroporation for prostate cancer: results of an ablate and resect study[J]. *J Urology*, 2016.
- [63] Golberg A, Bruinsma BG, Uygun BE, et al. Tissue heterogeneity in structure and conductivity contribute to cell survival during

irreversible electroporation ablation by "electric field sinks"[J]. *SCI REP-UK*, 2015, 5:8485.

- [64] Cohen E I, Field D, Lynskey GE, et al. Technology of irreversible electroporation and review of its clinical data on liver cancers[J]. *Expert Rev Med Devic*, 2018, 15(2):99-106.
- [65] Au JT, Wong J, Mittra A, et al. Irreversible electroporation is a surgical ablation technique that enhances gene transfer[J]. *Surgery*, 2011, 150(3):474-479.
- [66] Yang Y, Qin Z, Du D, et al. Safety and short-term efficacy of irreversible electroporation and allogenic natural killer cell immunotherapy combination in the treatment of patients with unresectable primary liver cancer[J]. *Cardiovasc Inter Rad*, 2019, 42(1):48-59.
- [67] Sano MB, Fan RE, Cheng K, et al. Reduction of muscle contractions during irreversible electroporation therapy using high-frequency bursts of alternating polarity pulses: a laboratory investigation in an ex vivo swine model[J]. *JVIR*, 2018, 29(6):893-898.e4.

本文编辑 崔丽君