



## · 专题论著 ·



薛恩生，主任医师，硕士研究生导师，福建医科大学附属协和医院超声科主任。担任福建省医师协会超声医学分会会长，福建省超声医学质控中心主任，福建省超声医学研究所所长，中国医师协会超声医师分会常务委员，中国超声医学工程学会浅表器官及外周血管专业委员会副主任委员等。主要从事浅表器官疾病的超声诊断及介入治疗。

## C-TIRADS联合人工智能辅诊系统对甲状腺结节良恶性的诊断价值

郑梅娟，薛恩生，俞悦，陈舜，陈聪，方贞燕

福建医科大学附属协和医院超声科，福建省超声医学研究所，福建福州 350000

**[摘要]** 目的：探讨甲状腺结节超声恶性危险分层中国甲状腺影像报告和数据系统（Chinese-Thyroid Imaging Reporting and Data System, C-TIRADS）联合甲状腺结节人工智能（artificial intelligence, AI）辅诊系统对良恶性结节的诊断价值，并分析桥本甲状腺炎（Hashimoto thyroiditis, HT）背景对诊断结果的影响。方法：纳入817个甲状腺结节，以组织病理学检查结果为金标准，对比AI辅诊系统、不同年资医师应用C-TIRADS及联合应用的诊断效能；将结节分为HT背景组/非HT背景组，比较结节背景对AI辅诊系统、不同年资医师诊断结果的影响。结果：817个甲状腺结节中，恶性462个，良性355个。与高年资医师应用C-TIRADS相比，AI辅诊系统诊断的特异度较高（89.58% vs 81.69%， $P=0.003$ ），准确度两者相当（92.29% vs 90.58%， $P=0.151$ ）；与低年资医师应用C-TIRADS相比，AI辅诊系统诊断的特异度较高（89.58% vs 56.90%， $P=0.002$ ），准确度较高（92.29% vs 80.05%， $P<0.001$ ）。通过联合AI辅诊系统，高年资医师诊断的特异度提高（81.69% vs 89.01%， $P=0.006$ ），而准确度、灵敏度的差异均无统计学意义（ $P>0.05$ ）；低年资医师诊断的特异度（56.90% vs 86.76%， $P<0.001$ ）及准确度（80.05% vs 92.66%， $P<0.001$ ）均明显提高，灵敏度的差异无统计学意义（ $P=0.526$ ）。HT背景下，高年资医师的诊断特异度低于非HT组（69.64% vs 83.95%， $P<0.05$ ），低年资医师的诊断特异度也低于非HT组（42.86% vs 59.53%， $P<0.05$ ），但HT背景对AI辅诊系统诊断灵敏度、特异度及准确度均无影响。结论：C-TIRADS对甲状腺结节性质的鉴别具有较高的灵敏度；联合AI辅诊系统对高年资医师价值有限，但可提高低年资医师应用C-TIRADS的诊断特异度及准确度，避免过度诊断。

**[关键词]** 甲状腺结节；超声；人工智能；中国甲状腺影像报告和数据系统；诊断

DOI: 10.19732/j.cnki.2096-6210.2022.05.005

中图分类号: R736.1 文献标志码: A 文章编号: 2096-6210(2022)05-0484-07

**Evaluation of C-TIRADS combined with artificial intelligence-assisted diagnosis in differential diagnosis of thyroid nodules** ZHENG Meijuan, XUE Ensheng, YU Yue, CHEN Shun, CHEN Cong, FANG Zhenyan (Department of Ultrasound, Fujian Medical University Union Hospital, Fujian Medical Ultrasound Research Institute, Fuzhou 350000, Fujian Province, China)

基金项目: 福建省卫生健康科技计划项目资助 (2019-1-33); 福建省科技厅科技创新联合资金项目 (2019Y9104)

通信作者: 薛恩生 E-mail: xuees01@163.com

Correspondence to: XUE Ensheng E-mail: xuees01@163.com

[ **Abstract** ] **Objective:** To evaluate the application value of Chinese-Thyroid Imaging Reporting and Data System ( C-TIRADS ) combine artificial intelligence (AI)-assisted diagnosis in the differential diagnosis of benign and malignant thyroid nodules, and to analyze the effect of Hashimoto thyroiditis (HT) on the diagnostic efficacy. **Methods:** This study included 817 thyroid nodules confirmed histopathologically after thyroidectomy and compared the diagnostic efficacy of AI, C-TIRADS and combined application. Grouping the nodules according to the background of HT, then compared the diagnosis results of AI and sonographer with different seniority. **Results:** Of the 817 thyroid nodules, 462 were malignant and 355 were benign. Compared with senior sonographer using C-TIRADS, AI system had higher specificity (89.58% vs 81.69%,  $P=0.003$ ), comparable accuracy (92.29% vs 90.58%,  $P=0.151$ ). Compared with junior sonographer using C-TIRADS, AI system had higher specificity (89.58% vs 56.90%,  $P=0.002$ ), higher accuracy (92.29% vs 80.05%,  $P<0.001$ ). Combined with AI system, the specificity of senior sonographer improved (81.69% vs 89.01%,  $P=0.006$ ), while the difference in accuracy and sensitivity were not statistically significant ( $P>0.05$ ); and the specificity (56.90% vs 86.76%,  $P<0.001$ ) and accuracy (80.05% vs 92.66%,  $P<0.001$ ) of junior sonographer improved, while the difference in sensitivity was not statistically significant ( $P=0.526$ ). With the background of HT, the diagnostic specificity of senior physicians was lower than that of non-HT group (69.64% vs 83.95%,  $P<0.05$ ), and the diagnostic specificity of junior physicians was also lower than that of non-HT group (42.86% vs 59.53%,  $P<0.05$ ), but the background of HT had no influence on the diagnostic sensitivity, specificity and accuracy of AI system. **Conclusion:** C-TIRADS has high sensitivity for the differential diagnosis of thyroid nodules, but low diagnosis specificity. Combined AI system has limited value for senior sonographer, but can improve the diagnostic specificity and accuracy of C-TIRADS for junior sonographer and avoid overdiagnosis.

[ **Key words** ] Thyroid nodule; Ultrasound; Artificial intelligence; Chinese-Thyroid Imaging Reporting and Date System; Diagnosis

甲状腺癌是近年来全球发病率增长最快的恶性肿瘤<sup>[1]</sup>, 中国甲状腺癌发病率每年以20.1%的速度增长<sup>[2]</sup>。甲状腺癌的诊疗问题也越来越受人们关注。应用超声对甲状腺结节进行恶性风险分级评估, 在甲状腺疾病的规范化诊疗方面有着不可替代的作用。甲状腺结节超声恶性危险分层中国甲状腺影像报告和数据系统 ( Chinese-Thyroid Imaging Reporting and Data System, C-TIRADS )<sup>[3]</sup> 得到国内甲状腺专家的广泛认可, 它也强调了超声图像质量和医师经验对于甲状腺结节危险分层的重要性。深度学习技术越来越多地应用于医学影像学领域, 利用计算机自动客观地提取、分析影像学特征, 可以减少医师扫查手法、诊断经验等主观因素对诊断结果的影响, 辅助医师提高诊断的效率及准确度<sup>[4]</sup>。本研究旨在研究C-TIRADS联合基于深度学习技术的甲状腺结节人工智能 ( artificial intelligence, AI ) 辅诊系统对甲状腺结节良恶性的诊断价值, 并探讨桥本甲状腺炎 ( Hashimoto thyroiditis, HT ) 背景对AI辅诊系统、不同年资医师诊断结果的影响。

## 1 资料和方法

### 1.1 研究对象

收集2021年1—7月于福建医科大学附属协和医院行甲状腺外科手术的患者共870例。纳入标准: ① 术后病理学诊断明确; ② 术前1个月内的超声检查资料完善。排除标准: ① 目标结节超出切面的观察范围; ② 检查时已接受化学药物治疗; ③ 检查前2周内接受穿刺活检。最终入选患者486例, 其中男性108例, 年龄18~73岁, 平均 ( 43.64 ± 12.63 ) 岁; 女性378例, 年龄20~77岁, 平均 ( 43.24 ± 11.35 ) 岁。

对于一侧腺叶内并发多种性质结节的患者, 则每种病理学类型选择最大结节作为研究对象, 共有817个结节纳入本研究, 左侧叶结节331枚, 右侧叶结节448枚, 峡部结节38枚。

### 1.2 仪器与方法

#### 1.2.1 超声检查仪器和方法

采用韩国SAMSUNG公司的RS80A、法国SuperSonic公司的Aixplorer、荷兰Philips公司的

EPI Q7、美国GE公司的Logiq E9等彩色多普勒超声诊断仪, 线阵探头频率为10~15 MHz。调节超声成像频率、成像深度、总增益、深度增益补偿曲线、焦点数目、焦点位置等, 使甲状腺灰阶图像达到最佳成像效果。

术前超声检查由数名甲状腺超声检查经验较丰富(工作经验在5年以上)的超声科医师完成。记录患者年龄、性别、甲状腺疾病相关病史等一般资料, 同时留存各结节的最大横切面、最大纵切面及能体现相关特征的图像资料。

纳入研究的结节图像资料由2名超声科医师(1名为通过超声专业住院医师规范化培训结业考试入职1年的初级职称医师, 1名为具有15年以上甲状腺超声诊断经验的高级职称医师)依据C-TIRADS对结节进行声像特征评估、分类。采用单盲法控制其主观因素的影响, 即2名医师在分析图像时均不知情患者的其他临床资料。

C-TIRADS是基于计数法的分类: 阳性指标(+1分)包括实性、微钙化、极低回声、垂直位以及边缘模糊/边缘不规则或甲状腺外侵犯5个甲状腺结节声像特征; 点状强回声(彗星尾伪像)为阴性指标(-1分)。分类标准: 2类(-1分)、3类(0分)、4a类(1分)、4b类(2分)、4c类(3~4分)和5类(5分)。

### 1.2.2 AI辅诊系统检测结节

采用浙江德尚韵兴医疗科技有限公司的智能辅助诊断系统AI-SONIC Throid(以下简称AI辅诊系统)进行检测。操作时将无任何标记的甲状腺结节灰阶图导入AI辅诊系统, 记录风险概率值。风险概率值范围为0.00~1.00, 其中0.00~0.40为偏良性; 0.41~0.60为可疑恶性; 0.61~1.00为偏恶性。

### 1.3 统计学处理

采用SPSS 22.0软件进行统计学分析, 所有假设检验均为双侧检验,  $P < 0.05$ 为差异有统计学意义。计量资料以 $\bar{x} \pm s$ 表示。计数资料以频数(例数或结节数)表示。正态分布数据采用独立样本 $t$ 检验, 非正态分布数据采用Mann-Whitney  $U$ 检验。以组织病理学结果为金标准, 采用灵敏

度、特异度、准确度、Kappa系数等指标评价各个方案的诊断效能, 率的比较采用Pearson  $\chi^2$ 检验( $n > 40$ 且所有 $T \geq 5$ )或Fisher精确概率检验法检验( $T < 1$ )。

## 2 结果

### 2.1 病理学检查结果

486例患者817个结节中, 良性355个, 其中结节性甲状腺肿265个, 腺瘤样结节45个, 滤泡性腺瘤22个, 桥本结节19个, 亚急性甲状腺炎3个, 纤维性结节1个; 恶性462个, 其中乳头状癌450个, 滤泡癌10个, 髓样癌1个, 低分化癌1个。

### 2.2 C-TIRADS不同年资医师、AI辅诊系统各分类的恶性率

良性组病灶超声测量最大径( $1.10 \pm 1.15$ ) cm, 恶性组病灶最大径( $0.98 \pm 0.82$ ) cm, 差异有统计学意义(两组数据不服从正态分布, 采用两独立样本的Mann-Whitney  $U$ 检验, 结果 $Z = -2.24$ ,  $P = 0.025$ )。

2名医师依据C-TIRAD对817个结节进行分类, 高年资医师诊断结果中, 2、3、4a、4b、4c及5类结节各23、278、76、98、324及18个, 恶性率分别为0.00%、4.32%、42.11%、84.69%、97.84%和100.00%; 低年资医师诊断结果中, 2、3、4a、4b、4c及5类结节各22、190、133、117、349及6个, 恶性率分别为0.00%、5.26%、12.03%、76.07%、97.71%和100.00%。AI辅诊系统诊断结果中, 偏良性结节344个, 可疑恶性12个, 偏恶性461个, 恶性率分别为7.56%、83.33%、92.41%。

### 2.3 C-TIRADS不同年资医师、AI辅诊系统以及两者联合应用的诊断效能对比

以AI评估风险概率值0.41、C-TIRADS 4a类作为良恶性的诊断界值, 以病理学检查结果为金标准, 分别计算AI辅诊系统及不同年资医师应用C-TIRADS对817个结节分类诊断的灵敏度、特异度、准确度、Kappa系数(表1)。与高年资医师单独应用C-TIRADS的诊断结果相

比, AI辅诊系统诊断的灵敏度略低, 特异度较高 ( $P=0.003$ ), 准确度两者相当, 其差异无统计学意义 ( $P=0.151$ ); 与低年资医师单独应用C-TIRADS的诊断结果相比, AI辅诊系统诊断的灵敏度略低 ( $P=0.020$ ), 特异度和特异度较高 ( $P=0.002, P<0.001$ )。

联合应用的总体原则是根据AI判定结果调整C-TIRADS 3和4类结节, 若结节的AI诊断恶性风险概率值 $\geq 0.41$ , 则升一级, 若恶性风险概率值 $< 0.41$ , 则降一级(图1)。联合AI辅诊系统调整分类后诊断的灵敏度、特异度、准确度、Kappa系数详见表1。低年资医师联合AI辅诊系统诊断后的特异度 ( $P<0.001$ ) 及准确度 ( $P<0.001$ ) 均明显提高, 灵敏度差异无统计学意义 ( $P=0.526$ ); 高年资医师联合AI辅诊系统

诊断后的特异度提高 ( $P=0.006$ ), 而准确度、灵敏度差异无统计学意义 ( $P=0.057$ )。

#### 2.4 HT背景对AI辅诊系统、不同年资医师诊断甲状腺结节效能的影响

本研究486例甲状腺结节患者中, 经病理学检查证实合并HT, 同时术前超声甲状腺表现为弥漫性病变的有83例归为HT组, 该组共135个结节, 其中恶性结节79个, 良性结节56个。另403例患者的682个结节归为非HT组, 其中恶性结节383个, 良性结节299个。AI辅诊系统、不同年资医师对HT组及非HT组的诊断灵敏度、特异度、准确度详见表2。比较AI辅诊系统对两组结节诊断的灵敏度、特异度及准确度差异均无统计学意义; 而不同年资医师非HT组的诊断特异度均高于HT组。

表1 C-TIRADS不同年资医师、AI辅诊系统及联合应用的诊断效能

诊断结果	病理学检查		灵敏度/%	特异度/%	准确度/%	Kappa系数
	恶性	良性				
高年资医师			97.40	81.69	90.58	0.805
恶性	450	65				
良性	12	290				
低年资医师			97.84	56.90	80.05	0.574
恶性	452	153				
良性	10	202				
AI辅诊系统			94.37	89.58	92.29	0.843
恶性	436	37				
良性	26	318				
高年资医师联合AI辅诊系统			96.32	89.01	93.15	0.860
恶性	445	39				
良性	17	316				
低年资医师联合AI辅诊系统			97.19	86.76	92.66	0.849
恶性	449	47				
良性	13	308				

AI辅诊系统与高年资医师单独应用C-TIRADS的灵敏度、特异度、准确度相比,  $P$ 值分别为0.007、0.003、0.151; AI辅诊系统与低年资医师单独应用C-TIRADS的灵敏度、特异度、准确度对比,  $P$ 值分别为0.020、0.002、0.000。低年资医师联合诊断前后的灵敏度、特异度、准确度对比,  $P$ 值分别为0.526、0.000、0.000; 高年资医师联合诊断前后的灵敏度、特异度、准确度对比,  $P$ 值分别为0.345、0.006、0.057。

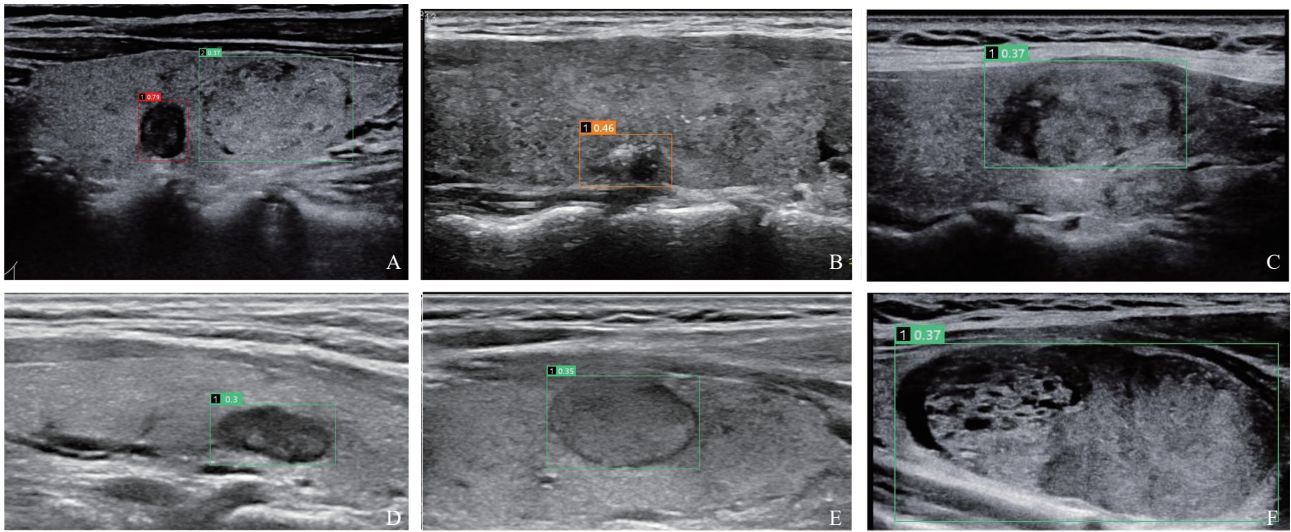


图1 C-TIRADS不同年资医师、AI辅诊系统及联合诊断结果

A: 腺体上部结节为甲状腺乳头状癌, AI辅诊系统评估风险概率值0.79 (红色框表示恶性), 高、低年资医师均诊断为C-TIRADS 4c类; 下部结节为结节性甲状腺肿, AI辅诊系统评估风险概率值0.37 (绿色框表示良性), 高年资医师判为C-TIRADS 3类、低年资医师判为C-TIRADS 4a类 (联合AI调整分类为3类); B: 甲状腺微小乳头状癌, 伴周围甲状腺呈HT改变, AI辅诊系统评估风险概率值0.46 (橙色框表示可疑恶性), 高、低年资医师均诊断为C-TIRADS 4b类; C: HT伴腺瘤样增生性结节形成, AI辅诊系统评估风险概率值0.37, 高、低年资医师分别判为C-TIRADS 4a、4b类 (联合AI调整分类为3、4a类); D: 甲状腺腺瘤, AI辅诊系统评估风险概率值0.3, 高、低年资医师均判为C-TIRADS 4a类 (联合AI调整分类为3类); E: HT伴腺瘤样增生性结节形成, AI辅诊系统评估风险概率值0.35, 高、低年资医师均判为C-TIRADS 4a类 (联合AI调整分类为3类); F: 甲状腺滤泡癌 (弥漫浸润型), AI辅诊系统评估风险概率值0.37, 高、低年资医师均诊断为C-TIRADS 4a类 (联合AI调整分类为3类而造成误诊)。

表2 不同年资医师与AI辅诊系统对HT组、非HT组的诊断结果

分组及诊断结果	病理学检查		灵敏度/%	特异度/%	准确度/%
	恶性	良性			
AI辅诊系统					
HT组			97.47	87.50	93.33
恶性	77	7			
良性	2	49			
非HT组			93.73	89.97	92.08
恶性	359	30			
良性	24	269			
高年资医师					
HT组			100.00	69.64 <sup>*</sup>	87.41
恶性	79	17			
良性	0	39			
非HT组			96.87	83.95 <sup>*</sup>	91.20
恶性	371	48			
良性	12	251			
低年资医师					
HT组			98.73	42.86 <sup>#</sup>	75.56
恶性	78	32			
良性	1	24			
非HT组			97.65	59.53 <sup>#</sup>	80.93
恶性	374	121			
良性	9	178			

AI辅诊系统对HT组与非HT组的诊断灵敏度、特异度、准确度比较, 均 $P>0.05$ 。高年医师对HT组与非HT组诊断情况比较, <sup>\*</sup>:  $P<0.05$ ; 低年医师对HT组与非HT组诊断情况比较, <sup>#</sup>:  $P<0.05$ 。

### 3 讨 论

在临床实践中,超声检查在甲状腺结节随访监测、术前评估、治疗后评估等各个环节均起到不可替代的作用。由于甲状腺结节声像表现的复杂性和多样性,不同国家的研究团队或协会提出了多个版本的TI-RADS分类<sup>[5-8]</sup>,旨在规范化和标准化甲状腺结节的描述和诊断。C-TIRADS是国内第一个由行业协会提出的分类方法,也被证实可以很好地对甲状腺结节进行评估和危险分层<sup>[9]</sup>。本研究纳入的817个结节均可应用C-TIRADS进行分类,分析高年资医师和低年资医师的诊断结果,不同年资医师诊断的灵敏度均较高(均大于97%),低年资医师诊断的特异度(56.90%)及准确度(80.05%)显著低于高年资医师(特异度81.69%、准确度90.58%)。李潜等<sup>[10]</sup>的研究中,医师应用C-TIRADS的诊断灵敏度、特异度及准确度分别为97.37%、47.62%及75.53%,与本研究中低年资医师的诊断结果吻合,即灵敏度高,特异度、准确度较低。另外,本研究低年资医师诊断结果与病理学检查结果一致性较差,Kappa系数0.57,高年资医师诊断结果与病理学检查结果一致性较好,Kappa系数0.81。低年资医师的诊断结果判为4a类的结节133个,其中良性结节117个,对其诊断记录进行分析,结果显示112个结节因呈实性判为4a类,4个结节因边缘模糊判为4a类,1个结节因微钙化判为4a类。虽然C-TIRADS中实性是一个阳性指标,但同时提出可依据医师经验修正分类,低年资医师由于缺乏经验,机械应用C-TIRADS,导致较多呈实性的良性结节被误诊为恶性;而高年资医师依据其丰富的经验分辨出更多的良性结节,因此诊断特异度及准确度高于低年资医师。

在AI研发应用已经上升为国家战略的时代背景下,AI相关技术在医学影像学领域的研究也开展得广泛而深入,在肺部、乳腺等脏器已经进入临床应用阶段<sup>[11-12]</sup>。本研究显示,AI辅诊系统诊断甲状腺结节良恶性的灵敏度(94.37%)低于高、低年资医师,特异度(89.58%)高于高、低

年资医师,准确度(92.29%)优于低年资医师、与高年资医师相当,与病理学检查结果的一致性较好,Kappa系数0.84。当高、低年资医师分别结合AI诊断结果重新对817个结节进行TI-RADS分类,结果显示,高年资医师的诊断特异度及准确度略提高,低年资医师的诊断特异度及准确度明显提高,接近于高年资医师。换言之,低年资医师在AI辅诊系统的辅助下,诊断效能可提高至接近高年资医师的水平,而对高年资医师辅助价值有限。该系统利用智能算法对目标区域进行降噪、增强、细化图像特征等预处理,对甲状腺结节良恶性的诊断准确度较好,弥补了低年资医师因缺乏经验引起的过度诊断,提高了诊断准确度。但是,AI辅诊系统的不足之处在于仅对图像进行分析,而在实际的应用中,医师诊断除了分析图像,尚需结合患者病史、实验室检查等资料进行综合判断。

HT是一种常见的自身免疫性疾病,其病理学改变包括弥漫性淋巴细胞和浆细胞浸润,间质纤维化和腺体萎缩,并伴腺泡嗜酸性退行性变等,导致甲状腺声像表现呈弥漫不均匀或结节样改变<sup>[13]</sup>。在HT背景下结节的良恶性鉴别工作是甲状腺超声检查的难点。本研究中,高、低年资医师对HT组结节的诊断特异度均明显低于非HT组,即HT组中有更大比例的良性结节被误诊为恶性。多项研究<sup>[14-15]</sup>表明,合并HT背景时良性结节呈现边缘不规则或模糊的比例较正常甲状腺背景高,而结节边缘不规则或者模糊是C-TIRADS的恶性指标之一,这使得HT组良性结节声像表现趋向恶性而导致医师诊断假阳性率较高。而本研究中AI辅诊系统对HT组、非HT组的诊断结果对比差异无统计学意义,这得益于该系统所应用的深度学习理论和变分能量泛函结合的图像分析方法,可以较好地处理伴有复杂背景、弱边界的医学图像分割问题,从而保证了AI辅诊系统的稳定性<sup>[16]</sup>。许敏等<sup>[17]</sup>对另一种需要医师手动勾勒甲状腺边缘轮廓的计算机辅诊系统的研究中,结果也显示,在HT背景和非HT背景下,该系统诊断结果的差异无统计学意义,具有较好的诊断稳定性。

对AI辅诊系统误判结节进行分析,我们注意到AI辅诊系统误判为偏良性的结节26个,其中9个为滤泡癌,而本研究入组的结节中共有10个滤泡癌结节,误诊率高达90%(图1F),而高、低年资医师对滤泡癌的误诊率均为40%。滤泡癌声像表现上缺乏恶性征象<sup>[18]</sup>可能是误诊的首要原因,而AI辅诊系统误诊率尤其高可能还与该系统训练时此类病理学类型少见的甲状腺结节数据输入较少有关。

本研究存在一定局限性:①患者图像的采集由多名医师共同完成,可能存在一定的操作者间差异;②本研究样本量不够大,尤其是HT组样本量较小,后续将进一步扩大HT组样本量,深入分析AI辅诊系统对伴HT背景的甲状腺结节良恶性的诊断效能。

综上所述,C-TIRADS对甲状腺结节性质的鉴别具有较高的灵敏度,但对于缺乏经验的低年资医师需加强培训以免出现过度诊断的问题;甲状腺结节AI辅诊系统对甲状腺结节的诊断准确度接近高年资医师,优于低年资医师,且诊断结果不受HT背景的影响,但是单独应用AI辅诊系统作诊断,须警惕AI辅诊系统漏诊恶性结节,尤其是滤泡癌;联合AI辅诊系统对高年资医师应用C-TIRADS的总体诊断准确度影响不显著,但是可提高低年资医师的诊断准确度。

## [参 考 文 献]

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. *CA Cancer J Clin*, 2021, 71(3): 209-249.
- [2] 滕卫平. 我国面临甲状腺癌“过度诊断”的挑战[J]. *中华内分泌代谢杂志*, 2016, 32(7): 531-532.
- [3] 中华医学会超声医学分会浅表器官和血管学组, 中国甲状腺与乳腺超声人工智能联盟. 2020甲状腺结节超声恶性危险分层中国指南: C-TIRADS [J]. *中华超声影像学杂志*, 2021, 30(3): 185-200.
- [4] 王海星, 杨志清, 郭玲玲, 等. 基于大数据和人工智能的超声医学发展现状及问题研究 [J]. *肿瘤影像学*, 2020, 29(4): 410-413.
- [5] KWAK J Y, HAN K H, YOON J H, et al. Thyroid Imaging Reporting and Data System for US features of nodules: a step in establishing better stratification of cancer risk [J]. *Radiology*, 2011, 260(3): 892-899.
- [6] HAUGEN B R, ALEXANDER E K, BIBLE K C, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer [J]. *Thyroid*, 2016, 26(1): 1-133.
- [7] GHARIB H, PAPINI E, GARBER JR, et al. American Association of Clinical Endocrinologists, American College of Endocrinology, and associazione medici endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules--2016 update [J]. *Endocr Pract*, 2016, 22(5): 622-639.
- [8] TESSLER F N, MIDDLETON W D, GRANT E G, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): white paper of the ACR TI-RADS committee [J]. *J Am Coll Radiol*, 2017, 14(5): 587-595.
- [9] 郑琳琳, 李世岩, 许立龙, 等. 甲状腺结节超声恶性风险分层: C-TIRADS、ACR-TIRADS和EU-TIRADS的对比研究 [J]. *中华超声影像学杂志*, 2021, 30(9): 785-791.
- [10] 李 潜, 丁思悦, 郭兰伟, 等. 甲状腺结节超声恶性危险分层中国指南 (C-TIRADS) 联合人工智能辅助诊断对甲状腺结节鉴别诊断的效能评估 [J]. *中华超声影像学杂志*, 2021, 30(3): 231-235.
- [11] 罗 晓, 李安华. 人工智能在乳腺癌诊治中的应用与思考 [J]. *中华医学超声杂志 (电子版)*, 2019, 16(4): 247-251.
- [12] 王 爽, 雷 盛, 谷 涛, 等. 人工智能肺结节辅助诊断软件在肺CT检查结节分析中的临床应用 [J]. *中华肺部疾病杂志 (电子版)*, 2019, 12(6): 757-759.
- [13] 宋艳芳, 田 敏. 桥本甲状腺炎背景下甲状腺结节的超声弹性成像分析 [J]. *中国超声医学杂志*, 2018, 34(11): 979-982.
- [14] 胡 梅, 李明星, 王世界, 等. 伴桥本甲状腺炎的甲状腺良恶性结节: 超声特征及甲状腺超声征象报告与数据系统诊断价值 [J]. *中国医学影像技术*, 2019, 35(6): 828-832.
- [15] WANG D, DU L Y, SUN J W, et al. Evaluation of thyroid nodules with coexistent Hashimoto's thyroiditis according to various ultrasound-based risk stratification systems: a retrospective research [J]. *Eur J Radiol*, 2020, 131: 109059.
- [16] 郭芳琪, 赵佳琦, 陈 蕊, 等. 2.0版人工智能自动检测系统对甲状腺结节术前超声评估效能的初步探讨 [J]. *第二军医大学学报*, 2020, 41(10): 1077-1083.
- [17] 许 敏, 韩 峰, 罗 晓, 等. 计算机辅助诊断系统鉴别甲状腺结节良恶性的诊断效能及其影响因素 [J]. *中华医学超声杂志 (电子版)*, 2019, 16(4): 252-256.
- [18] 张吉臻, 胡 兵. 甲状腺滤泡癌的超声声像图特征 [J]. *中华医学超声杂志 (电子版)*, 2013, 10(6): 484-488.

(收稿日期: 2022-08-09 修回日期: 2022-09-13)