

论医学人工智能时代的移动健康终端质量控制

王浩, 李澍, 孟祥峰, 任海萍* (中国食品药品检定研究院, 北京 100050)

摘要 目的: 梳理医学人工智能时代的移动健康终端在质量控制方面的要求和问题。方法: 查阅国内外监管政策文献, 从国外公开的医疗器械审批资料中查询带有移动健康终端的智能化医疗器械在验证与确认环节开展的具体评价工作。结果: 移动健康终端的质量评价要求在国内存在差异, 国内在电气安全方面主要参考医用电气设备, 而美国可执行其他标准。结论: 本文讨论了移动健康终端对人工智能医疗器械的影响, 明晰了移动健康终端的质量评价思路, 更好地保证基于移动健康终端的人工智能医疗器械的有效性和安全性。

关键词: 人工智能医疗器械; 质量控制; 移动健康终端

中图分类号: TP181; TH77 文献标识码: A 文章编号: 1002-7777(2019)09-1009-06

doi:10.16153/j.1002-7777.2019.09.008

Discussion on Quality Control of Mobile Health Terminals in the Era of Medical Artificial Intelligence

Wang Hao, Li Shu, Meng Xiangfeng, Ren Haiping* (National Institutes for Food and Drug Control, Beijing 100050, China)

Abstract Objective: To study the requirement and issues with quality control of mobile health terminals in the era of medical artificial intelligence. **Methods:** Regulation policy literatures of China and other countries were searched and reviewed. Specific evaluation of validation and verification of intelligent medical devices with mobile health terminals was collected from published review documents of medical devices of other countries. **Results:** The requirement of quality evaluation of mobile health terminals in China was different from that in other countries. Compliance with medical electrical equipment is required in China, while USA applies other standards. **Conclusion:** This paper discusses the impact of mobile health terminals on artificial intelligence medical device and clarifies the requirement of quality evaluation of mobile health terminals so as to ensure the safety and effectiveness of the artificial intelligence medical device based on mobile health terminals.

Keywords: artificial intelligence medical device; quality control; mobile health terminal

近年来, 在深度学习算法^[1]的推动下, 人工智能医疗器械发展迅速, 预期用途多种多样, 在辅助诊断^[2-3]、辅助检测^[4-5]、辅助分诊^[6]、临床辅助决策^[7]等领域的研究不断取得突破, 在国外已有一批产品获批上市。随着人工智能芯片的发展与移动终

端产品硬件配置的提高, 人工智能算法在移动健康终端上的部署和应用成为现实, 使人工智能医疗器械的应用领域扩展到院外和家庭, 处理对象扩展到非医疗条件下的个人健康数据, 而这给人工智能医疗器械的质量控制带来新的挑战。

基金项目: 国家重点研发计划课题(编号 2017YFC0111004)

作者简介: 王浩; 研究方向: 生物医学工程、人工智能、有源医疗器械检定

通信作者: 任海萍; 研究方向: 生物医学工程、医疗器械检定; E-mail: renhaiping@nifdc.org.cn

由于人工智能算法的性能与数据质量的关系紧密,人工智能医疗器械的质量控制应当综合考虑硬件、软件、数据集等多种因素。由于移动健康终端产品尚未完全纳入医疗器械监管范畴,在质量标准、技术要求、合规性等方面的强制要求不同于医疗器械。当移动健康终端成为人工智能医疗器械的数据来源或运行平台时,终端本身的质量风险可能影响该类医疗器械的有效性与安全性。

为加强人工智能时代基于移动健康终端的医疗器械的质量控制,本文对国内外移动健康终端产品监管思路^[8-10]进行了回顾,对基于移动健康终端的医疗器械产品进行了案例分析,对移动健康终端的质量风险进行了讨论,对加强质控工作、更好地保障相关人工智能医疗器械的有效性和安全性提出了建议。

1 移动健康终端产品监管思路回顾

简单地说,移动健康终端产品与医疗器械之间的界定主要由预期用途决定。当移动健康终端产品的预期用途符合医疗器械定义时,应作为医疗器械进行监管。在具体实施方面,国内外思路有所差别。

1.1 国外监管思路

美国食品药品监督管理局(Food and Drug Administration, FDA)对移动健康产品发布了两个指导原则^[8-9],分别解释了低风险通用健康产品^[8](Low Risk General Wellness Product)和移动医疗应用^[9](Mobile Medical Application)的监管要求,体现了移动健康终端的监管思路。

前者提出了通用健康产品的概念,明确通用健康产品不作为医疗器械进行监管。通用健康产品的定义包含两个要素:1)仅用于通用健康目的;2)使用者的安全风险较低。符合定义的预期用途包括对健康状态的维持、对健康活动的鼓励、对健康生活方式的鼓励、心理压力缓解等。具体表现形式包括体重管理、饮食管理、运动管理、精神鼓励,以及用于降低慢性疾病(比如糖尿病、高血压等)风险的生活方式调整等。其中,用于运动状态监控的心率测量功能也被列入通用健康产品行列,这意味着日常生活中常见的能测量心率的手环、手表等产品属于通用健康产品,可以不作为医疗器械监管。

后者提出了移动计算平台和移动医疗应用的

概念。根据相应指导原则^[9],日常生活中的智能手机、平板电脑、智能手表等移动终端都属于移动计算平台,不按照医疗器械监管。部署在这些平台上的APP符合医疗器械定义时,称之为移动医疗应用。移动医疗应用的预期用途有两种:1)用于成为其他医疗器械的组件;2)把移动计算平台转化为医疗器械。第二种预期用途尤其值得关注,它的意义在于当移动计算平台成为移动医疗应用的载体时,应当按照医疗器械进行评价,而不是继续游离于医疗器械监管之外。

1.2 国内监管思路

我国的监管思路同样从预期用途出发^[10],发布过《移动医疗器械注册技术审查指导原则》,其中规定符合医疗器械定义的移动计算设备或软件属于移动医疗器械,具体包括移动医疗设备、移动独立软件和移动医疗附件等形式。与美国类似,预期用于健康人群的健康管理、生活方式管理、基于移动计算平台的健康信息统计等功能不属于医疗器械,而预期用于医护人员的疾病状态管理、对医疗器械的控制和数据交换等功能属于医疗器械。从硬件角度,对移动医疗设备的监管,除考虑其等效的传统医疗器械的要求之外,还需综合考虑移动计算技术的特点和风险^[10],比如显示屏、电池、数据传输等问题,以及网络安全、云计算、大数据等合规性要求。从软件角度,移动独立软件的风险管理应考虑传统独立软件和移动计算平台的特殊风险。移动医疗附件需遵照移动医疗器械整体的申报要求。

相对于美国,我国对移动医疗器械用到的移动健康终端的考量更严格,不仅明确了移动医疗设备应满足等效的传统医疗器械要求,还增加了显示屏、电池、数据传输等特殊要求。关于移动医疗器械终端产品的质控研究也在推进^[11-12]。

2 基于移动健康终端的智能化医疗器械案例与问题分析

随着人工智能算法的发展,基于移动健康终端的智能化医疗器械在美国已有多款产品获批上市,其公开资料为研究移动健康终端的质控要求提供了有益的参考,因此本文列举近年来的若干案例进行分析,并讨论了智能化移动健康终端的特殊质控问题。

2.1 软硬件结合案例

软硬件结合案例见表1。

表1 软硬件结合的智能化医疗器械

产品名称	预期用途	产品组成	计算平台	建议关注的部分评价内容
DIP.IO ^[13]	尿液分析 ^[13]	手机APP、比色板、尿液分析试剂条 ^[13] (FDA已单独批准)	LG Nexus 5 智能手机 ^[13]	性能: 重复性、复现性、准确率、线性度、探测极限等; 软件特性; 可用性; 临床试验 ^[13]
Embrace ^[14]	监控皮肤电活动, 检测癫痫发作的征兆 ^[14]	可穿戴硬件(皮肤电活动传感器、加速传感器、蓝牙通信、数据处理芯片)、手机APP ^[14]	自身(数据处理)、手机(发送预警) ^[14]	电气安全(医用电气设备) ^[15] 、电磁兼容 ^[16] 、医疗器械生物学评价 ^[14] 、临床试验 ^[14]
Kardia Band ^[17]	单导联心电信号的采集、传输、显示和心律失常分析 ^[17]	硬件: 表带(电极) 软件: Apple Watch app; iPhone app ^[17]	Apple Watch、iPhone ^[17]	电气安全 ^[15] 、电磁兼容 ^[16] 、心电安全与性能专标 ^[18] 、临床试验 ^[17]

表1所示为3个有代表性的、软硬件结合的智能化医疗器械, 每一行描述同一个产品的信息。其共同特征是产品组成既包含移动APP(以及智能算法), 又包含硬件组件; 产品的评价均参照了传统医疗器械的要求, 涉及软件特性、硬件性能与安全、临床试验等环节, 涉及的具体标准都是医疗器械领域的常见标准, 表1列出了笔者建议关注的部分内容。其中DIP.IO的比色板、Embrace的传感器和Kardia Band的表带等硬件组件都进行了上市前的验证与确认, 而DIP.IO的尿液分析试剂条属于已批准的医疗器械, 不再重复验证。

从硬件的质控角度来看, Embrace和Kardia Band相对全面, 而DIP.IO存在一定的薄弱, 问题在

于尿液分析环节需要使用手机对试纸和比色板进行拍照, 但并未对手机进行单独考量, 而是把手机、试纸、比色板和算法作为一个整体进行性能评价。从光学角度和算法性能角度看, 手机光源的色温、显色指数、照度、图像质量等因素对于最终的尿液分析结果可能有重要影响, 不同品牌、不同使用年限、不同状态的手机可能带来结果的变化。这可能是制造商规定手机型号的原因之一。从全生命周期的角度看, 手机的软硬件更新、电磁兼容等问题都有可能影响预期用途的实现, 有必要加强质控。

2.2 独立软件/APP案例

独立软件/APP案例见表2。

表2 独立的移动医疗应用

产品名称	预期用途	信号来源	计算平台	建议关注的部分评价内容
ECG App ^[19]	检测房颤和窦性心律 ^[19]	Apple Watch ^[19] (手表本身在美国不作为医疗器械)	Apple Watch (数据分析)、iPhone (数据管理与传输) ^[19]	电磁兼容 ^[16] 、电气安全(视频、音频和信息通讯设备) ^[20] 、心电设备专标 ^[18] 、软件特性 ^[19] 、心电算法 ^[19] 、可用性 ^[19] 、临床确认 ^[19]
EMurmur ID ^[21]	心音信号的记录、显示、杂音分析和存储 ^[21]	电子听诊器, 型号Littmann 3200 (已批准的医疗器械) ^[21]	智能手机 ^[21]	电气安全(医用电气设备) ^[15] 、电磁兼容标准 ^[16] 、无线共存、软件特性、人工智能算法性能、临床确认 ^[21]
CardioLogs ECG Analysis Platform ^[22]	心电信号的下载、处理和心律失常分析 ^[22]	动态心电图机、12导联心电图机、事件记录器等各种用于心律失常诊断的医疗器械 ^[22]	电脑、云平台、心电设备 ^[22]	心电专标 ^[18] 、软件生命周期 ^[23] 、心电图机电气安全 ^[24] 、可用性、心电算法、临床确认 ^[22]

表2所示为3个有代表性的智能化移动医疗应用,每一行描述同一个产品的信息。它们都作为独立软件获得批准,其评价要求都包括了对算法的评价(使用数据集)、软件特性评价、临床确认。与传统医疗器械独立软件不同,它们的申报资料中还包含对相关硬件进行的评价,但评价方式与角度不同,表2列出了笔者建议关注的部分内容。

EMurmur ID和CardioLogs ECG Analysis Platform各自使用了FDA已批准的医疗器械进行数据采集,但在申报资料中仍然体现了对硬件的评价。前者的申报资料中明确了电气安全、电磁兼容、无线共存的符合性,也对性能进行了评价。后者的申报资料体现了心电图机电气安全IEC 60601-2-25、心电性能与安全标准IEC 60601-2-47,但并未考虑电磁兼容和无线共存。出现这种差异的原因可能是EMurmur ID配合指定型号的医疗器械硬件,有条件进行全部试验,而CardioLogs ECG Analysis Platform未指定硬件,在电磁兼容和无线共存方面难以选择典型性的型号开展试验。

ECG App使用Apple Watch进行数据采集和数据分析,Apple Watch本身并不属于医疗器械。在申报资料中,电气安全的评价使用了音视频、信息与通讯设备的电气安全标准IEC 62368-1,未执行医用电气设备的IEC 60601-1标准,反映了与医疗器械硬件的差别。但另一方面,对心电采集功能的评价遵循了心电性能与安全专标IEC 60601-2-47,电磁兼容也使用了IEC 60601-1-2标准,说明从数据采集功能、电磁抗扰性的角度,Apple Watch仍然按照医疗器械的要求进行评价。在这个案例当中,ECG App实质上把移动健康终端Apple Watch变成了移动医疗设备,根据国内的监管思路,应当按照其等效的传统医疗器械进行评价,意味着在电气安全等方面的要求与国外不同。

2.3 智能化移动健康终端特殊质控问题分析

随着人工智能芯片的发展,传统的通用移动健康终端与人工智能算法进一步融合,形成智能化移动健康终端,对产品质量评价带来一些新的问题。

第一是算力问题。移动健康终端一般体积较小,由电池供电,自身功率和续航能力对人工智能芯片的功耗提出了较大的限制,因而人工智能芯片在设计时强调性能功耗比,但其算力难以达到深度

学习服务器的水平。这要求质量控制工作关注效率、最大吞吐量、最大并发数等指标,充分验证移动健康终端的算力能否满足设计输入。

第二是数据问题。智能化移动健康终端容易植入预处理算法,在进行产品检验时,需关注原始数据和预处理数据的区别,尤其是性能测试和电磁抗扰测试试验结果的比对,以充分评估预处理算法对于数据质量、产品抗扰性能的影响和对人工智能算法性能的影响。

第三是算法评价问题。移动终端的人工智能算法运行环境与服务器端或云端不同,较多采用移动平台专有的框架,以硬件底层编程的形式实现,算法的可解释性、可用性差异较大。因此,对智能化移动健康终端算法的评价有别于人工智能医疗器械独立软件,需要开发特殊的方法和测试用例。

3 对加强移动健康终端质控的建议

从以上案例可以看出,对基于移动健康终端的智能化医疗器械,无论是软硬件结合的设备,还是独立软件或APP,其质量评价都把智能算法、移动医疗应用、传感器和计算平台考虑在内,但不同产品的具体评价方面存在一定的差异和薄弱。产品组成中不属于医疗器械的部分尤其值得关注,涉及到是否进行评价、是否依据医疗器械标准进行评价等关键问题的判断,影响质量评价的完整性、科学性,有必要建立明确、统一的思路。

由于移动健康终端在这一类医疗器械中主要扮演数据采集、数据处理的角色,是人工智能算法的数据来源和算法运行的平台。其质量不仅影响到当前算法的性能表现,也影响算法的研发、训练与测试,以及产品上市后的后续表现与不良事件。从全生命周期质量控制的角度看,预期用于医学人工智能的移动健康终端,即使自身不作为医疗器械监管,也应当在性能方面进行全面的评价,以揭示和预防移动健康终端的质量风险对数据质量、移动医疗器械的最终性能与安全的影响,并提倡在复杂的环境下观测和加强智能算法的鲁棒性。同时,有必要在产品质量标准方面开展比对研究,分析通用健康产品的研发标准与医疗器械标准的区别,确定符合行业实际的技术要求,兼顾监管尺度的灵活性。

4 结论

在医学人工智能蓬勃发展的今天,移动健康终端作为人工智能算法硬件载体的可能性越来越

大,对人工智能医疗器械质量的影响也越来越重要。在以往的监管思路中,移动健康终端与移动医疗器械的界定较为模糊。从已批准的相关医疗器械申报资料来看,即使按照医疗器械进行评价,移动健康终端的质量控制仍然缺乏明确而统一的要求。本文建议从全生命周期质量控制的角度,对移动健康终端开展全面的质量评价,并积极探索合理的评价尺度,以促进行业健康快速发展。

参考文献:

- [1] Lecun Y, Bengio Y, Hinton G. Deep Learning[J]. Nature, 2015, 521 (7553): 436.
- [2] Tufail A, Rudisill C, Egan C, et al. Automated Diabetic Retinopathy Image Assessment Software: Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders[J]. Ophthalmology, 2017, 124 (3): 343-351.
- [3] Becker A S, Marcon M, Ghafoor S, et al. Deep Learning in Mammography: Diagnostic Accuracy of a Multipurpose Image Analysis Software in the Detection of Breast Cancer[J]. Investigative Radiology, 2017, 52 (7): 434.
- [4] Misawa M, Kudo S E, Mori Y, et al. Artificial Intelligence-assisted Polyp Detection for Colonoscopy: Initial Experience[J]. Gastroenterology, 2018: S0016508518304153.
- [5] Yamada Y, Shiomi E, Hashimoto M, et al. Value of a Computer-aided Detection System Based on Chest Tomosynthesis Imaging for the Detection of Pulmonary Nodules[J]. Radiology, 2017, 287 (1): 333-339.
- [6] Kwon J, Lee Y, Lee Y, et al. Validation of Deep-learning-based Triage and Acuity Score Using a Large National Dataset[J]. PLOS ONE, 2018, 13 (10): 1-10.
- [7] Barbieri C, Molina M, Ponce P, et al. An International Observational Study Suggests That Artificial Intelligence for Clinical Decision Support Optimizes Anemia Management in Hemodialysis Patients[J]. Kidney International, 2016, 90 (2): 422-429.
- [8] Center for Devices and Radiological Health. Guidance for Industry and Food and Drug Administration Staff: General Wellness: Policy for Low Risk Devices [S]. Washington DC: Food and Drug Administration, 2016.
- [9] Center for Devices and Radiological Health. Guidance for Industry and Food and Drug Administration Staff: Mobile Medical Applications[S]. Washington DC: Food and Drug Administration, 2015.
- [10] 国家食品药品监督管理总局. 移动医疗器械注册技术审查指导原则[S]. 2017.
- [11] 李澍, 王浩, 罗维娜, 等. 基于预期用途的移动健康产品监管分类要求[J]. 中国医疗设备, 2017, 32 (6): 6-8, 19.
- [12] 王浩, 李澍, 王晨希, 等. 移动医疗器械终端产品质控技术研究[J]. 中国医疗设备, 2017, 32 (6): 9-11, 19.
- [13] Center for Devices and Radiological Health. Premarket Approval (510k) Summary K173327[S]. Washington DC: Food and Drug Administration, 2018.
- [14] Center for Devices and Radiological Health. Premarket Approval (510k) Summary K172935[S]. Washington DC: Food and Drug Administration, 2017.
- [15] TC 62/SC 62A - Common Aspects of Electrical Equipment Used in Medical Practice. IEC 60601-1:2005 Medical Electrical Equipment—Part 1: General Requirements for Basic Safety and Essential Performance[S]. Geneva: International Electrotechnical Commission, 2005.
- [16] TC 62/SC 62A - Common Aspects of Electrical Equipment Used in Medical Practice. IEC 60601-1-2: 2014 Medical Electrical Equipment - Part 1-2: General Requirements for Basic Safety and Essential Performance - Collateral Standard: Electromagnetic Disturbances - Requirements and Tests[S]. Geneva: International Electrotechnical Commission, 2014.
- [17] Center for Devices and Radiological Health. Premarket Approval (510k) Summary K171816[S]. Washington DC: Food and Drug Administration, 2017.
- [18] TC 62/SC 62D - Electromedical Equipment. IEC 60601-2-47 Medical Electrical Equipment - Part 2-47: Particular Requirements for the Basic Safety and Essential Performance of Ambulatory Electrocardiographic Systems[S]. Geneva: International Electrotechnical Commission, 2012.
- [19] Center for Devices and Radiological Health. De Novo Summary DEN180044 [S]. Washington DC: Food and Drug Administration, 2018.

- [20] TC 108 – Safety of Electronic Equipment within the Field of Audio/video, Information Technology and Communication Technology. IEC 62368-1 Audio/video, Information and Communication Technology Equipment – Part 1: Safety Requirements[S]. Geneva: International Electrotechnical Commission, 2018.
- [21] Center for Devices and Radiological Health. Premarket Approval (510k) Summary K181988[S]. Washington DC: Food and Drug Administration, 2019.
- [22] Center for Devices and Radiological Health. Premarket Approval (510k) Summary K170568[S]. Washington DC: Food and Drug Administration, 2017.
- [23] TC 62/SC 62A – Common Aspects of Electrical Equipment Used in Medical Practice. IEC 62304 Medical Device Software – Software Life Cycle Processes[S]. Geneva: International Electrotechnical Commission, 2006.
- [24] TC 62/SC 62D – Electromedical Equipment. IEC 60601-2-25 Medical Electrical Equipment – Part 2-25: Particular Requirements for the Basic Safety and Essential Performance of Electrocardiographs[S]. Geneva: International Electrotechnical Commission, 2011.

(收稿日期 2019年6月24日 编辑 王雅雯)