

人工智能乳腺超声的发展现状及上市前评价制度研究

许晓萍¹ 曾祥卫^{1*} 刘婷²

¹ 广东省药品监督管理局审评认证中心 (广东 广州 510080)

² 深圳迈瑞生物医疗电子股份有限公司 (广东 深圳 518057)

文章编号: 1006-6586(2022)17-0001-03

中图分类号: R445.1

文献标识码: A

内容提要: 人工智能是目前医学研究的热点,超声医学与人工智能技术相结合,在乳腺疾病检查领域的研究日益增多。介绍和讨论人工智能乳腺超声的发展现状以及上市前评价的相关政策,为相关技术的进一步发展与应用提供有益的指导。

关键词: 乳腺超声 人工智能 发展现状 上市前评价

Research on the Development Status and Pre-Market Evaluation System of Artificial Intelligence Breast Ultrasound

XU Xiao-ping¹ ZENG Xiang-wei^{1*} LIU Ting²

¹ Center for Certification and Evaluation, Guangdong Drug Administration (Guangdong Guangzhou 510080)

² Shenzhen Mindray BioMedical Electronics, Co., Ltd., Shenzhen, China (Guangdong Shenzhen 518057)

Abstract: At present, artificial intelligence (AI) is the hotspot of medical research. The combination of ultrasonic medicine and AI technology is increasing in the field of breast disease examination. This paper introduces and discusses the development and the relevant policies of pre-market evaluation, and provides useful guidance for the further development and application of related technologies.

Key words: breast ultrasound, artificial intelligence, development status, pre-market evaluation

超声检查是中国广泛使用的乳腺病变检查方法,但常规超声检查有较强的操作者依赖性及主观性,且往往对早期病变的诊断灵敏度欠佳,容易出现漏诊误诊。随着计算机技术的发展,人工智能逐步从前沿技术转变为现实应用,人工智能结合超声进行乳腺检查的研究日益增多,相关研究致力于辅助超声医生诊断,以提高临床诊断能力和诊断效率,减少误诊漏诊。本文将介绍和讨论人工智能乳腺超声的发展以及主要国家和地区上市前评价的相关政策,为技术进一步的发展与应用提供有益的指导。

1. 人工智能乳腺超声的发展现状

临床上用于乳腺超声检查的成像模式通常包括二维灰阶成像、三维容积成像、弹性成像、造影成像等。人工智能乳腺超声的研究主要基于二维灰阶影像和三维容积影像,大多用于乳腺病变的检测和分割、特征分析、风险评估和诊断。

1.1 乳腺病变检测和分割

手工勾画病灶的边界繁琐且耗时,自动勾画乳腺超声图像中病灶的边界,可以减少医生的工作量,提高其工作效率。目前大多数基于深度学习神经网络分割乳腺病灶的算法是

采用先降采样提取特征再上采样恢复图像尺寸的“U”型网络实现的。通过在图像中检测异常或可疑区域,能提高病变区域的检出率,该应用在乳腺超声三维容积成像上的优势尤为明显。乳腺超声三维容积成像数据量大,阅片时间长,容易造成医生疲劳从而导致漏诊。有研究将辅助检测人工智能系统应用于三维容积乳腺超声影像,发现能显著改善三维容积成像系统的不足,明显缩短超声医师对患者的诊断时间。此外,辅助检测人工智能系统可以在二维乳腺超声扫查过程中快速提示可疑病灶区域并追踪定位,减少漏诊可能^[1]。

1.2 乳腺病变特征分析

根据美国放射学会(American College of Radiology, ACR)制定的《乳腺影像报告与数据系统》(Breast Imaging Reporting and Data System, BI-RADS)指南内容,临床主要关注二维图像中病灶的形态、边缘、内部回声、后方回声等特征。但由于BI-RADS指南中没有给出定量的指标描述各种特征,医生在进行分析时往往带有主观因素,导致不同医生之间的分析结果一致性较低。人工智能算法通过计算乳腺结节的形态学和纹理特征等,能够实现病灶特征的量化、客观化表达,包括肿块形态是否规则、边缘是否光滑、内部回声特性(低回声、无回声)、后方回声衰减情况、包膜完

收稿日期: 2022-08-09

作者简介: 许晓萍,从事医疗器械审评检查工作;曾祥卫,通信作者。

基金项目: 国家重点研发计划资助课题(课题名称: 人工智能医学信息系统软件审评指导体系构建,课题编号: 2019YFC0118801)。

整度、有无钙化、纵横比等，可以反映病变间更细微的差异，在鉴别乳腺良恶性中可表现出比医生更高的特异度^[2]。通过提取病变形态学和纹理学特征，人工智能可以帮助医生对图像进行更细致的分析，增强人工智能诊断结果的可解释性，规范超声诊断并提升观察者之间的一致性，辅助医生做出更准确的评价。

1.3 乳腺病变风险评估及诊断

在乳腺超声病灶风险评估领域，良恶性分类和BI-RADS分级是两大热门的研究方向。常见将乳腺肿块分为BI-RADS 2、3、4和5四级，和（或）分为良性和恶性两类。BI-RADS 4级乳腺肿块由于常介于良性与恶性之间，是临床诊断的难点。有研究以病理结果作为金标准，分别比较人工智能系统与2、4、6年资超声医生的诊断准确性，发现人工智能系统对乳腺良恶性肿瘤诊断的敏感度和特异度均高于不同年资医生。在恶性特征表现不明显的小肿块超声图像中，人工智能系统由于不受主观认知的影响表现尤为出色，因此对提高小乳腺癌的检出率有一定帮助。但对某些复诊病例，人工智能系统由于未能结合病史进行综合分析，仅依据当前的超声图像做出判断，容易出现偏差，这也是人工智能系统诊断的局限性^[3]。

2. 人工智能乳腺超声产品的上市情况及相关政策

人工智能乳腺超声产品用于医疗用途，在我国医疗器械监管体系中属于中高风险医疗器械，在美国、欧盟等地区也属于严格监管的产品，需要获得上市前批准才能投入临床使用。人工智能医疗器械用于临床仍处于发展初期，主要国家和地区的监管和评审部门都在加速医疗行业中人工智能技术的落地及合法化。下面分别介绍我国、美国和欧盟三个地区的人工智能医疗器械上市情况及相关政策。

2.1 我国人工智能乳腺超声的上市情况及相关政策

2019年，国家食品药品监督管理局（NMPA）为加强人工智能医疗器械注册申报工作的指导，制定了《深度学习辅助决策医疗器械软件审评要点》（2019年第7号）^[4]。为进一步加强人工智能医用软件类产品监督管理规范，NMPA又陆续发布了《人工智能医用软件产品分类界定指导原则》^[5]（2021年第47号）和《人工智能医疗器械注册审查指导原则》^[6]（2022年第8号），以下简称《导则》。根据产品分类，乳腺超声人工智能产品可分为辅助决策和流程优化两类，目前国内没有对应的审评指导原则或者行业标准，但有相关团体标准已立项编制中。

从上市产品来看，尽管有多款人工智能软件在近几年陆续获批，如视网膜病变眼底图像辅助诊断软件、肺炎CT影

像辅助分诊与评估软件等，但没有明确采用人工智能算法的超声乳腺人工智能软件或超声设备获批。

2.2 美国人工智能乳腺超声的上市情况及相关政策

美国食品药品监督管理局（FDA）已有多年各类型人工智能医疗器械审评审批的经验，包括已审批多个采用机器学习算法的超声乳腺人工智能产品，超声乳腺辅助诊断类如Koios DS for breast, AI辅助可佩戴3D乳腺超声产品，都可实现自动病灶识别和分析。但目前也没有发布针对乳腺超声人工智能产品的法规和协调标准。FDA于2019年4月发表了《基于人工智能/机器学习（AI/ML）的医疗仪器软件修改的拟议规管架构-讨论文件及反馈要求》^[7]描述了试想建立一个为人工智能和机器学习驱动的软件变更上市前审查方法，希望官方和制造商评估和监测人工智能/机器学习软件产品从上市前的开发到上市后的性能。2021年FDA发布了《基于人工智能的医疗器械软件学习行动计划》^[8]，并且与加拿大卫生部和英国药品和医疗产品监管署和药物管理局共同确定了以患者为中心的良好机器学习方法的十条指导原则，有助于促进使用人工智能和机器学习的医疗器械的安全、有效和高质量保证^[9]。

2.3 欧盟人工智能乳腺超声的上市情况及相关政策

随着2021年5月26日欧盟医疗器械法规（Medical Device Regulation, MDR）的实施，人工智能医疗器械软件，尤其是辅助决策类影像产品的上市前监管日趋严格^[10]。尽管MDR并没有对人工智能医疗器械提出具体的要求，国际标准化组织ISO已着手在制定多个人工智能产品开发过程质量控制的标准，譬如ISO/IEC AWI 5259系列，ISO/TR 24291:2021，ISO/IEC AWI 25059等。2021年初，欧盟委员会发布了人工智能法规（Artificial Intelligence Act, AIA）提议，预计在2024年~2025年实施^[11]。在该提议中人工智能医疗器械被分为II a、II b和III类风险等级的产品都将属于高风险人工智能产品，要求其上市符合性路径需要持有人工智能产品审核资格的第三方公告机构审批。在上述监管要求发布实施之前，超声乳腺人工智能医疗器械软件合格评定需要依赖于对已有法规要求（譬如MDCG 2020-1^[12]）和最新技术SOTA分析进行解读。早年不少在欧盟获批的医疗器械产品安全有效性的验证是无法通过FDA认证的，但从近两年同时获批欧盟和美国上市的Koios DS for Breast和Atusa可佩戴3D乳腺超声产品来看，对于超声乳腺人工智能医疗器械软件而言，如今欧盟审评的严格程度与美国FDA已不相上下。

3. 讨论

人工智能技术在乳腺超声领域中的研究如火如荼，但多

为单中心研究的局限性数据, 尚未见国内多中心大样本研究的相关报道。纵观全球, 研究成果转化成商业应用的产品也屈指可数, 人工智能乳腺超声在智能化算法落地方面仍存在着局限性。

3.1 超声影像对人工智能算法的性能要求高

不同于CT、MRI等影像可通过设置标准化的扫描参数规范成像质量, 超声成像有明显的操作者依赖性。同一病变, 不同医生的手法, 采集到不同的切面, 使用不同的参数调节图像等, 得到的图像可能会不一样, 这给人工智能算法提取特征带来了巨大的挑战。因此, 在扫描顺序、留图方式及使用超声设备方面, 需要建立统一的乳腺超声图像标准。同时, 二维乳腺超声要求医生边扫描边读图, 对算法的实时性有很高的要求。由于临床上使用的超声设备品牌众多, 图像质量差异较大, 对开发通用乳腺超声人工智能软件的厂家而言, 如何适应不同品牌的超声设备是个挑战, 这对人工智能算法的泛化能力和鲁棒性提出了更高的要求。

此外, 实际临床过程中, 医生往往结合多种模态的图像(如: 超声血流图、剪切波弹性成像等)以及患者信息(如: 年龄、家族病史等)综合分析病灶, 所以, 乳腺超声人工智能诊断不是简单的图像识别问题, 而是多信息、多模态图像信息处理和分析的问题, 如何对上述信息进行有效的融合进而得出更全面、准确的诊断, 是人工智能乳腺超声未来的发展方向。

3.2 数据库构建难度大

目前尚无乳腺超声影像诊断相关数据库及其统一使用标准或指南。由于亚洲女性的致密型腺体特点, 其他地区的数据不能简单地移植使用。而人工智能模型运行的关键步骤是收集数据构建数据库, 训练集、验证集、测试集等不同数据库之间还需要做到两两无交集, 因此构建数据库的工作量是巨大的。而且医学影像数据的收集和标注十分困难, 多数数据不易获得, 且数据标记不正确或达不到质量要求、数据

采集不标准和各种超声设备的差异都可产生有差别的数据集合。另外, 由于自然人群中患有严重乳腺病变(如乳腺癌)的人数远少于正常人群, 获取严重病变的数据又比较微病变或健康人群的数据更加困难, 这就容易产生数据类别不平衡等问题, 数据不平衡会导致人工智能模型的训练更加倾向于样本数量较多的类别, 降低实际准确率。因此, 建立中国女性乳腺多中心超声声像图数据库, 为训练人工智能模型提供病史详细、病种丰富、图像或视频质量规范的数据集, 对提升模型的鲁棒性和泛化能力至关重要。

3.3 隐私保护与责任归属

在智能化算法落地方面, 因医学影像数据牵扯到患者隐私, 人工智能系统在医学领域的开展需要相关部门加强监管。同时, 当超声医学智能化算法具备临床应用的可能性, 从研究成果转成商业应用时, 既需要考虑算法部署到超声设备上的实时性问题, 也需要严肃对待智能算法进入诊疗环节后的众多问题, 如人工智能辅助诊疗性能评估、法律法规的制定、医疗伦理、智能诊疗的责任归属、智能诊疗系统质量控制、诊疗系统的信息安全等。

4. 小结

乳腺超声与人工智能技术的结合, 有望解决我国乳腺超声行业长期存在的超声医生资源缺口大、诊断主观性强、工作量大、诊断效率低等实际问题。但要让人工智能算法商业化落地, 真正为临床所用, 这是一个多维度协作的复杂过程, 获得上市许可则是迈入商业化的敲门砖。因此, 结合我国的政策法规, 构建科学高效的上市前评价体系, 指导人工智能乳腺超声从技术研发走向现实应用, 对加速产业化落地, 助力产业高质量发展具有重要的现实意义。

参考文献

- [1] 杨意. 超声新技术在乳腺良恶性病变诊断中的应用进展[J]. 肿瘤预防与治疗, 2020, 33(11): 894-900.
- [2] 高思琦, 牛司华, 黄剑华, 等. 超声人工智能在乳腺良恶性疾病鉴别诊断中的应用[J]. 中国超声医学杂志, 2021, 37(7): 752-755.
- [3] 臧爱华, 姜明, 孟聪, 等. 人工智能系统评估BI-RADS 4类乳腺肿块的应用价值[J]. 中华医学超声杂志(电子版), 2021, 18(8): 795-799.
- [4] 国家食品药品监督管理局, 医疗器械技术审评中心. 关于发布深度学习辅助决策医疗器械软件审评要点的通告(2019年第7号)[EB/OL]. 2019-07-03. <https://www.cmde.org.cn/xwdt/shpgzg/gztg/20190703141714991.html>.
- [5] 国家食品药品监督管理局, 国家药监局关于发布人工智能医用软件产品分类界定指导原则的通告(2021年第47号)[EB/OL]. 2021-07-01. <https://www.nmpa.gov.cn/xxgk/ggtg/qtggtg/20210708111147171.html>.
- [6] 国家食品药品监督管理局, 医疗器械技术审评中心, 国家药监局器审中心关于发布人工智能医疗器械注册审查指导原则的通告(2022年第8号)[EB/OL]. 2022-03-07. <https://www.cmde.org.cn/xwdt/shpgzg/gztg/20220309090800158.html>.
- [7] U.S. 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[EB/OL]. 2019.
- [8] U.S. food and drug administration. Artificial Intelligence/Machine Learning(AI/ML)-Based Software as a Medical Device(SaMD) Action Plan[EB/OL]. 2021.
- [9] U.S. food and drug administration. Good Machine Learning Practice for Medical Device Development: Guiding Principles[EB/OL]. 2021.
- [10] Regulation(Eu) 2017/745 of the European Parliament and of the Council Of 5 April 2017[EB/OL]. 2017-05.
- [11] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence(Artificial Intelligence Act) and Amending Certain Union Legislative Acts[EB/OL]. 2021-04.
- [12] MDCG 2020-01. Guidance on Clinical Evaluation(MDR)/Performance Evaluation(IVDR) of Medical Device Software[EB/OL]. 2020-03.