

· 临床研究 · doi:10.3969/j.issn.1671-8348.2024.05.015

网络首发 [https://link.cnki.net/urlid/50.1097.R.20240227.1615.020\(2024-02-27\)](https://link.cnki.net/urlid/50.1097.R.20240227.1615.020(2024-02-27))

不同重建层厚 CT 图像对 AI 诊断肋骨骨折效能的影响*

敖平,朱丽,修志刚[△],肖涵,李为民
(四川大学华西医院龙泉医院放射科,成都 610100)

[摘要] **目的** 探讨不同重建层厚 CT 图像对人工智能(AI)诊断肋骨骨折效能的影响。**方法** 选取 100 例肋骨骨折患者的首次 CT 图像,分别以 0.625、1.250、2.500、5.000 mm 的层厚进行无间隔重建,应用 AI 的肋骨骨折筛查功能对 4 组层厚 CT 图像进行自动检测,比较不同重建层厚条件下 AI 对肋骨骨折诊断效能的差异。**结果** AI 在 0.625、1.250、2.500、5.000 mm 层厚诊断肋骨骨折的灵敏度分别为 99.32%(436/439)、98.41%(432/439)、89.52%(393/439)、83.60%(367/439),假阳性率分别为 4.80%(22/458)、0.92%(4/436)、0.76%(3/396)、0.27%(1/368)。AI 在 0.625 mm 及 1.250 mm 层厚的诊断灵敏度高于 2.500 mm 及 5.000 mm,差异有统计学意义($P < 0.05$),而 0.625 mm 与 1.250 mm 层厚比较差异无统计学意义($P > 0.05$)。AI 在 0.625 mm 层厚诊断的假阳性率高于 1.250、2.500 及 5.000 mm,差异有统计学意义($P < 0.05$),而 1.250、2.500 及 5.000 mm 层厚之间比较差异无统计学意义($P > 0.05$)。**结论** 1.250 mm 层厚 CT 图像对 AI 诊断肋骨骨折效能优于 0.625、2.500 及 5.000 mm 层厚。

[关键词] 人工智能;多层螺旋 CT;重建;层厚;肋骨骨折;诊断效能

[中图分类号] R445.3

[文献标识码] A

[文章编号] 1671-8348(2024)05-0723-04

Diagnostic efficacy of AI in rib fracture under CT images with different reconstruction slice thickness*

AO Ping, ZHU Li, XIU Zhigang[△], XIAO Han, LI Weimin

(Department of Radiology, West China Longquan Hospital Sichuan University,
Chengdu, Sichuan 610100, China)

[Abstract] **Objective** To investigate the diagnostic efficiency of artificial intelligence (AI) in rib fracture under the computed tomography (CT) images with different reconstruction slice thickness. **Methods** The first CT images of 100 patients with rib fractures were selected, and the interval-free reconstruction was carried out with the thickness of 0.625 mm, 1.250 mm, 2.500 mm and 5.000 mm, respectively. The rib fracture screening function of AI was used to automatically detect the CT images of four groups, and the diagnostic efficiency of AI for rib fracture under different reconstruction thickness conditions was compared. **Results** The sensitivity of AI in the diagnosis of rib fracture at 0.625 mm, 1.250 mm, 2.500 mm and 5.000 mm thickness was 99.32% (436/439), 98.41% (432/439), 89.52% (393/439) and 83.60% (367/439), respectively. The false positive rate was 4.80% (22/458), 0.92% (4/436), 0.76% (3/396) and 0.27% (1/368). The diagnostic sensitivity of AI in 0.625 mm and 1.250 mm thickness was higher than that in 2.500 mm and 5.000 mm, and the difference was statistically significant ($P < 0.05$), while there was no significant difference in the thickness of 0.625 mm and 1.250 mm. The false positive rate of AI in the diagnosis of 0.625 mm slice thickness was higher than that of 1.250 mm, 2.500 mm and 5.000 mm, and the difference was statistically significant ($P < 0.05$), while there was no significant difference in the thickness of 1.250 mm, 2.500 mm and 5.000 mm ($P > 0.05$). **Conclusion** The diagnostic efficiency of AI in 1.250 mm CT images is better than that in 0.625 mm, 2.500 mm and 5.000 mm CT images.

[Key words] artificial intelligence; multi-slice spiral CT; reconstruction; slice thickness; rib fracture; diagnostic efficiency

肋骨骨折是胸部外伤患者最常见的骨损伤,准确 诊断肋骨骨折对患者合理诊治、改善预后具有重要意

义^[1-3]。随着多层螺旋 CT (multi-slice spiral CT, MSCT) 技术的快速发展及设备的广泛普及, 因其可以快速完成检查并准确有效诊断肋骨骨折, 已被公认为肋骨骨折诊断的最佳检查方法^[4-7]。肋骨呈双侧弧形走向, 诊断肋骨骨折需要逐侧、逐肋地反复观察, 加上肋骨扫描产生的 CT 图像数量繁多, 放射医师的阅片负担十分繁重, 极易引起视觉疲劳, 造成骨折尤其是细微骨折的漏诊^[7-8]。近年来, 随着人工智能 (artificial intelligence, AI) 技术的迅猛发展, AI 在医疗领域逐步开始广泛应用, 基于深度学习的 AI 辅助诊断技术在乳腺癌、肺结节、甲状腺结节等疾病的筛查、诊断方面取得了较好的效果, 获得临床的认可^[9-16], AI 辅助诊断肋骨骨折也逐步进入大家的研究视线^[17-19]。

已有的肺结节 AI 辅助诊断相关研究显示, 不同重建层厚的 CT 图像会对 AI 检测结果产生影响^[20-21], 但其是否会对 AI 诊断肋骨骨折效能产生影响目前鲜有报道。本研究拟就不同重建层厚 CT 图像对 AI 诊断肋骨骨折效能的影响进行分析, 以期寻找适合 AI 诊断肋骨骨折的 CT 重建层厚。

1 资料与方法

1.1 一般资料

收集 2022 年 9 月至 2023 年 5 月在本院诊治的 100 例肋骨骨折患者资料, 男 69 例, 女 31 例, 年龄 18~71 岁, 平均 (43.82±13.12) 岁, 患者均进行肋骨 MSCT 检查。纳入标准: (1) 有明确胸部外伤病史, 首次 MSCT 检查距受伤不超过 48 h; (2) 初次检查后 2~8 周有至少 1 次肋骨 MSCT 复查随访资料; (3) AI 软件能成功计算肋骨骨折结果。排除标准: (1) 图像质量差, 影响诊断; (2) 明显胸廓畸形; (3) 肋骨骨质破坏或骨肿瘤; (4) 复查期间再次出现胸部外伤者。本研究经医院医学伦理委员会批准 (AF-KY-2022003)。

1.2 方法

1.2.1 仪器和参数

扫描设备采用美国 GE 公司 64 层螺旋 CT 机, 取仰卧位, 头先进, 双手上举或抱头, 无法上举者置于身体两侧, 患者吸气后屏气状态下一次完成扫描 (病情严重无法配合者除外), 由头侧向足侧扫描, 扫描范围自胸廓入口至肋弓下缘。扫描参数: 管电压 120 kV, 管电流 100~300 mAs, 层厚及层距均为 5.000 mm。将扫描获得的原始图像再分别以 0.625、1.250、2.500 mm 的层厚进行无间隔轴位重建, 将重建图像传输至图像存储与传输系统 (picture archiving communication system, PACS) 工作站。

1.2.2 诊断方法

采用上海联影智能医疗科技有限公司基于深度学习模型的 AI 肋骨骨折诊断软件 (uAI-BoneCare) 分别对 0.625、1.250、2.500、5.000 mm 层厚的 CT 图像进行检测, 记录检测结果。所有 CT 图像由两名具有 15 年以上诊断经验的放射科副主任医师借助 AI 软件结合 0.625 mm 层厚图像共同阅片分析, 以初次检

查后 2~8 周复查发现拟诊断肋骨骨折处出现折端错位或骨痂生长作为骨折诊断“金标准”。

1.2.3 诊断结果判定

分别计算 AI 基于 0.625、1.250、2.500、5.000 mm 层厚的 CT 图像对诊断肋骨骨折的灵敏度及假阳性率。灵敏度 = 真阳性骨折数 / (真阳性骨折数 + 假阴性骨折数) × 100%, 假阳性率 = 假阳性骨折数 / (真阳性骨折数 + 假阳性骨折数) × 100%, 因本研究无法获得真阴性骨折数, 故假阳性率采用此法计算。

1.3 统计学处理

采用 SPSS24.0 软件进行统计学分析。计数资料以例数或百分比表示, 采用 χ^2 检验, 以 $P < 0.05$ 为差异有统计学意义。

2 结果

100 例患者经“金标准”确认共有 439 处肋骨骨折, AI 在 0.625 mm 层厚共检出 458 处, 其中正确 436 处, 误诊 22 处, 漏诊 3 处; 1.250 mm 层厚共检出 436 处, 其中正确 432 处, 误诊 4 处, 漏诊 7 处; 2.500 mm 层厚共检出 396 处, 其中正确 393 处, 误诊 3 处, 漏诊 46 处; 5.000 mm 层厚共检出 368 处, 其中正确 367 处, 误诊 1 处, 漏诊 72 处。

AI 在 0.625、1.250、2.500、5.000 mm 层厚诊断肋骨骨折的灵敏度分别为 99.32%、98.41%、89.52%、83.60%, AI 在 0.625 mm 及 1.250 mm 层厚的诊断灵敏度高于 2.500 mm 及 5.000 mm, 差异有统计学意义 ($P < 0.05$), 而 0.625 mm 和 1.250 mm 层厚之间比较差异无统计学意义 ($P > 0.05$), 见表 1、2。

AI 在 0.625、1.250、2.500、5.000 mm 层厚诊断肋骨骨折的假阳性率分别为 4.80%、0.92%、0.76%、0.27%。AI 在 0.625 mm 层厚诊断肋骨骨折的假阳性率高于 1.250、2.500 及 5.000 mm, 差异有统计学意义 ($P < 0.05$), 而 1.250、2.500 及 5.000 mm 层厚之间比较差异无统计学意义 ($P > 0.05$), 见表 1、2。

表 1 AI 在不同重建层厚诊断肋骨骨折的灵敏度、假阳性率 [% (n/n)]

层厚(mm)	灵敏度	假阳性率
0.625	99.32(436/439)	4.80(22/458)
1.250	98.41(432/439)	0.92(4/436)
2.500	89.52(393/439)	0.76(3/396)
5.000	83.60(367/439)	0.27(1/368)

表 2 AI 在不同重建层厚诊断肋骨骨折灵敏度和假阳性率两两比较的统计量

层厚(mm)	灵敏度		假阳性率	
	χ^2	P	χ^2	P
0.625 vs. 1.250	1.722	0.088	3.931	<0.001
0.625 vs. 2.500	2.001	0.048	4.088	<0.001
0.625 vs. 5.000	2.001	0.047	4.396	<0.001

续表 2 AI 在不同重建层厚诊断肋骨骨折灵敏度和假阳性率两两比较的统计量

层厚(mm)	灵敏度		假阳性率	
	χ^2	<i>P</i>	χ^2	<i>P</i>
1.250 vs. 2.500	2.001	0.048	1.421	0.158
1.250 vs. 5.000	2.009	0.047	1.750	0.083
2.500 vs. 5.000	1.988	0.056	1.421	0.158

3 讨 论

随着社会的发展进步,人们生活节奏加快,意外事故频发,急性胸部创伤在临床工作中日益增多^[1-2],作为胸部创伤中最常见的骨性损伤,肋骨骨折患者也随之增多。肋骨骨折的严重程度可作为评判患者创伤和预后的指标,准确诊断肋骨骨折可为临床制订合理的治疗方案提供重要依据,改善患者预后^[1-3]。同时,部分胸部创伤患者因涉及工伤认定、伤情司法鉴定等相关问题,肋骨骨折的有无及骨折的准确数量对伤情鉴定等有重要影响^[4,8]。

MSCT 扫描速度快,扫描范围大,能在短时间内完成胸部乃至全腹部的扫描,尤其适用于危急重症患者快速完成检查^[6-8]。MSCT 采用容积扫描,所获得的数据实现了各向同性,可在此基础上进行不同层厚的图像重建及多种功能强大的图像后处理,得到清晰度高、立体直观的高质量图像,对肋骨骨折进行多角度、全方位的观察。MSCT 在清晰地显示肋骨骨折的同时,还能提供合并存在的胸腹部脏器损伤等信息,随着 MSCT 在各级医院的广泛普及,已逐步成为胸部外伤的首选检查方法^[4-7]。为了更好地显示肋骨细微骨折,肋骨 MSCT 多需进行薄层重建,从而产生大量的 CT 图像数据,放射医师阅片工作量明显增加。同时,由于肋骨特殊的解剖形态特点,同一根肋骨会在多个层面的 CT 断层图像显示,而同一 CT 断层图像当中又可以同时显示多根不同的肋骨,对肋骨的序数认定需要花费大量时间,加上部分肋骨骨折断端分离错位明显,多发肋骨骨折的诊断尤为耗时且烦琐,极易产生视觉疲劳,造成骨折漏诊^[17,19]。

近年来,以大数据为依托的 AI 辅助诊断技术开始应用于临床并显示出良好的效果。AI 可以快速对肋骨骨折进行识别和标记,其客观性及可重复性强。相关研究显示,AI 能够辅助放射医师提高肋骨骨折尤其是细微骨折的检出率,同时明显缩短诊断时间,提高了工作效率^[22-27]。

已有相关肺结节 AI 辅助诊断研究报道^[20-21]不同重建层厚对病灶检出效果存在一定的影响,CT 图像重建层厚越厚,容积效应越明显,随着层厚的增加,AI 对病灶的检出能力减弱,灵敏度降低。本研究中 AI 在 0.625、1.250、2.500 及 5.000 mm 层厚的 CT 图像均有漏诊,且随着层厚增加,漏诊数增多。究其原因主要与图像重建时骨折线能否充满层厚相关。当骨折较细微,仅有骨皮质发生细小断裂或骨折线位于重

建图像层面的一部分时,会因为层厚增加产生的部分容积效应所掩盖。而当骨折位于重建层面以外时,则会被直接遗漏。采用相对较厚的层厚 CT 图像诊断虽可减少放射医师的阅片数量,有利于缩短诊断时间,但却常常造成细微骨折的漏诊。因此在实际工作中,为了提高肋骨骨折的检出率,不建议采用较厚的层厚,如 2.500 mm 与 5.000 mm。

本研究结果显示,AI 在 0.625 mm 及 1.250 mm 层厚 CT 图像对肋骨骨折的诊断灵敏度较 2.500 mm 及 5.000 mm 高,差异有统计学意义($P < 0.05$),说明层厚越薄越能提高 AI 对肋骨骨折的检出率,这与 AI 诊断肺结节相关研究一致^[21]。重建层厚越薄,肋骨骨折能够充满层厚的机会越多,骨折显示概率就越大,越有利于 AI 识别病灶,对骨折诊断能力越强。但 CT 重建层厚并非越薄越好,本研究中虽然 AI 在 0.625 mm 层厚的 CT 图像检出肋骨骨折的灵敏度较 1.250 mm 层高,但二者差异并无统计学意义($P > 0.05$)。层厚越薄,重建所需时间越长,生成的薄层图像越多,导致 AI 所需运算量增加,相应的对医院网络建设、PACS 的传输速度、硬盘的存储空间等软件及硬件性能也提出更高的要求,成本相应增加^[28]。同时,在实际工作中,AI 完成骨折筛查以后,放射医师还需对 AI 所做出的诊断进行复核,层厚越薄,图像数量越多,医师阅片分析时间亦随之增加。此外,随着层厚变薄,虽然改善了部分容积效应的影响,但同时往往伴随着图像噪声的增加,对图像质量造成影响,对一部分细微病灶判定造成干扰^[20-21]。本研究中 0.625 mm 层厚的 CT 图像 AI 诊断肋骨骨折的假阳性率较其他层厚明显增高,差异有统计学意义($P < 0.05$)。本研究也存在一定局限性。纳入研究的均为医师回顾性发现骨折的病例,临床中漏诊的病例无法通过回顾性选择获得,同时本研究样本量较小,未能克服不同类别的样本量不均衡问题,因此结果可能存在一定的选择性偏倚。此外,本研究中所有病例均采用美国 GE 公司 CT 设备扫描,其图像重建层厚为 0.625 mm、1.250 mm、2.500 mm、5.000 mm 组合,得到 1.250 mm 为综合效能较优层厚,而其他公司的 CT 设备图像重建层厚可能是 0.750、1.000、2.000、3.000 mm 或 0.500、1.000、2.000、3.000 mm 等组合,较优层厚有待进一步研究。

综上所述,AI 对肋骨骨折诊断的效能跟 CT 重建层厚关系密切。1.250 mm 层厚的综合效能相对较优,在获得高灵敏度的同时,可减轻放射医师工作负担,更具实用性。

参考文献

- [1] 王伯珉,杨永良,贾宏磊,等.多发肋骨骨折的治疗进展[J].创伤外科杂志,2021,23(12):948-951.

- [2] 冉隆强,李勇,贾萌. 胸腔镜辅助内固定术对多发肋骨骨折合并胸部创伤患者肺功能及预后的影响[J]. 中国医刊,2023,58(5):517-520.
- [3] 靳贺,危朝辉,张琦,等. 肋骨骨折评分和肋骨评分对创伤性肋骨骨折患者伤情的评估价值及对不良预后事件的预测效能研究[J]. 创伤外科杂志,2023,25(6):437-442.
- [4] 张光霞,何英,杨绍光,等. 司法鉴定中肋骨骨折影像诊断符合率影响因素的研究[J]. 中国医药科学,2022,12(20):24-29.
- [5] 余开冠,孙芳仁. MSCT 与 DR 检查外伤性肋骨骨折的临床应用对比分析[J]. 医学影像学杂志,2022,32(7):1260-1263.
- [6] 付哲祥,李昌松. 多层螺旋 CT 的 MPR、SSD 及 VRT 重建技术在肋骨骨折中的临床应用[J]. 中国 CT 和 MRI 杂志,2021,19(7):164-166.
- [7] 马芳芳,李相生,方红,等. MSCT 三维重建、数字化 X 线摄影对多发性肋骨骨折的诊断价值对比[J]. 中国 CT 和 MRI 杂志,2022,20(4):166-168.
- [8] 姚越,杜北珏. 急诊与复查多层螺旋 CT 对肋骨骨折诊断价值的对比分析[J]. 安徽医药,2019,23(6):1128-1130.
- [9] 赵正凯,梁勇,周建收,等. 人工智能对肺磨玻璃结节检出及定性诊断价值研究[J]. 大连医科大学学报,2021,43(3):235-239.
- [10] DU W, HE B, LUO X, et al. Diagnostic value of artificial intelligence based on CT image in benign and malignant pulmonary nodules [J]. J Oncol,2022,2022:5818423.
- [11] 兰梓涵,彭玉兰. 基于乳腺超声图像的人工智能诊断[J]. 中南大学学报(医学版),2022,47(8):1009-1015.
- [12] FREEMAN K, GEPPERT J, STINTON C, et al. Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy [J]. BMJ,2021,374:n1872.
- [13] BALKENENDE L, TEUWEN J, MANN R M. Application of deep learning in breast cancer imaging [J]. Semin Nucl Med,2022,52(5):584-596.
- [14] 李晓宇,刘利平,辛雨薇,等. 人工智能软件联合超声造影鉴别诊断甲状腺良、恶性结节[J]. 中国医学影像学杂志,2023,31(3):226-230.
- [15] LUDWIG M, LUDWIG B, MIKUŁA A, et al. The use of artificial intelligence in the diagnosis and classification of thyroid nodules: an update [J]. Cancers (Basel),2023,15(3):708.
- [16] 林少坤,曾志雄,刘中华,等. 人工智能超声辅助诊断系统联合超声造影对 ACR TI-RADS 4 类结节的诊断价值[J]. 临床超声医学杂志,2023,25(2):142-144.
- [17] 熊山,陈博,毛杰,等. 基于深度学习的计算机辅助诊断系统在肋骨骨折诊断中的应用[J]. CT 理论与应用研究,2022,31(5):617-622.
- [18] AKIFUMI N, KOUZOU M, REI K, et al. Development of an artificial intelligence-assisted computed tomography diagnosis technology for rib fracture and evaluation of its clinical usefulness [J]. Sci Rep,2022,12(1):8363.
- [19] 李星宇,雷禹,黄晓旗,等. 基于 AI 与人工阅片对肋骨骨折性质诊断价值的比较[J]. 中国医疗设备,2022,37(8):32-36.
- [20] 周围,胡富碧,刘亚斌,等. 基于 AI 对肺小结节 CT 图像层厚与结节特征的检出效能影响[J]. 医学影像学杂志,2021,31(2):234-238.
- [21] 崔兆国,吴昊,汤敏,等. 重建层厚(1/2/3 mm)对人工智能检测肺结节效能的影响[J]. 中国医疗设备,2020,35(10):103-105.
- [22] 董浩,经齐峰,邱勇刚,等. 基于深度学习人工智能辅助 CT 检测肋骨骨折的价值[J]. 浙江临床医学,2022,24(6):914-915,919.
- [23] WANG S, WU D, YE L, et al. Assessment of automatic rib fracture detection on chest CT using a deep learning algorithm [J]. Eur Radiol,2023,33(3):1824-1834.
- [24] 刘想,谢辉辉,许玉峰,等. 人工智能在胸部创伤肋骨骨折 CT 诊断中应用的初步研究[J]. 上海交通大学学报(医学版),2021,41(7):920-925.
- [25] ZHOU Q Q, WANG J, TANG W, et al. Automatic detection and classification of rib fractures on thoracic CT using convolutional neural network: accuracy and feasibility [J]. Korean J Radiol,2020,21(7):869-879.
- [26] 贾春雪,张彬,吴润泽,等. 基于深度学习的人工智能在肋骨骨折检测中的应用价值[J]. 实用放射学杂志,2020,36(11):1861-1864.
- [27] 谭辉,田占雨,潘宁,等. 基于深度学习的计算机辅助诊断系统在提高急性肋骨骨折诊断效能上的价值[J]. 临床放射学杂志,2020,39(12):2493-2497.
- [28] 王伟军,习羽,田宏哲. CT 层厚对肺炎人工智能诊断软件诊断准确率的影响[J]. 医疗卫生装备,2022,43(2):60-63,69.