

人工智能技术在医学影像诊断学 肺结节教学中的初步应用

陈建华¹, 孙斌¹, 葛慧婷¹, 林立², 李坚³

- (1. 福建医科大学附属协和医院 放射科, 福建 福州 350001;
2. 福建医科大学 临床医学部, 福建 福州 350122;
3. 福建医科大学附属第一医院 医学影像科, 福建 福州 350004)

摘要:通过比较人工智能(AI)技术辅助教学组和传统 PACS 教学组学生实习后的肺结节影像诊断理论成绩、报告书写时间、结节发现率和诊断报告评分,评价 AI 技术在辅助影像本科实习生肺结节实习教学的效果。结果显示,AI 辅助教学组的理论成绩、报告书写速度、结节发现率和报告评分都优于传统 PACS 教学组。因此,基于 AI 技术的肺结节教学有助于巩固学生的理论知识,快速掌握肺结节病变报告的规范书写,使肺结节影像教学效果得到全面提高,值得推广应用。

关键词:肺结节;人工智能(AI);PACS;医学教育;医学影像诊断学;实习

中图分类号:R-4;G642

文献标志码:A

文章编号:1009-4784(2023)04-0063-05

人工智能(artificial intelligence, AI)指研发用于模拟、延伸和扩展人类智能的理论、方法、技术及应用系统的一门信息科学,在计算机学、心理学、哲学和语言学等领域都有涉及,是一门多学科、多领域交叉的前沿学科^[1]。AI 已被医学教育专家应用于辅助医学教学,为不同专业、不同层次的住院医师提供个性化、智能化的人机交互教学模式,以期能够在短时间内培训学员有效掌握医学技能^[1]。

当前,医学影像学教育模式仍依赖于以教师授课、学生听讲为主的传统教学模式。传统教学模式的学习效果受病例数和病例的多样性的影响,且随着影像科医师工作量的快速增长、受训人员的日益增多,传统模式的教学已日渐无法满足医学影像人才培养的需求^[2]。医学影像数据的数字化给 AI 辅助影像实践提供了契机,医学影像与 AI 的结合已被认为是最具发展前景的领域之一^[3]。AI 在医学影像方面的应用主要包括病变检测、图像分割、图像质量评估、自然语言处理、协议改进等,人工智能在医学影像教育中的应用研究还处于萌芽阶段^[4-6]。

近年来,AI 在医学影像辅助诊断实践中的应用不断成熟,给 AI 辅助影像教学的发展带来了更多的可能。本研究通过开展对比研究,探讨肺结节 AI 系统辅助教学在学生肺结节影像诊断教学中的应用效果。

一、研究方法

(一) 研究对象及分组

选取福建医科大学 2016 级和 2017 级于福建医科大学附属协和医院影像科实习的本科生为研究对象。入组标准:(1)大学本科 5 年全日制医学影像学专业学生;(2)大学本科 5 年级初入影像科实习者。排除标准:(1)非首次接触 CT 阅片实践者;(2)中途退出实习者。

按实习轮转批次,将 2016 级本科实习生纳入传统影像归档和通信系统(picture archiving and communication system, PACS)教学组(以下简称传统教学组),2017 级本科实习生纳入 AI 辅助教学组。

收稿日期:2022-09-09

资助项目:2020 年福建省本科高校教育教学改革研究一般项目(FBJG20200028)

作者简介:陈建华,男,副主任医师,医学硕士。研究方向:CT 和 MR 影像诊断。

(二) 教学方法

所有学生均由同一个教师带教,教学时间均为 2 个课时(90 min);随后开始进入为期 1 周的实践学习,所有学生每日实践学习时间均为 8 个学时,并于实习第 2 天开始每日书写 3 份胸部 CT 报告。带教教师回答学生学习过程中存在的问题,并对所有书写报告进行审核并指出不足之处,具体如下。

1. 传统教学组。带教教师通过调阅 PACS 系统中典型肺结节病变的 CT 图像,为学生讲解常见肺结节病变的影像表现(位置、大小、CT 值、边界情况、病变性质等)。随后学生自主于 PACS 中结合影像图像和既往的诊断报告学习 1 周。

2. AI 教学组。带教教师结合 PACS 系统和 AI 肺结节辅助诊断系统(推想科技)进行教学。带教教师选择同一批典型病例,调阅图像结合 AI 辅助系统给学生讲解常见肺结节病变的影像表现。随后学生自主浏览图像结合 AI 诊断系统和既往报告学习 1 周。其中 AI 系统可提供的功能有:快速自动识别肺部结节,提供肺部结节位置、大小(包括体积)、CT 值、边界情况、恶性概率等结节特征。

(三) 教学效果评价

1. 理论考试。所有学生在实习前均进行肺结节影像诊断理论摸底考试,并在 1 周的实习后进行出科理论考试,考题从题库中随机抽取,分别有 60 道选择题和 6 道简答题,总分 100 分。

2. 阅片技能考试。考试方法:1 周实习结束后,在医院题库中随机选择 20 份胸部 CT 影像资料对学生进行阅片技能考核,每个病例的肺结节均 ≥ 5 个。传统教学组和 AI 教学组学生均不依靠任何工具独立书写 20 份肺结节诊断报告,诊断报告的内容包括所有结节的层面(大小、形态、密度、边缘等)征象的描述,并在诊断中指出肺结节的危险程度(以高危、中危、低危表示)。由 2 名有 5 年以上胸部 CT 影像诊断工作经验的主治医师记录学生的书写报告时间;并于考试结束后对学生的结节检出率、诊断报告质量进行评分,取二者的平均值作为最终结果。计时方法:学生于 PACS 中完全打开患者的图像开始到提交所有 20 份报告结束计时,时间单位为 s,计算学生出具每份报告的平均用时。结节检出率:20 份报告中学生检出结节的总数目占实际结节总数目的比例。诊断报告质量评估方法:CT 诊断报告采用 5 分制进行评分,计算 20 份报告的总得分作为最终的诊断报告质量评分依据。5 分制的评分

细则根据“胸部 CT 诊断报告评分标准”改良如下^[7]:(1)5 分为肺结节无漏诊且定位准确,结节特征(大小、形态、密度、边缘等)描述清晰、用语规范,结节的危险程度诊断准确。(2)4 分为肺结节无漏诊且定位准确,结节的危险程度诊断准确,但结节特征描述欠完整和/或用语欠规范;或有 1 个低危肺结节定位错误或漏诊。(3)3 分为肺结节无漏诊且定位准确、描述清晰,有 1 个中危结节诊断错误;或有 2 个低危肺结节定位错误、诊断错误或漏诊。(4)2 分为介于 3 分和 1 分之间的其他情况。(5)1 分为 3 个以上肺内结节诊断错误;或遗漏高危结节。

(四) 数据录入与统计

数据录入采用双人录入法,在 Excel 输入数据形成数据库,经过对比核对无误后,使用 SPSS 24.0 软件进行数据分析。计量资料:本研究中所有计量资料均符合正态分布,以 $\bar{X} \pm S$ 检验进行统计学分析;来源独立的样本采用独立样本 t 检验进行统计学分析。计数资料以率(%)表示,采用 Fisher 检验进行统计学分析;以 $P < 0.05$ 为 2 组数据间的差异有统计学意义。本研究的样本量较小,故拟通过 PASS 11 软件计算检验效能 $1 - \beta$ 来评估上述统计结果的准确率,一般认为 $1 - \beta$ 达到 0.9 以上可视为结果可靠^[8]。

二、研究结果

2016 级和 2017 级福建医科大学 5 年制影像医学专业的学生各有 60 名,2 个年级在医院实习的学生各有 18 名且全部入组。所有学生在实习前均未使用过 AI 辅助阅片。传统教学组与 AI 教学组学生的一般资料差异均无统计学意义($P > 0.05$,表 1)。本文检验效能结果均为可靠。

(一) 2 组学生肺结节理论考试成绩对比分析

传统教学组和 AI 教学组学生实习前的肺结节影像诊断理论成绩(基线成绩)分别为 (79.33 ± 5.39) 分和 (79.11 ± 4.94) 分,差异无统计学意义($P = 0.898$,表 2)。传统教学组和 AI 教学组学生实习后的肺结节影像诊断理论考试成绩分别为 (87.39 ± 3.43) 分和 (94.39 ± 2.43) 分,成绩均较实习前有所提高,差异有统计学意义($P < 0.001$);实习后 AI 教学组学生的肺结节理论考试成绩优于传统教学组,差异有统计学意义($P < 0.001$,表 2)。

表 1 研究的 2 组学生入科时的基本资料

项 目	传统教学组	AI 教学组	<i>t</i>	<i>P</i>
性别 <i>n</i> (%)			—	1.0
男	7 (38.9)	8 (44.4)		
女	11 (61.1)	10 (55.6)		
年龄/岁	23.17±0.38	23.22±0.43	0.41	0.684
大学四年平均绩点	2.94±0.30	2.79±0.49	1.05	0.303
影像诊断学成绩/分	74.88±5.92	76.39±5.83	-0.75	0.459

注:*n*=18;表中年龄为 $\bar{X}\pm S$;AI:人工智能。

表 2 研究的 2 组学生实习前、后肺结节理论知识考试成绩比较

项 目	传统教学组	AI 教学组	<i>t</i>	<i>P</i>
实习前理论考试成绩/分	79.33±5.39	79.11±4.94	-0.13	0.898
实习后理论考试成绩/分	87.39±3.43	94.39±2.43	-7.60	<0.001
<i>t</i>	-6.51	-12.95	—	—
<i>P</i>	<0.001	<0.001	—	—

注:*n*=18;表中数据为 $\bar{X}\pm S$;AI:人工智能。

(二)阅片技能考试成绩

传统教学组的报告书写时间长于 AI 教学组,分别为(496.61±72.29)s 和 (275.78±13.60)s, AI 教学组肺结节诊断报告评分高于传统教学组,分

别为(78.28±4.63)分和(94.06±1.26)分,差异有统计学意义(*P* < 0.001)。AI 教学组学生的结节发现率高于传统教学组,分别为(83±3.33)%和(97.06±1.47)%,差异有统计学意义(*P* < 0.001,表 3)。

表 3 研究的 2 组学生阅片考试成绩比较

项 目	传统教学组	AI 教学组	<i>t</i>	<i>P</i>
报告书写时间/s	496.61±72.29	275.78±13.60	12.74	<0.001
结节发现率/%	83.00±3.33	97.06±1.47	-16.39	<0.001
诊断报告评分/分	78.28±4.63	94.06±1.26	-13.97	<0.001

注:*n*=18;表中年龄数据为 $\bar{X}\pm S$;AI:人工智能。

三、分析与讨论

(一)AI 辅助教学对学生学习效果的影响

AI 辅助肺结节诊断是医工结合的产物,受到医师和计算机工程师的广泛关注。AI 在辅助影像临床实践诊疗中的研究,如辅助诊断肺结节的病变检出、病变性质诊断准确率以及诊断效率等方面的研究层出不穷, AI 也被证实能切实提高影像诊断的准确性和诊断效率^[9-11]。在教学方面, AI 技术可为影像教学的个性化、交互性、自主性提供技术支撑,有望让学生的各方面能力得到提升^[12-14]。本研究对比了传统教学组和 AI 教学组在肺结节方面的教学效果,结果表明 AI 教学组的学生在实习后无论理论知识还是阅片技能都优于传统教学组。

本研究中 2 组实习生在实习前的理论成绩相

仿,表明学生在入科前的知识储备是相似的。但是在用不同的方法进行阅片教学后,在阅片技能方面, AI 教学组无论在结节的检出率、影像诊断报告评分,还是在报告书写时间方面都优于传统教学组。可能的原因是在传统教学过程中,学生只能通过自己阅片,然后与既往的报告对比或询问教师来证实自己的阅片结果是否正确。而事实上,对于多发的结节而言,诊断报告中无法详尽描述,更多细节只能通过询问带教教师来获取每个结节的准确描述,效率相对较低。而 AI 教学组中,由于 AI 系统可以快速地自动识别结节,并实时描述每个结节的影像特征,出具初步诊断报告, AI 教学组的学生可以借助 AI 系统进行自主学习,在学习过程中实时对比自身和 AI 系统的阅片结果。相对于传统教学中完全依赖于带教教师的解答而言, AI 辅助教学让学生在有限的时间内学习到更多病例,同时还能

在一定程度上减少因为带教教师主观因素造成的影响。因此, AI 辅助教学在医学影像实习教学中有其独特的优势, 有助于实习生快速掌握肺结节阅片技能。

本研究还发现, AI 辅助教学不仅在帮助学生掌握阅片技能方面有优势, 在提高学生的理论知识水平方面也同样优于传统教学法。首先, 在阅片实习后, 所有学生的理论成绩均较之前有所提高, 表明了理论结合实际有利于巩固和提高理论知识, 这与刘星君等^[15]的观点一致。另外, 虽然实习后 2 组学生成绩均有所提高, 但是提高的程度不同, AI 教学组学生的理论成绩明显高于传统教学组。可能的原因是, AI 的实时反馈性充分调动学生积极性并提高学习效率, 让学生在有限的时间内汲取更多的知识, 将理论和实践不断结合, 同时提高理论和实践的成绩^[16-17]。

需要指出的是, 本研究样本量较小, AI 辅助教学和传统教学组不是将所有学生进行随机分组, 而是分别取自不同年级的学生在不同时间进行。(1) 针对随机问题, 将 2 个年级或多个年级的学生合并, 随机分成 AI 辅助教学和传统教学组, 这在理论上可行, 但是实际可操作性不强, 因为在 AI 软件已经安装的情况下, 学生对 AI 软件的使用是不可控的, 可能会导致结果不可靠。此外, 将学生随机分成 AI 辅助教学和传统教学组, 可能也会涉及教育的公平性, 不符合伦理, 故笔者选择了在 AI 辅助软件引进前的 2016 级和引进后的 2017 级的学生进行研究。由于每年实习的学生由福建医科大学随机分配, 故在一定程度上符合随机原则。(2) 针对可比性问题, 虽然传统教学组和 AI 辅助教学组分别为 2016 级和 2017 级的学生, 并非同一年入学, 但所有学生都是本科 5 年制, 经过 4 年理论学习, 刚进入影像科实习的学生前期接受的教育也基本相当, 具有一定的可比性。(3) 针对样本量问题, 受上述因素影响, 选择 2016 级和 2017 级学生作为研究对象是本研究的最优选, 而每年实习学生人数基本固定, 所以在本研究中样本量基本属于不可控因素。针对样本量小可能存在结果不可靠的问题, 笔者对研究结论进行了准确性分析, 结果显示, 当前的样本量可 100% 支持研究结论。后续可通过多中心的研究来扩充样本量, 从而进一步验证本研究的结论。

(二) AI 辅助教学存在的不足

AI 技术有望推进影像的高质量、高效率教育, 但 AI 技术辅助影像教育的研究仍处于初级阶段, 无论设备还是方法都还需进一步完善。目前存在最突出的一个问题是 AI 系统虽然能快速自动识别病变, 对病变进行描述和出具初步诊断报告, 但仍存在一定的假阳性和假阴性。荷兰学者^[18]对肺结节检测系统进行了临床评价, 研究表明 AI 系统肺结节检测敏感度为 88%, 每次扫描中平均检出 1.04 个假阳性结节, 人工智能还不足以取代影像科医师。依靠 AI 辅助诊断系统学习, 虽然能在有限的时间内学到更多病例, 但如若不注意甄别假阳性、假阴性病例, 迷信 AI 辅助诊断结果, 可能会吸收错误的知识, 从而影响学习效果。

四、提高 AI 辅助影像教学效果的对策建议

(一) 建立医学影像教学与 AI 教学模式融合体系

在信息爆炸的大数据时代, 筛选出高质量的数据是在有限的精力和时间里尽可能汲取高质量知识的基础, 在医学影像教育中也是同理。通过高质量的数据筛选, 建立传统医学影像+AI 技术的融合课程体系, 可形成更科学的教学和学习模式。对于教师而言, 使用 AI 技术弥补传统教学的不足, 可使教师从传统的“照本宣科”过渡到与 AI 技术相结合, 教师的带教过程更有层次感, 从单一角色向复合型角色成长, 以适应日趋发展的医学教学, 实现自我提升。对于学生而言, 高质量、成熟的医学影像教学与 AI 教学模式融合体系的应用, 让学生可选用适合自己的个体化的学习资料, 将理论与 AI 辅助诊断软件相结合, 使影像知识的学习不再枯燥, 有利于实现知识的高效获取和巩固。

(二) 建立医学影像 AI 教学的实践机制

传统影像医学与现代化 AI 技术相结合是未来医学的发展方向, 注重 AI 技术的革新与应用, 针对影像诊断报告制定有关规范, 包括病灶的规范描述及准确测量、患者随诊资料的规范对比、医学诊断术语的规范使用等将进一步促进影像医学教育的发展。此外, 还应针对如何科学利用 AI 系统进行专门教导, 包括但不限于: (1) 教导学生在使用 AI

辅助系统学习过程中,应善于总结问题,结合自身的理论知识,对AI的诊断结果进行甄别,对于存疑的病例应该勤学多问,结合教师的答疑,更高效、快速地提升自身的理论和实践技能;(2)教导学生将AI辅助系统作为提高学习能力的工具,可借鉴和利用,但不可过度依赖,应重视培养自身独立阅片的能力。通过建立和完善医学影像AI教学的实践机制,让影像实践技能与课本知识教育有机结合,提高新时代影像医学生的职业素质能力。

参考文献:

- [1]熊瑶,陈敏.人工智能在医疗领域应用现状探讨[J].医学信息学杂志,2018,39(4):24-28.
- [2]张冉冉,蒲利红,张文,等.人工智能在医学影像教育中的应用现状和前景[J].现代预防医学,2019,46(24):4527-2609.
- [3]CARLOS R C, KAHN C E, HALABI S. Data science: big data, machine learning, and artificial intelligence[J]. J Am Coll Radiol, 2018, 15(3 Pt B):497-498.
- [4]KAHN C E JR. Artificial intelligence in radiology: decision support systems[J]. Radiographics, 1994, 14(4):849-861.
- [5]LAKHANI P, PRATER A B, HUTSON R K, et al. Machine learning in radiology: applications beyond image interpretation[J]. J Am Coll Radiol, 2018, 15(2):350-359.
- [6]KIM D W, JANG H Y, KIM K W, et al. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers[J]. Korean J Radiol, 2019, 20(3):405-410.
- [7]严陈晨,麦筱莉,辛小燕,等.人工智能技术在医学影像实践教学中的探索[J].江苏卫生事业管理,2021,32(11):1534-1538.
- [8]MACHIN D, CAMPBELL M, FAYERS P, et al. Sample size tables for clinical studies[M]. 2nd ed. Malden: Blackwell Science, 1997.
- [9]ATHER S, KADIR T, GLEESON F. Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications[J]. Clin Radiol, 2020, 75(1):13-19.
- [10]CHEN L, GU D, CHEN Y, et al. An artificial-intelligence lung imaging analysis system (ALIAS) for population-based nodule computing in CT scans[J]. Comput Med Imaging Graph, 2021, 89:101899.
- [11]CHASSAGNON G, VAKALOPOULOU M, PARAGIOS N, et al. Artificial intelligence applications for thoracic imaging[J]. Eur J Radiol, 2020, 123:108774.
- [12]何炼图,汤庆,汤佳馨,等.人工智能+医学影像教育新模式的作用与构想[J].中国继续医学教育,2018,10(20):31-33.
- [13]王霄英.人工智能在医学影像中的进展:2017年RSNA参会感受[J].放射学实践,2018,33(2):101-103.
- [14]RABBANI M, KANEVSKY J, KAFI K, et al. Role of artificial intelligence in the care of patients with nonsmall cell lung cancer[J]. Eur J Clin Invest, 2018, 48(4):e12901.
- [15]刘星君,刘艳,郭辉,等.本科生诊断学实践技能考核结果及其与理论成绩相关性分析[J].现代医药卫生,2017,33(19):3036-3038.
- [16]文亮,赵迅冉,何利平.调查研究医学本科生对AI的态度[J].继续医学教育,2019,33(7):12-13.
- [17]祝叶华.“弱人工智能+”时代来了[J].科技导报,2016,34(7):67-69.
- [18]JARNALO C O M, LINSSEN P V M, BLAZÍS S P, et al. Clinical evaluation of a deep-learning-based computer-aided detection system for the detection of pulmonary nodules in a large teaching hospital[J]. Clin Radiol, 2021, 76(11):838-845.

(编辑:陈越,陈典)

欢迎投稿《福建医科大学学报(社会科学版)》