

AI 与不同年资放射医师对新冠肺炎 CT 诊断效能比较

应伟峰, 陈穹, 王钢, 路明, 翁慧超, 张莹

【摘要】 目的:比较和分析人工智能(AI)技术、低年资放射医师、高年资放射医师、AI联合低年资放射医师对新冠肺炎CT诊断效能,探讨AI对新冠肺炎CT辅助诊断价值。**方法:**分析经AI新冠诊断软件诊断新冠概率 $>0\%$ 的53例患者图像,统计计算A组(低年资医师,参与放射诊断工作 <10 年)、B组(高年资医师,参与放射诊断工作 ≥ 10 年)、C组(AI)、D组(低年资医师联合AI)初次CT诊断新冠肺炎信心评分受试者特征曲线(ROC),Z检验比较各组诊断效能;并以AI诊断新冠肺炎最大约登指数相应截断值为标准,比较各组敏感度、特异度。**结果:**各组诊断效能曲线下面积比较(AUC)。A组 $<$ B、D组($Z=2.406, 2.013$),C组 $<$ D组($Z=2.031$),B组 $>$ C组($Z=2.113$),差异均具有统计学意义($P<0.05$);余组间差异均无统计学意义($P>0.05$)。以AI诊断新冠肺炎最大约登指数相应截断值 >3 分为标准,各组对新冠肺炎CT诊断效能比较。敏感性:A组 $<$ B、D组($\chi^2=3.854, 13.781$),C组 $<$ D组($\chi^2=7.726$),差异均具有统计学意义($P<0.05$);余组间差异均无统计学意义($P>0.05$)。特异性:A组 $<$ B、D组($\chi^2=6.876, 6.876$),差异均具有统计学意义($P<0.05$);余组间差异均无统计学意义($P>0.05$)。**结论:**基于深度学习AI技术在新冠肺炎CT诊断中具有较高临床应用价值,AI可辅助低年资医师提高对新冠肺炎CT诊断效能。

【关键词】 人工智能;诊断;肺炎病毒感染;新型冠状病毒;体层摄影术,X线计算机

【中图分类号】 R-05;R44;R512.99;R563.1;R814.4 **【文献标识码】** A

【文章编号】 1000-0313(2020)10-1231-05

DOI:10.13609/j.cnki.1000-0313.2020.10.003

开放科学(资源服务)标识码(OSID):



Comparison of the diagnostic performance in COVID-19 pneumonia between AI and radiologists based on CT images YING Wei-feng, CHEN Qiong, WANG Gang, et al. Department of Radiology, Dahua Hospital, Xuhui District, Shanghai 200237, China

【Abstract】 Objective: To investigate the diagnostic value of artificial intelligence (AI) in coronavirus disease 2019 (COVID-19) pneumonia by comparing and analyzing the diagnostic performance of AI technology, junior radiologists, senior radiologists, and AI combined with junior radiologists. **Methods:** The images of 53 patients with a probability $>0\%$ for COVID-19 pneumonia diagnosed by AI-aided software were retrospectively analyzed. The receiver operating characteristic (ROC) curves for diagnosing COVID-19 pneumonia were analyzed in group A (junior radiologists with less than 10 years of experience), group B (senior radiologists with more than 10 years of experience), group C (AI-aided system) and group D (junior radiologists combined with AI) based on the initial CT images. Z test was used to compare the diagnostic performance between different groups. Referring to the cutoff value corresponding to the maximum Youden index for AI aided model, the sensitivity and specificity of each group were compared. **Results:** Regarding the area under the curve (AUC), group A $<$ group B and D ($Z=2.406, 2.013$), group C $<$ group D ($Z=2.031$), group B $>$ group C ($Z=2.113$), these differences were all statistically significant ($P<0.05$), and there were no statistically significant differences between the other groups ($P>0.05$). When using the cut-off value corresponding to the maximum Youden index >3 as a standard, for sensitivity: group A $<$ group B and D ($\chi^2=3.854, 13.781$), group C $<$ group D ($\chi^2=7.726$), for specificity: group A $<$ group B and D ($\chi^2=6.876, 6.876$), these differences had statistical significant ($P>0.05$), and other groups have no significant differences ($P>$

作者单位:200237 上海市徐汇区大华医院放射科

作者简介:应伟峰(1978-),男,上海人,主治医师,主要从事CT、MRI诊断工作。

通讯作者:陈穹, E-mail: cq1444@sina.com

0.05). **Conclusion:** AI technology based on deep learning has a significant clinical application value in the diagnosis of COVID-19 pneumonia using CT, and AI could help junior radiologists to improve the diagnostic performance of COVID-19 pneumonia.

【Key words】 Artificial intelligence; Diagnosis; Pneumovirus infections; Novel coronavirus; Tomography, X-ray computed

材料与方法

1. 临床资料

搜集 2020 年 1—2 月来自 4 中心由人工智能 (artificial intelligence, AI) 判定新冠肺炎概率 > 0% 53 例 CT 扫描患者, 男 35 例、女 18 例, 年龄 29~82 岁, 平均年龄 (51.94 ± 11.08) 岁。53 例中 48 例发热, 45 例咳嗽, 19 例不同程度胸闷。53 例中 9 例有近期疫区旅居史, 8 例有近期外地旅居史, 26 例与新冠患者密切接触史, 10 例无明确与新冠患者密切接触史。

2. 检查方法

联影 (uCT760) 128 层螺旋 CT 机, 扫描参数: 探测器 128 × 0.625 mm, 管电压 120 kV, 管电流为自动毫安技术, 重建层厚 1.5 mm, 重建层间距 1.5 mm, 滤波函数 B-SOFT-B。

3. 图像分析

A 组: 低年资医师组, 由 2 名从事放射诊断 < 10 年医师组成; B 组: 高年资医师组, 由 2 名从事放射诊断工作 ≥ 10 年医师组成; C 组: AI 软件为卫宁健康基于深度学习新冠肺炎 CT 诊断软件; D 组: 低年资医师联合 AI 阅片组。各医师组内意见不一致需协调后判定结果。

医师诊断信心评分标准: 1 分, 几乎不可能是新冠肺炎; 2 分, 较小可能是新冠肺炎; 3 分, 可能是新冠肺炎; 4 分, 较大可能是新冠肺炎; 5 分, 极可能是新冠肺炎。AI 软件评分标准: 新冠概率 < 20% 为 1 分, 20%~39% 为 2 分, 40%~59% 为 3 分, 60%~79% 为 4 分, ≥ 80% 为 5 分。A、B、C、D 组对 CT 影像资料进行分析, 比较各组对新冠肺炎诊断效能。

4. 统计学方法

采用 Medcalc 软件计算各组诊断受试者工作特征曲线 (ROC) 并行统计学比较 (Z 检验); 采用 SPSS 23 软件运用 X 检验对各组敏感度、特异度进行比较。

结 果

53 例中经逆转录聚合酶链反应 (reverse transcription-polymerase chain reaction, RT-PCR) 检测阳性 36 例 (金标准), ROC 曲线下面积 (area under curve, AUC) 比较 (图 1)。

A 组 (0.717) < B 组 (0.902) ($Z = 2.406, P < 0.05$), A 组 < D 组 (0.853) ($Z = 2.013, P < 0.05$), B 组 > C 组 ($Z = 2.113, P < 0.05$), C 组 < D 组 ($Z = 2.031, P < 0.05$), 差异均具有统计学意义; A 组 < C 组 (0.753) ($Z = 0.363, P > 0.05$), B 组 > D 组 ($Z = 0.744, P > 0.05$), 差异均无统计学意义。* C 组 (AI) ROC 最大约登指数相应截断值为 > 3 分。

以 AI 诊断新冠肺炎最大约登指数相应的截断值 > 3 分为标准, 各组对新冠肺炎 CT 诊断效能比较 (表 1)。

敏感性: A 组 < B 组 ($\chi^2 = 3.854, P < 0.05$), A 组 < C 组 ($\chi^2 = 1.049, P > 0.05$), A 组 < D 组 ($\chi^2 = 13.781, P < 0.05$), B 组 > C 组 ($\chi^2 = 0.907, P > 0.05$), B 组 < D 组 ($\chi^2 = 3.556, P > 0.05$), C 组 < D 组 ($\chi^2 = 7.726, P < 0.05$)。

特异性: A 组 < B 组 ($\chi^2 = 6.876, P < 0.05$), A 组 < C 组 ($\chi^2 = 2.976, P > 0.05$), A 组 < D 组 ($\chi^2 = 6.876, P < 0.05$), B 组 > C 组 ($\chi^2 = 0.827, P > 0.05$), B 组 = D 组 ($\chi^2 = 0.000, P > 0.05$), C 组 < D 组 ($\chi^2 = 0.827, P > 0.05$)。

AI (C 组) 在具有典型 CT 征象新冠肺炎诊断中对低年资医师 (A 组) 具有辅助诊断作用, 且 AI 对病灶体积精确测量功能对新冠肺炎转归判断具有辅助价值 (图 2), 高年资医师 (B 组) 结合病史、实验室指标、自身临床经验等对不典型新冠肺炎诊断、鉴别诊断能力部分优于其他组 (图 3)。

表 1 各组对新冠肺炎 CT 诊断效能比较

指标	A 组	B 组	C 组	D 组
敏感性 (%)	75	86.11	80.56	94.44
特异性 (%)	52.94	70.59	64.71	70.59

讨 论

1. 新冠肺炎 CT 诊断

胸部 CT 对新冠肺炎漏诊率较低^[1], 可作为新冠肺炎快速诊断标准方法^[2], 并被推荐为临床诊断主要依据^[3], 典型新冠肺炎主要表现为双侧及周边毛玻璃样病变 (ground-glass opacity, GGO)、伴有或不伴有血管增厚^[4], 多灶性病变大部分分布于中下肺及后肺区^[5], 可有“疯狂铺路石征”、“反晕征”^[6]等。而肺水肿、间质性肺炎、过敏性肺炎等也表现为肺部弥漫性磨

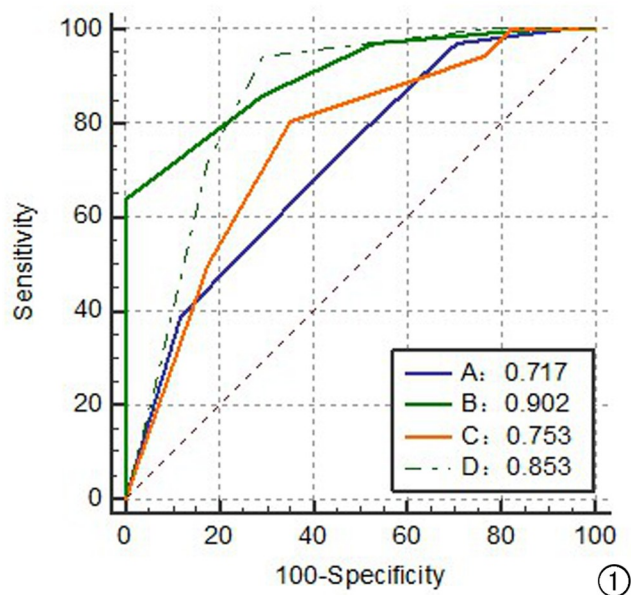


图 1 各组 ROC 曲线下面积(AUC)比较

玻璃密度影(diffuse ground-glass opacity, DGGO)为主^[7], H1N1 流感、H7N9 流感肺炎等病毒性肺炎 CT 表现也存在毛玻璃密度影、小叶间隔增厚^[8,9]等与新冠肺炎相似的征象,甚至临床表现都有一定类似;不典型新冠肺炎 CT 表现各异,甚至存在核酸检测阳性而

人工 CT 诊断“阴性”^[10]现象,所以对于新冠肺炎诊断存在着放射科医师所能运用影像特征较多却又鉴别困难现象,且在新冠疫情中对大量患者筛查工作难度、繁重度较高。AI 技术已研究和广泛应用于 CT^[11]、MRI^[12]、乳腺 X 线诊断^[13],胸部 CT 中已应用于肺部小结节筛选^[14]、小结节良恶性鉴别^[15]等。

2. AI 对新冠肺炎 CT 辅助诊断应用特点

本次研究 AI(C 组)诊断新冠肺炎 AUC 达到了中等度的 0.753,敏感度、特异度分别达 80.56%、64.71%,显示其有一定诊断应用价值,这 3 项指标虽高于低年资医师(A 组),但结果不具统计学意义,显示 C、A 组诊断效能相差不大。蔡雅倩等^[15]研究发现住院医师联合 AI 后降低了对肺毛玻璃结节筛查及定性诊断中误诊率、漏诊率,敏感度也从 65.20% 上升到 96.02%。本组联合低年资医师与 AI 的 D 组对新冠肺炎诊断敏感度、特异度达到 94.44%、70.59%,明显高于 A 组 75%、52.94%($P < 0.05$),接近蔡雅倩等^[15]研究数据,且本次 D 组 AUC 为 0.853,也明显高于 A 组 0.717($P < 0.05$)。充分说明联合 AI 后对低年资医师诊断新冠肺炎具有显著提升作用。我们分析主要原因包括①AI 已“深度学习”了数千例新冠肺炎病例,积累“数据”量远远超过了低年资医师;②刘小玉等^[16]发现

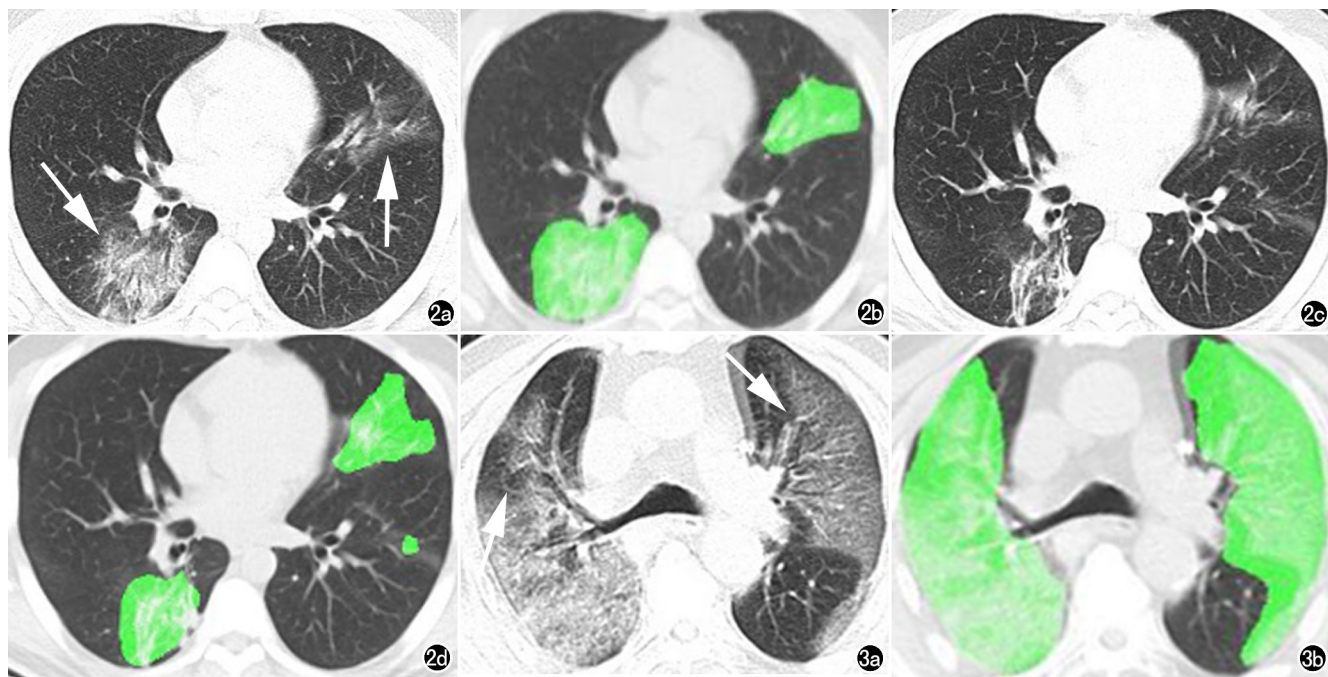


图 2 男,52 岁,发热、咳嗽就诊,无明确新冠肺炎患者接触史。a)两肺胸膜下为主毛玻璃密度影(箭),A 组新冠肺炎诊断信心评分 2 分,B 组 5 分;b)C 组评 5 分,病灶体积占肺 7.46%。RT-PCR 检测阳性,新冠肺炎。D 组评 3.5 分;c)1 周后病灶部分实变;d)AI 示病灶体积占肺体积 5.71%,提示部分吸收。图 3 男,82 岁,发热、咳嗽、胸闷,明确疫区人员接触史,有慢性间质性肺炎既往史,白细胞值正常、淋巴细胞值处于下限、血清淀粉样蛋白酶 $\uparrow$ 。a)两肺胸膜下为主磨玻璃密度影伴间质增厚(箭);b)C 组根据影像特征示新冠诊断评分 3 分,病灶体积占肺 24.95%。A 组评 3 分,B 组结合实验室指标及临床病史对其评 5 分。最终 RT-PCR 阳性,新冠肺炎。

首次 CT 新冠肺炎筛查敏感性可以高于核酸检测,本组 53 例筛选出具有新冠肺炎概率病例是在数千名常规胸部 CT 检查中筛选出来的, AI 没有低年资医师视觉、脑力疲劳等问题^[15]; ③低年资医师对新冠肺炎诊断及鉴别诊断经验积累不足。所以低年资医师联合 AI 诊断可弥补自身所没有大量数据经验。

本次研究中由于部分病例肺内病灶分布于非胸膜下为主,与典型新冠肺炎征象不甚相同^[4],且部分病患具有一些肺外新冠征象如纵膈气肿、肝实质密度减低等,这与丁义等^[17]研究结果相似,而 AI 不能发现肺外征象;侯可可等^[18]研究发现中性粒细胞/淋巴细胞比值、T 淋巴细胞计数等联合 CT 影像特征对早期识别新冠肺炎具有重要意义,而本组各年资医师较 AI 拥有结合实验室指标、影像特征综合诊断优势;且此次样本中部分病例伴有基础性肺部疾病如间质性肺炎、肺水肿等具有磨玻璃病灶的征象^[7]掩盖了新冠肺炎病灶征象;这些都突出了 AI 诊断机械性缺陷,而其联合低年资医师共同阅片可以很好解决这类缺陷。

另外, D 组对新冠肺炎诊断 AUC 低于 B 组,而敏感性高于 B 组,特异性两者相仿,虽然结果均不具有统计学意义,但也显示低年资医师联合 AI 后诊断新冠肺炎部分效能接近甚至高于高年资医师,特别适用于大多数医院急诊通常只有 1 名放射医师值班时, AI 具有“预警”和提醒医师新冠肺炎可能存在作用。

除此之外,由于新冠肺炎大多呈多发病灶,随病程进展病灶影像变化较大^[4],学者研究发现可以有早期→消散期;早期→进展期→消散期;早期→进展期→重症期→消散期等类型^[19],病灶占整肺体积百分比是预测新冠肺炎转归重要指标之一^[20], AI 可以自动测量占整肺体积百分比,有利于对新冠肺炎转归判断。

3. AI 应用伦理学

AI 加入使原本医-患关系变成医-AI-患的三角性关系^[21],而医生是承担医疗及法律责任主体,所以确定医生为主、AI 为辅主次关系非常重要。

参考文献:

- [1] Yan Li, Liming Xia. Coronavirus disease 2019 (COVID-19): role of chest CT in diagnosis and management[J]. AJR, 2020, 214(6): 1280-1286.
- [2] Zi Yue Zu, Meng Di Jiang, Peng Peng Xu, et al. Coronavirus disease 2019 (COVID-19): a perspective from China[published online ahead of print, 2020 Feb 21]. Radiology, 2020, 200490. DOI: 10.1148/radiol.2020200490.
- [3] Tao Ai, Zhenlu Yang, Hongyan Hou, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases[published online ahead of print, 2020

- Feb 26]. Radiology, 2020, 200642. DOI: 10.1148/radiol.2020200642
- [4] Harrison X Bai, Ben Hsieh, Zeng Xiong, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT[published online ahead of print, 2020 Mar 10]. Radiology, 2020, 200823. DOI: 10.1148/radiol.2020200823
- [5] Shuchang Zhou, Yujin Wang, Tingting Zhu, et al. CT features of coronavirus disease 2019 (COVID-19) pneumonia in 62 patients in Wuhan, China[J]. AJR, 2020, 214(6): 1287-1294.
- [6] Adam Bernheim, Xueyan Mei, Mingqian Huang, et al. Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection[published online, 2020 Feb 20]. Radiology, 2020, 295(3): 200463.
- [7] 雷志丹, 葛英辉, 史大鹏. 肺部弥漫性磨玻璃阴影的 CT 诊断与鉴别诊断[J]. 中国医学影像技术, 2007, 23(8): 1147-1151.
- [8] 施裕新, 黎淑娟, 周粟, 等. 甲型 H1N1 流感胸部高分辨率 CT 表现[J]. 中华放射学杂志, 2010, 44(2): 127-129.
- [9] 云永兴, 王立非, 杨根东, 等. H7N9 病毒性肺炎胸部 CT 随访分析[J]. 中国中西医结合影像学杂志, 2017, 15(3): 277-279.
- [10] Wenjie Yang, Fuhua Yan. Patients with RT-PCR confirmed COVID-19 and normal chest CT[J]. Radiology, 2020, 295(2): E3.
- [11] Yoshiko Aiji, Motoki Fukuda, Yoshitaka Kise, et al. Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using a deep learning system of artificial intelligence[J]. Oral Surg Oral Med Oral Pathol Oral Radiol, 2019, 127(5): 458-463.
- [12] Clinton J Wang, Charlie A Hamm, Lynn J Savic, et al. Deep learning for liver tumor diagnosis Part II: convolutional neural Network interpretation using radiologic imaging features[J]. Eur Radiol, 2019, 29(7): 3348-3357.
- [13] 李欣, 梁森, 黄正南, 等. 乳腺 X 线 AI 智能病灶检测[J]. 放射学实践, 2018, 33(10): 1029-1032.
- [14] 金文忠, 陆耀, 汪阳. 基于人工智能的胸部 CT 智能辅助诊断系统在 LDCT 数据集上的应用研究[J]. 中国医学计算机成像杂志, 2018, 24(5): 373-377.
- [15] 蔡雅倩, 张正华, 韩丹, 等. AI 对肺磨玻璃结节筛查及定性的临床应用研究[J]. 放射学实践, 2019, 34(9): 958-962.
- [16] 刘小玉, 谢元亮, 张树桐, 等. 肺部 CT 在新型冠状病毒感染性肺炎中的诊断价值[J]. 放射学实践, 2020, 35(4): 433-436.
- [17] 丁义, 黄增发, 赵胜超, 等. 新型冠状病毒肺炎(COVID-19)临床及影像学特征分析[J]. 放射学实践, 2020, 35(3): 281-285.
- [18] 侯可可, 张娜, 李桃, 等. 新型冠状病毒肺炎不同时期 CT 表现及中性粒细胞/淋巴细胞比值、T 淋巴细胞亚群变化[J]. 放射学实践, 2020, 35(3): 272-276.
- [19] 汪汉林, 梁奕, 王佳, 等. 91 例新型冠状病毒肺炎 CT 演变特征[J]. 放射学实践, 2020, 35(6): 701-706.
- [20] 刘长华, 余晓洁, 杨先春, 等. 胸部 CT 特征预测普通型新型冠状病毒肺炎患者转归的价值[J]. 放射学实践, 2020, 35(5): 587-591.
- [21] 陈默. 人工智能辅助医疗的医患关系伦理机制重构研究[J]. 医学与哲学, 2018, 39(9A): 39-41.

(收稿日期: 2020-04-12 修回日期: 2020-07-16)