

【临床研究】

作者简介:耿 然(1982-),男,河南洛阳人,学士,主治医师,研究方向:呼吸系统影像诊断。

film-reading method, and 93 nodules (54.38%) were detected by AI film-reading method but not by manual film-reading method; the detection rate of nodules detected by AI film-reading method was significantly higher than that of manual film-reading method ($\chi^2 = 4.317, P < 0.01$). Of the 83 nodules with the diameter from 5 mm \leq to ≤ 10 mm, 83 nodules (100.00%) were detected by AI film-reading method, and 45 nodules (54.21%) were detected by manual film-reading method; among them, 0 nodules (0.00%) was detected by manual film-reading method but not detected by AI film-reading method, and 38 nodules (45.78%) were detected by AI film-reading method but not by manual film-reading method; the detection rate of AI film-reading method was significantly higher than that of manual film-reading method ($\chi^2 = 1.744, P < 0.01$). The detection rate of AI film-reading method and manual film-reading method were 100.00% (36/36) in 36 nodules with the diameter > 10 mm. Of the 290 true nodules, 281 nodules were detected by AI film-reading method, including 99 low-risk nodules, 52 medium low-risk nodules, 38 medium-risk nodules, and 92 high-risk nodules (malignant nodules); 159 nodules were detected by manual film-reading method, including 59 low-risk nodules, 24 medium low-risk nodules, 21 medium-risk nodules and 55 high-risk nodules (malignant nodules). There were 66 malignant nodules confirmed by pathology, of which 62 nodules were correctly confirmed by AI film-reading method, and 4 nodules were not detected by AI film-reading method, with an accuracy rate of 93.93% (62/66); there were 52 malignant nodules were correctly confirmed by manual film-reading method, and 14 nodules were not detected by manual film-reading method, with an accuracy rate of 78.79% (52/66); the accuracy rate of malignant nodules correctly confirmed by AI film-reading method was significantly higher than that of manual film-reading method ($\chi^2 = 3.216, P < 0.05$). The sensitivity and false positive rate of AI film-reading method in the diagnosis of pulmonary malignant nodules were significantly higher than those of manual film-reading method, and the specificity and false negative rate were significantly lower than those of manual film-reading method ($\chi^2 = 2.311, 4.165, 7.896, 2.311; P < 0.05$). **Conclusion** AI image aided diagnostic system has a high detection rate and sensitivity for pulmonary nodules, however, but its specificity is lower than that of manual film-reading method, with a certain false positive rate. In clinical work, the mode of AI image aided diagnosis system film-reading combined with manual film-reading should be applied to diagnose pulmonary nodules, so as to effectively improve the accuracy of early detection of pulmonary nodules and differential diagnosis between benign and malignant, and reduce the rate of missed diagnosis.

Key words: artificial intelligence; computed tomography; pulmonary nodules; diagnostic efficiency; specificity; false positive rate

肺癌是最常见的恶性肿瘤之一,据美国癌症学会 2020 年 1 月公布的数据,肺癌的发病率和病死率均居恶性肿瘤首位^[1],早期诊断与治疗可有效提高肺癌患者的生存率^[2]。肺结节是肺癌的早期表现,其检出率随着薄层计算机断层扫描(computed tomography, CT)技术的发展而提高,但明显增多的 CT 图像数据量也增加了影像诊断医师的阅片负担,从而可能导致肺结节的漏诊^[3]。随着人工智能(artificial intelligence, AI)技术的进步,通过深度学习算法对肺结节进行识别、分类、重建及降噪处理,能够提取影像图像上肺结节的主要特征,对肺结节的早期发现及病变良恶性质的判断等进行预测^[4-5]。本研究旨在通过比较 AI 影像辅助诊断系统与常规人工阅片在肺结节检出率以及对肺恶性结节定性的灵敏度、特异度、假阳性率、假阴性率等方面的差异,探讨 AI 影像辅助诊断系统诊断胸部 CT 肺结节的价值。

1 资料与方法

1.1 一般资料 回顾性分析 2016 年 1 月至 2019 年 12 月于洛阳市第一人民医院经术前 CT 检查及

病理检查结果确诊的 85 例肺结节患者的胸部 CT 影像资料。其中男 33 例,女 52 例;年龄 25 ~ 79 (55.23 \pm 11.02) 岁。结节的定性以病理检查结果为金标准。病例纳入标准:(1) CT 检查前未接受过穿刺活检和手术、放射治疗等相关治疗;(2) 对于高度疑似恶性结节者,CT 检查后 1 个月内行 CT 引导下穿刺活检或手术,获得明确的病理诊断;(3) 对于判定为良性结节者,连续随访 1 a,1 a 内肺结节无变化者定为良性结节,1 a 内肺结节有进展者以穿刺活检或手术切除后的病理诊断判定其良恶性。排除标准:(1) 其他部位原发性恶性肿瘤肺部转移者;(2) 缺乏层厚 1.0 mm 及以下的薄层 CT 图像;(3) 图像有呼吸运动伪影及金属伪影干扰病变者。

1.2 CT 检查方法 所有患者均于术前行胸部 CT 扫描。使用联影 UCT528 40 排 CT(上海联影医疗科技股份有限公司)及西门子 Definition AS + 128 层 CT 机[西门子(中国)有限公司]分别进行扫描。扫描范围:肺尖至肺底水平,扫描参数:层厚 5.0 mm,层距 5.0 mm,管电压 120 kV,管电流 auto dose。扫描结束后对全肺行肺窗薄层 CT 重建,重建层厚 1.0 mm。

1.3 结节危险度诊断标准 根据结节的形态、大小、边缘状态、结节内实性成分的位置及结节邻近胸膜改变等影像特征作为诊断结节危险度的标准。(1)结节形态:斑片状形态可见于任何良恶性结节中;条片状形态多见于良性结节中;形态不规则以恶性结节居多,且病灶恶性程度越高形态不规则越明显。根据结节的形态进行计分:规则为 0 分,不规则为 1 分,高度不规则为 2 分。(2)结节大小:排除炎症病变后,结节体积越大性质越趋于恶性,且恶性程度趋于更高。根据结节的平均直径进行计分:直径 < 5 mm 为 0 分,5 mm ≤ 直径 ≤ 10 mm 为 1 分,直径 > 10 mm 为 2 分。(3)结节边缘状态:观察结节边缘是否模糊及有无毛刺征,边缘清晰的结节以炎症结节居多。根据结节边缘的状态进行计分:清晰为 0 分,模糊为 1 分,毛刺为 2 分。(4)结节内实性成分的位置:结节内无实性成分或实性成分位于结节边缘者多为良性病变,实性成分位于中心或偏中心位置的结节多为恶性病变。根据结节内实性成分的位置进行计分:结节内无实性成分或实性成分位于边缘计为 0 分,结节内实性成分偏中心位置计为 1 分,结节内实性成分位于结节中心计为 2 分。(5)结节邻近部位胸膜的变化:结节邻近部位胸膜的变化包括胸膜增厚、胸膜凹陷、胸膜牵引,其在良、恶性结节中的分布差异不明显,但胸膜牵引以恶性结节居多^[6],宜作为辅助征象。根据结节邻近部位胸膜的变化进行计分:胸膜无改变计为 0 分,胸膜增厚或凹陷计为 1 分,胸膜牵引计为 2 分。将每个结节的各项评分进行合计,根据其总分进行定性诊断:0 ~ 2 分为低危结节,3 ~ 4 分为中低危结节,5 ~ 7 分为中危结节,≥ 8 分为高危结节,将高危者定义为恶性结节。

1.4 图像分析 分别采用常规人工阅片和 AI 影像辅助诊断系统对患者的 CT 图像进行分析和诊断。(1)常规人工阅片:采用单纯人工阅片方式,由 3 位从事 CT 诊断工作满 5 a 的主治医师组成诊断组,依据“1.3”项的结节危险度诊断标准分别对检出的肺结节进行定性诊断。如组内医师对肺结节的定性诊断有分歧,则经过讨论达成一致意见,最终判定结节的性质。(2)AI 影像辅助诊断系统阅片:将 85 例患者的肺结节 CT 图像输入医准智能 Direct AI 系统(北京医准智能科技有限公司),通过网络学习自动提取结节的形态、大小、边缘状态、结节内实性成分的位置及结节邻近胸膜改变的特征,依据“1.3 项”的结节危险度诊断标准对检出的肺结节进行定性诊断。

1.5 诊断效能评价 通过比较常规人工阅片与 AI 影像辅助诊断系统阅片对肺结节诊断的灵敏度、假阳性率、假阴性率及评定结节性质的特异度,评估这 2 种阅片方法对肺结节定性的诊断效能。

1.6 统计学处理 应用 SPSS 22.0 软件进行统计学分析。计数资料以例数和百分率表示,组间比较采用 χ^2 检验, $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 常规人工阅片与 AI 影像辅助诊断系统阅片对肺结节检出情况比较 85 例患者共检出真性结节 290 枚,其中平均直径 < 5 mm 结节 171 枚,5 mm ≤ 直径 ≤ 10 mm 结节 83 枚,直径 > 10 mm 结节 36 枚。171 枚平均直径 < 5 mm 的小结节中,AI 检出 162 枚(94.74%),人工阅片检出 78 枚(45.61%),其中人工阅片检出而 AI 阅片未检出 9 枚(5.26%),人工阅片未检出而 AI 阅片检出 93 枚(54.38%),AI 阅片检出率显著高于人工阅片检出率,差异有统计学意义($\chi^2 = 4.317, P < 0.01$)。83 枚 5 mm ≤ 直径 ≤ 10 mm 的结节中,AI 阅片检出 83 枚(100.00%),人工阅片检出 45 枚(54.21%),其中人工阅片检出而 AI 阅片未检出 0 枚(0.00%),人工阅片未检出而 AI 阅片检出 38 枚(45.78%),AI 阅片检出率显著高于人工阅片检出率($\chi^2 = 1.744, P < 0.01$)。36 枚平均直径 > 10 mm 的结节中,AI 阅片检出率和人工阅片检出率均为 100.00%(36/36)。

2.2 常规人工阅片与 AI 影像辅助诊断系统阅片对肺结节良恶性的评定结果比较 在检出的 290 枚真性结节中,AI 阅片检出 281 枚,其中 AI 阅片评定为低危结节 99 枚、中低危结节 52 枚、中危结节 38 枚、高危结节(恶性结节)92 枚,另有 9 枚未检出;人工阅片检出 159 枚,其中人工阅片评定为低危结节 59 枚、中低危结节 24 枚、中危结节 21 枚、高危结节(恶性结节)55 枚,另有 131 枚未检出。经病理证实的恶性结节共 66 枚,其中 AI 判定正确 62 枚,4 枚未检出,准确率为 93.93%(62/66),AI 判定为恶性而实际为良性者 30 枚;人工阅片判定正确者 52 枚,14 枚未检出,准确率为 78.79%(52/66),人工阅片判定为恶性而实际为良性者 3 枚;AI 判定准确率显著高于人工阅片,差异有统计学意义($\chi^2 = 3.216, P < 0.05$)。

2.3 2 种阅片方法对肺结节良恶性判定的效能比较 结果见表 1。85 例患者中,病理诊断为良性 24 例,其中结核 13 例,隐球菌 3 例,炎症结节 8 例;病

理诊断为恶性 61 例,其中肺鳞癌 7 例,腺癌 54 例。AI 阅片对肺恶性结节检查的灵敏度、假阳性率显著高于人工阅片,特异度和假阴性率显著低于人工阅片,差异均有统计学意义($\chi^2 = 2.311、4.165、7.896、2.311, P < 0.05$)。

表 1 2 种阅片方法对肺结节良恶性判定的效能比较
Tab. 1 Comparison of efficacy of the two film-reading methods in the judgement of benign and malignant pulmonary nodules

阅片方法	灵敏度/%	特异度/%	假阳性率/%	假阴性率/%
AI 阅片	93.9	64.5	32.6	6.0
人工阅片	78.7	85.2	4.5	21.2
χ^2	2.311	4.165	7.896	2.311
P	<0.05	<0.05	<0.05	<0.05

3 讨论

AI 技术是随着大数据的进步而兴起的一门新兴技术,其在医学领域中得到广泛应用,尤其在肺结节的检测方面具有较大进展^[6]。基于 AI 的计算机辅助诊断(AI computer-aided diagnosis, CAD)系统方案大多比较复杂和耗时,需要更多的图像处理模块,且相关敏感度及特异度较低^[7]。目前,深度学习是 AI 技术的研究热点,其敏感度和特异度较传统的 CAD 高,在肺结节检出及定性方面应用较多。但以往的研究对不同模型算法的 AI 的探讨大多侧重于其原理、算法步骤等^[8],很少从临床应用角度评价其对肺结节的诊断价值。深度学习通过算法使 AI 可以从影像资料中提取肺结节的各项特征数据,通过大量的数据分析,从中提取到规律,从而对肺结节进行识别并判断肺结节的良恶性。基于此,本研究结合病理和随访结果,对比常规人工阅片与 AI 影像辅助诊断系统阅片对肺结节的诊断效果,探讨 AI 影像辅助诊断系统诊断肺结节的价值。

本研究结果显示,对于直径 <5 mm 的微小结节及 5 mm ≤ 直径 ≤ 10 mm 的小结节, AI 检出率分别为 94.74%、100.00%,显著高于常规人工阅片。对于直径 >10 mm 的结节, AI 及常规人工阅片均能全部检出。在对肺结节的定性分析中, AI 的特异度为 64.5%,而常规人工阅片的特异度为 85.2%。分析其原因如下:(1) AI 对与肺纹理位置相近、形态相似病灶的检出率高于常规人工阅片。(2) AI 对于平均直径 <5 mm 的微小结节及 5 mm ≤ 直径 ≤ 10 mm 的小结节检出率高于常规人工阅片,尤其是对微小结节的诊断效果更显著,可以明显降低临床诊断的漏诊率及假阴性率,但假阳性率也明显增高。对于肺

结节的定性, AI 的特异度显著低于人工阅片,尤其是对直径 <5 mm 的微小结节定性中特异度更低。李欣菱等^[2]研究认为, AI 对较小结节(特别是 <5 mm 的结节)诊断的临床意义有待商榷,若完全按照 AI 的结果进行临床干预会增加医疗负担及造成过度治疗。(3)目前肺结节诊断主要依赖于薄层 CT,图像数据量大,影像医师相对短缺,导致工作压力过大,精神及注意力无法长时间集中,进而造成肺小结节的漏诊。(4)人工阅片对于肺小结节良恶性的鉴别有赖于影像医师自身的临床经验,具有较强的主观性,从而导致诊断精确度和可信度较低。

肺癌的早期诊断、早期治疗是延长患者生存期的最佳途径^[9-10]。肺结节作为肺癌的早期表现,其早期发现及良恶性的鉴别对肺癌的早期诊断、早期治疗有重大意义,是目前临床医师工作中面临的挑战。本研究结果显示, AI 阅片诊断肺结节的灵敏度、假阴性率高于常规人工阅片,但其假阳性率亦高于常规人工阅片,特异度低于常规人工阅片。分析其原因,主要是由于 AI 会将支气管内黏液栓、交叉的血管、肺纤维化灶、局部肺实变、肺尖部胸膜帽等误诊为高危结节,所以需要进一步加强 AI 软件数据集的质量管理,扩大数据集训练数量,以进一步减少误诊,提升恶性结节检出的准确性^[11]。虽然常规人工阅片的特异度高于 AI 阅片,但对于早期缺乏特异影像学表现的肺结节,影像医师无法客观评价其形态、大小、边缘、病灶内实性成分的位置及病灶邻近胸膜改变情况,只能依据既往经验进行主观判断,进而影响肺小结节良恶性的判断;而 AI 阅片是基于对大数据深度学习基础之上的进一步分析,能对早期恶性肺结节进行客观分析及诊断。

综上所述, AI 影像辅助诊断系统阅片有较高的检出率及灵敏度,能有效减少临床工作中的漏诊率;但其特异度低于常规人工阅片,有一定的假阳性率。因此,在临床工作中推荐运用 AI 影像辅助诊断系统阅片联合常规人工阅片的模式,以有效提高对肺结节的早期发现及良恶性鉴别的准确度,减少漏诊率。

参考文献:

[1] 中华医学会呼吸病学分会肺癌学组,中国肺癌放置联盟专家组. 肺结节诊治中国专家共识(2018 年版)[J]. 中华结核和呼吸杂志, 2018, 41(10): 763-771.
LUNG CANCER GROUP OF RESPIRATORY BRANCH OF CHINESE MEDICAL ASSOCIATION, EXPERT GROUP OF CHINA LUNG CANCER PREVENTION AND CONTROL ALLIANCE. Consensus of Chinese expert on diagnosis and treatment of pulmo-

nary nodules(2018 edition)[J]. *Chin J Tuberc Respir Dis*,2018,41(10):763-771.

[2] 李欣菱,郭芳芳,周振,等. 基于深度学习的人工智能胸部 CT 肺结节检测效能评估[J]. 中国肺癌杂志,2019,22(6):336-340.

LI X L, GUO F F, ZHOU Z, *et al.* Performance of deep-learning-based artificial intelligence on detection of pulmonary nodules in chest CT[J]. *Chin J Lung Cancer*,2019,22(6):336-340.

[3] 刘娜,赵正凯,邹佳瑜,等. 基于人工智能的胸部 CT 肺结节检出及良恶性诊断效能评估[J]. CT 理论与应用研究,2021,30(6):709-715.

LIU N, ZHAO Z K, ZOU J Y, *et al.* Evaluation of detection and diagnostic efficiency of pulmonary nodules by chest CT based on artificial intelligence[J]. *Comput Tomogr Theory Applic*,2021,30(6):709-715.

[4] 宋振春,于铁链. 基于人工智能肺结节 CT 特征性诊断优势分析[J]. 影像研究与医学应用,2020,4(20):67-68.

SONG Z C, YU T L. Advantage analysis of CT characteristic diagnosis of pulmonary nodules based on artificial intelligence[J]. *J Imaging Res Med Appl*,2020,4(20):67-68.

[5] CUI S J, MING S, LIN Y, *et al.* Development and clinical application of deep learning model for lung nodules screening on CT images[J]. *Scientific Reports*,2020,10(1):13657.

[6] WILSHIRE C L, LOUIE B E, MANNING K A, *et al.* Radiologic evaluation of small lepidic adenocarcinomas to guide decision making in surgical resection[J]. *ANN Thorac Surg*,2015,100(3):979-988.

[7] WANG Y, WU B, ZHANG N, *et al.* Research progress of computer aided diagnosis system for pulmonary nodules in CT images[J]. *J Xray Sci Technol*,2020,28(1):1-16.

[8] WANG X Q, MAO K M, WANG L Z, *et al.* An appraisal of lung nodules automatic classification algorithms for CT images[J]. *Sensors (Basel)*,2019,19(1):194.

[9] REVEL M P, MANNES I, BENZAKOUN J, *et al.* Subsolid lung nodule classification: a CT criterion for improving interobserver agreement[J]. *Radiology*,2018,286(1):316-325.

[10] MACMAHON H, NAIDICH D P, GOO J M, *et al.* Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner society 2017[J]. *Radiology*,2017,284(1):228-243.

[11] TRAVIS W D, BRAMBILLA E, NOGUCHI M, *et al.* International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma[J]. *J Thorac Oncol*,2011,6(2):244-285.

(本文编辑:李胜利)

(上接第 1030 页)

[23] 吴婉雯,李晓伟,崔立红. 脂肪肝诊断标准的变更对心血管疾病风险分层管理的价值分析[J]. 解放军医学杂志,2021,46(5):486-491.

WU W F, LI X W, CUI L H. The value of changes in diagnostic criteria for fatty liver in stratified management of cardiovascular risk[J]. *Med J Chin PLA*,2021,46(5):486-491.

[24] 郭亮,任璐. 慢性心力衰竭患者高密度脂蛋白胆固醇、高敏肌钙蛋白 T 水平与心功能的相关性[J]. 解放军预防医学杂志,2017,35(9):1091-1094.

GUO L, REN L. Relationships between high-density lipoprotein, high-sensitivity cardiac troponin T and cardiac function in patients with chronic heart failure[J]. *Med J Chin PLA*,2017,35(9):1091-1094.

[25] 吴祖飞,陈诗,刘叶红,等. 血清 CRP/白蛋白、HCY/HDL-C 与冠状动脉病变程度的相关性分析[J]. 解放军医学杂志,2021,46(7):678-686.

WU Z F, CHEN S, LIU Y H, *et al.* Relationship between CRP/albumin, HCY/HDL-C and the severity of coronary artery disease[J]. *Med J Chin PLA*,2021,46(7):678-686.

[26] SANCHES P L, DE PIANO A, CAMPOS R M, *et al.* Association of nonalcoholic fatty liver disease with cardiovascular risk factors in obese adolescents: the role of interdisciplinary therapy[J]. *J Clin Lipidol*,2014,8(3):265-272.

[27] SCHWIMMER J B, PARDEE P E, LAVINE J E, *et al.* Cardiovascular risk factors and the metabolic syndrome in pediatric nonalcoholic fatty liver disease[J]. *Circulation*,2008,118(3):277-283.

[28] SPAHIS S, DELVIN E, BORYS J M, *et al.* Oxidative stress as a critical factor in nonalcoholic fatty liver disease pathogenesis[J]. *Antioxid Redox Signal*,2017,26(10):519-541.

[29] SEO Y B, HAN A L. Association of the serum uric acid-to-creatinine ratio with nonalcoholic fatty liver disease diagnosed by computed tomography[J]. *Metab Syndr Relat Disord*,2021,19(2):70-75.

[30] 邵翠萍,徐有青. 血清尿酸/肌酐比值与非酒精性脂肪性肝病的相关性分析[J]. 临床肝胆病杂志,2021,37(10):2348-2351.

SHAO C P, XU Y Q. Association of serum uric acid-to-creatinine ration with nonalcoholic fatty liver disease[J]. *J Clin Hepatol*,2021,37(10):2348-2351.

(本文编辑:郭 潇)