doi 10.12122/j.issn.1674-4500.2026.05.06 分子影像学杂志, 2026, 49(5): 604-610

ACR-TIRADS、C-TIRADS与AI-TIRADS在甲状腺结节良恶性诊断中的应用效果

徐 杰[1],魏 杰[1],范伟健[2],马 芳[1]
合肥市第二人民医院(安徽医科大学附属合肥医院)[1]超声医学科,[2]介入血管疼痛科,安徽 合肥 230011

摘要:目的 对比美国放射学会甲状腺影像报告与数据系统(ACR-TIRADS)、中国甲状腺影像报告与数据系统(C-TIRADS)与人工智能甲状腺影像报告与数据系统(AI-TIRADS)对甲状腺结节良恶性的诊断效能。**方法** 纳入2022年4月~2025年5月合肥市第二人民医院(安徽医科大学附属合肥医院)收治的102例甲状腺结节患者,经病理检查确诊恶性的患者作为恶性组、良性患者作为良性组,比较两组临床资料、各超声征象,将恶性甲状腺结节作为阳性病例,将良性甲状腺结节作为阴性对照,采用ROC曲线分析3种检测方案对甲状腺结节良恶性的诊断效能。**结果** 相较良性组,恶性组结节内部构成实性、回声强度低/极低回声、边界不规则或外扩、形态纵横比≥1占比更高(P<0.05)。经病理确诊恶性甲状腺结节71例,ACR-TIRADS、C-TIRADS与AI-TIRADS正确确诊恶性甲状腺结节分别59例、61例、65例。ROC曲线分析结果显示,ACR-TIRADS、C-TIRADS与AI-TIRADS的曲线下面积(AUC)为0.835(95% *CI*:0.756~0.914,P<0.05)、0.865(95% *CI*:0.793~0.934,P<0.05)、0.909(95% *CI*:0.847~0.972,P<0.05),其敏感度为83.10%、85.92%、91.55%,特异度分别为83.87%、87.10%、90.32%。DeLong检验显示AI-TIRADS的AUC高于ACR-TIRADS(P=0.008,95% *CI*:0.020~0.129)和C-TIRADS(P=0.037,95% *CI*:0.003~0.086),但ACR-TIRADS和C-TIRADS的AUC差异无统计学意义(P=0.110,95% *CI*:-0.007~0.067)。Holm多重比较校正后,AI-TIRADS与ACR-TIRADS的AUC差异仍有统计学意义(校正后P=0.024),而AI-TIRADS与C-TIRADS的差异在校正后呈边缘显著(校正后P=0.074)。ACR-TIRADS与C-TIRADS之间的差异无统计学意义(校正后P=0.110)。**结论** AI-TIRADS对甲状腺结节良恶性的诊断效能相对较高,C-TIRADS临床可操作性强,可优化甲状腺结节的诊疗流程。

关键词:甲状腺结节;美国放射学会甲状腺影像报告与数据系统;中国甲状腺影像报告与数据系统;人工智能甲状腺影像报告与数据系统;受试者工作特征

The application effects of ACR-TIRADS, C-TIRADS and AI-TIRADS in the diagnosis of benign and malignant thyroid nodules

XU Jie[1], WEI Jie[1], FAN Weijian[2], MA Fang[1]
[1]Department of Ultrasound Medicine, [2]Department of Interventional Vascular Pain, Hefei Second People's Hospital, Hefei 230011, China

Abstract: Objective To compare the diagnostic efficiency of the American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TIRADS), the Chinese (C)-TIRADS, and the Artificial Intelligence (AI)-TIRADS for benign and malignant thyroid nodules, and to provide a basis for their clinical application. **Methods** A total of 102 thyroid nodules from 102 patients admitted to Hefei Second People's Hospital from April 2022 to May 2025 were included in the study. After pathological examination, diagnosed as malignant and included in the malignant group, and diagnosed as benign and included in the benign group. The clinical data, various ultrasound signs of the malignant and benign groups were compared. The malignant patients were regarded as positive cases, and the benign ones as negative controls. The ROC curve was used to analyze the diagnostic efficiency of ACR-TIRADS, C-TIRADS, and AI-TIRADS for benign and malignant thyroid nodules. **Results** Compared with the benign group, the malignant group had a higher proportion of solid internal composition, low/very low echo, irregular or expansive borders, and an aspect ratio≥1 (P<0.05). A total of 71 cases of malignant thyroid nodules were diagnosed by pathology. ACR-TIRADS, C-TIRADS and AI-TIRADS correctly diagnosed 59 cases, 61 cases and 65 cases of malignant thyroid nodules respectively. The results of ROC curve analysis showed that the area under the curve (AUC) of ACR-TIRADS, C-TIRADS and AI-TIRADS was 0.835 (95% *CI*: 0.756-0.914, P<0.05), 0.865 (95% *CI*: 0.793-0.934, P<0.05), and 0.909 (95% *CI*: 0.847-0.972, P<0.05), with sensitivities of 83.10%, 85.92%, 91.55% and specificities of 83.87%, 87.10%, 90.32%, respectively. The DeLong test revealed that the AUC area of AI-TIRADS was significantly higher than that of ACR-TIRADS (P=0.008, 95% *CI*: 0.020-0.129) and C-TIRADS (P=0.037, 95% *CI*: 0.003-0.086), but there was no statistically significant difference in the AUC between ACR-TIRADS and C-TIRADS (P=0.110, 95% *CI*: -0.007-0.067). After Holm multiple comparison correction, the difference in AUC between AI-TIRADS and ACR-TIRADS was still statistically significant (adjusted P=0.024), while the difference between AI-TIRADS and C-TIRADS was marginally significant after correction (adjusted P=0.074). There was no statistically significant difference between ACR-TIRADS and C-TIRADS (adjusted P=0.110). **Conclusion** For benign and malignant thyroid nodules, the diagnostic efficiency of AI-TIRADS was higher, C-TIRADS was had strong clinical operability and could optimize the diagnosis and treatment process of thyroid nodules.

Keywords: thyroid nodules; American College of Radiology Thyroid Imaging Reporting and Data System; Chinese Thyroid Imaging Reporting and Data System; Artificial Intelligence Thyroid Imaging Reporting and Data System; receiver operating characteristic

收稿日期:2025-09-16
基金项目:安徽省卫生健康科研项目(AHWJ2022c001);合肥市第二人民医院(安徽医科大学附属合肥医院)科研基金项目(2024ykc046)
作者简介:徐 杰,硕士,主治医师,E-mail: 17756072717@163.com
通信作者:马 芳,主任医师,E-mail: 912869028@qq.com

甲状腺结节作为临床常见疾病，在普通人群中的超声检出率高达19%~68%，其中5%~7%为恶性病变，准确鉴别结节良恶性对临床决策至关重要，可避免不必要的侵入性检查，同时确保恶性肿瘤的及时干预[1, 2]。目前，基于超声特征的甲状腺影像报告与数据系统（TI-RADS）已成为标准化评估工具，但其不同版本在临床应用中的表现仍存在差异[3]。美国放射学会甲状腺影像报告与数据系统（ACR-TIRADS）采用加权评分法，具有较好的标准化程度，但其对亚裔人群的特异性可能受限[4]。中国学者制定的中国甲状腺影像报告与数据系统（C-TIRADS）针对国人流行病学特点进行了优化，特别是对微钙化的判定标准更为严格[5]。同时，人工智能甲状腺影像报告与数据系统（AI-TIRADS）系统通过深度学习技术，在特征提取和模式识别方面展现出独特优势，但其临床可靠性仍需大规模验证。本研究拟同时纳入3种主流分类系统进行直接比较，分析ACR-TIRADS、C-TIRADS与AI-TIRADS对甲状腺结节良恶性的诊断效能，采用严格的质量控制措施，为临床选择最优评估策略提供循证依据，现报道如下。

1 资料与方法

1.1 一般资料

纳入2022年4月~2025年5月合肥市第二人民医院（安徽医科大学附属合肥医院）收治的102例甲状腺结节患者，经病理检查确诊恶性的患者纳入恶性组（n=71），良性患者纳入良性组（n=31）。

纳入标准：接受高频超声检查者；经病理结果确诊甲状腺结节良、恶性者[6]；年龄>18岁；均接受ACR-TIRADS、C-TIRADS及AI-TIRADS 3种分类系统评估者；单发结节或多结节中最可疑结节者；临床资料来源于手术队列且资料完整者等。排除标准：甲状腺结节的超声图像有明显标记而影响判读者；罹患其他恶性肿瘤者；颈部放射史或手术史者；已接受过射频消融、微波消融等介入治疗者；存在严重心、肝、肾等其他重要脏器功能衰竭者等。本研究经合肥市第二人民医院（安徽医科大学附属合肥医院）伦理委员会审批（审批号：2025-科研-075）；本研究为回顾性研究，豁免知情同意。

1.2 方法

1.2.1 超声检查规范 采用西门子ACUSON Sequoia型高端超声诊断系统进行甲状腺检查。设备配置宽频线阵探头，频率范围设置为6~18 MHz，可根据结节深度自动优化分辨率。由具备5年以上甲状腺超声专科技能的医师执行标准化扫查流程：全面探查双侧甲状腺及周围组织，对发现的结节进行多切面（矢状位、横断位）成像，存储包含所有关键超声特征的动态及静态图像，图像分析由2位低年资超声医师采用双盲法（判读超声医师完全不知晓患者的任何临床信息、AI分析结果及最终的病理诊断）独立完成，首先计算2位医师初始独立判读结果的一致性，κ=0.84（95% *CI*：0.73~0.92），一致性较高。对于判读不一致的病例，由双方共同复核图像并协商达成共识。评估内容包括：形态学特征：内部构成（实性、混合性）、回声强度（等高回声、低/极低回声）、边界（光整/模糊、不规则或外扩）、形态（纵横比≥1、纵横比<1）。AI系统为上海交通大学医学院附属瑞金医院超声医学科和复星杏脉联合打造的甲状腺超声AI产品“杏脉•瑞声超影”，产品已实现实时结节检出、分割及恶性概率评估，并嵌入多品牌超声设备，支持基层医疗机构部署，AI判读独立于人工，由AI做最终决定判定。

1.2.2 ACR TI-RADS分类标准[7] 通过以下维度进行量化评估。结构组成：囊性以及海绵状（0分）、囊实混合（1分）、实性（2分）；回声特征：无回声（0分）、等/高回声（1分）、低回声（2分）、极低回声（3分）；生长方位：平行生长（0分）、垂直生长（3分）；边缘特征：光整/模糊（0分）、不规则/分叶（2分）、甲状腺外侵犯（3分）；钙化表现：无/彗尾征（0分）、粗大钙化（1分）、边缘钙化（2分）、微钙化（3分）。阈值设定方面如下：首先以计分方式量化恶性概率，随后将分值映射至不同类别。基础规则中，0分对应1类，判定为良性；2分对应2类，提示倾向良性；3分对应3类，代表低度可疑；4~6分归入4类，视为中度可疑；≥7分则划入5类，表明高度可疑。ACR-TIRADS 5类作为恶性判断的依据。

1.2.3 C-TIRADS分类标准[8] 采用特征累加计分法。恶性征象（+1分/项）：垂直位生长（纵横比≥1）、实性结构、显著低回声、微钙化（≤1 mm点状强回声）、边界不规则/分叶、被膜外侵犯；良性征象（-1分）：彗星尾伪像。风险等级：阈值设定采用负分起点：-1分定义为2类，意为明确良性；0分对应3类，象征可能良性；1分划归4A类，标志低度可疑；2分归入4B类，属于中度可疑；3~4分对应4C类，提示高度可疑；≥5分则升至5类，极高度可疑。C-TIRADS 4B类、4C类、5类作为恶性判断依据。

1.2.4 AI-TIRADS分类标准[9] 以厂家内置分类结果特征权重为标准。结构：实性（1分）；回声：极低（1分）；方位：非平行（1分）；边缘：不规则（1分）/外侵（1分）；钙化：微钙化（1分）。阈值设定方面，无甲状腺结节为1类；2类取-1分，强调明确良性；3类锁定0分，表达可能良性；4A类、4B类、4C类依次对应1分、2分、3~4分，分别呈现低、中、高三级可疑程度；5类则以≥5分为界，代表极高度可疑。AI对最大横切面或者最大纵切面分别判读、最终以任一切面达到4B类、4C类、5类作为恶性判断依据。

AI系统版本号：2.0.24.30，软件最终输出形式为分级，输入切面选择规则包括：成像质量要求：灰阶/彩色

多普勒切面图像清晰、无明显运动伪影/声影遮挡/探头耦合不良，结节实质区域无模糊、过曝或欠曝；聚焦点精准落在结节主体层面，避免近场/远场虚焦。结节显示比例：结节在切面中占有效成像区域的30%~70%，既不缩小导致特征无法识别，也不放大导致解剖定位缺失。解剖定位明确：切面中需清晰显示结节所属甲状腺叶（左/右叶）、极部（上/中/下极）/峡部的标志性解剖结构（气管、甲状腺包膜、颈总动脉/颈内静脉、胸锁乳突肌等），至少1个标志性结构与结节相邻显示，便于系统定位。没有无关遮挡：避免气管、锁骨、胸骨等正常组织的重度声影遮挡结节主体（>1/3结节面积），可通过微调探头角度规避；禁止其他组织（如手指、衣物）入镜。扫查体位基准：所有切面均基于标准仰卧位（患者头后仰，颈下垫枕，充分暴露颈前区），探头轻贴颈前皮肤（避免加压导致结节形态变形，影响纵横比、边界判读）。人工干预（如框选/修正）：框可以干预（可以删除或者添加，以及可以调整框的大小），细分类也可以调整类别，可以由医生根据临床实际和经验进行修改和修正，但对于评级，甲状腺不能人工修改，会根据修改的细分类根据评级准则自动算出。系统判定失败/无法判读分5类核心情形：硬件/数据格式异常、成像质量不达标、切面不合规（缺正交切面/非标准扫查等）、核心特征提取失败、分级结果置信度<70%；处理遵循“先自助后人工、分级推进”原则，分为4级流程：用户端按系统精准提示自助修正重提→系统算法优化辅助复核→超声医生+算法工程师双人工专业复核→极端情况出具正式无法判读报告，建议临床人工综合诊断。针对多发结节、微小结节、紧急临床场景、系统故障等特殊情况做适配处理；所有案例全流程溯源记录，用于算法迭代和规范优化，且系统反馈均为结构化、可视化的具体指引，明确修正要求和操作入口。

1.2.5 病理检测 根据手术切除组织病理检测判断疾病良恶性，细胞病理学的诊断均由同一位病理诊断医生完成。

1.3 统计学分析

采用SPSS26.0统计软件用于数据分析，采用S-W法检验计量资料的正态性，符合正态分布的计量资料以均数±标准差表示，组间比较行独立样本t检验。计数资料以n(%)表示，组间比较行χ^2/校正χ^2检验；以P<0.05为差异有统计学意义。采用ROC曲线分析3种方案的诊断效能，计算曲线下面积（AUC）。AUC的比较通过R4.4.3软件的pROC包实现：采用DeLong检验进行两两比较，随后对得到的3个P值使用Holm-Bonferroni法进行多重比较校正，校正后P<0.05为差异有统计学意义。AUC差值的95%*CI*方向代表比较顺序（如AI *vs* ACR的CI为负，表示AI的AUC小于ACR）；若CI包含0，则差异不显著。根据各系统的既定恶性阈值计算敏感度、特异度、阳性预测值（PPV）、阴性预测值（NPV）及其对应的混淆矩阵（TP、FP、TN、FN）。为更全面评估在类别不均衡情况下的性能，计算以恶性为正类的精确率-召回率曲线下面积（PR-AUC）以及F1分数，计算公式：F1-score=2×PPV×敏感度/(PPV+敏感度)。

2 结果

2.1 恶性组、良性组临床资料比较

两组性别、年龄、结节数量等一般资料的差异无统计学意义（P>0.05，表1）。

表1 恶性组、良性组临床资料比较

Tab.1 Comparison of clinical data between the malignant group and the benign group

Feature		Benign (n=31)	Malignant (n=71)	χ^2/t	P
Gender	Male	21(29.58)	11(35.48)	0.350	0.554
	Female	50(70.42)	20(64.52)		
Age (year)		45.21±5.20	45.16±5.15	0.045	0.964
Nodule number	Single	26(36.62)	14(45.16)	0.660	0.416
	Multiple	45(63.38)	17(54.84)		

2.2 恶性组、良性组超声征象比较

相较良性组，恶性组内部构成实性、回声强度低/极低回声、边界不规则或外扩、形态纵横比≥1占比更高（P<0.05，表2）。

2.3 ACR-TIRADS、C-TIRADS与AI-TIRADS分类结果

经病理确诊恶性甲状腺结节71例，良性31例。ACR-TIRADS、C-TIRADS与AI-TIRADS正确确诊恶性甲状腺结节59、61、65例，良性结节26、27、28例（表3）。

2.4 ACR-TIRADS、C-TIRADS与AI-TIRADS对甲状腺结节良恶性的诊断价值

ROC曲线分析结果显示，ACR-TIRADS、C-TIRADS与AI-TIRADS的曲线下面积（AUC）为0.835（95% *CI*：0.756~0.914，P<0.05）、0.865（95% *CI*：0.793~0.934，P<0.05）、0.909（95% *CI*：0.847~0.972，P<0.05），

表2 恶性组、良性组超声征象

Tab.2 Various ultrasound signs in the malignant group and benign group [*n*(%)]

Feature		Benign (*n*=31)	Malignant (*n*=71)	OR	95% *CI*	χ^2	*P*
Composition	Mixed	8(25.81)	4(5.63)	1		6.627*	0.010
	Solid	23(74.19)	67(94.37)	5.826	1.603-21.172		
Echogenicity	Isoechoic	8(25.81)	5(7.04)	1		5.249*	0.022
	Hypo/markedlyhypo	23(74.19)	66(92.96)	4.591	1.364-15.458		
Margin	Smooth/obscure	24(77.42)	40(56.34)	1		4.103	0.043
	Irregular/extruded	7(22.58)	31(43.66)	2.657	1.014-6.966		
Shape	Anteroposterior/transverse<1	28(90.32)	45(63.38)	1		7.698	0.006
	Anteroposterior/transverse≥1	3(9.68)	26(36.62)	5.393	1.492-19.49		

* Adjusted χ^2.

表3 ACR-TIRADS、C-TIRADS与AI-TIRADS分类结果

Tab.3 Classification results of ACR-TIRADS, C-TIRADS and AI-TIRADS (*n*)

Method	Category	Truth		Total
		Malignant	Benign	
ACR-TIRADS predictive value	Malignant	59	5	64
	Benign	12	26	38
	Total	71	31	102
C-TIRADS predictive value	Malignant	61	4	65
	Benign	10	27	37
	Total	71	31	102
AI-TIRADS predictive value	Malignant	65	3	68
	Benign	6	28	34
	Total	71	31	102

ACR-TIRADS: American College of Radiology Thyroid Imaging Reporting and Data System; C-TIRADS: Chinese Thyroid Imaging Reporting and Data System; AI-TIRADS: Artificial Intelligence-based Thyroid Imaging Reporting and Data System.

其敏感度分别为83.10%、85.92%、91.55%，特异度分别为83.87%、87.10%、90.32%。DeLong检验发现，AI-TIRADS的AUC高于ACR-TIRADS（*Z*=2.669，*P*=0.008，95% *CI*：0.020~0.129）和C-TIRADS（*Z*=2.0882，*P*=0.037，95% *CI*：0.003~0.086），但ACR-TIRADS的AUC和C-TIRADS的AUC差异无统计学意义（*Z*=1.597，*P*=0.110，95% *CI*：-0.007~0.067）。采用Holm多重比较校正后，AI-TIRADS与ACR-TIRADS的AUC差异仍具有统计学意义（校正后*P*=0.024），而AI-TIRADS与C-TIRADS的差异在校正后呈边缘显著（校正后*P*=0.074）。ACR-TIRADS与C-TIRADS之间的差异无统计学意义（校正后*P*=0.110，表4、图1）。

表4 ACR-TIRADS、C-TIRADS与AI-TIRADS对良恶性的诊断价值

Tab.4 Diagnostic value of ACR-TIRADS, C-TIRADS and AI-TIRADS for benign and malignant

Index	AUC	95% *CI*	*P*	Sensitivity (%)	Specificity (%)
ACR-TIRADS	0.835	0.756-0.914	<0.001	83.10	83.87
C-TIRADS	0.865	0.793-0.934	<0.001	85.92	87.10
AI-TIRADS	0.909	0.847-0.972	<0.001	91.55	90.32

结果显示，AI-TIRADS在所有指标上均表现最佳，其PR-AUC和F1-score为三者中最高，提示其在恶性结节比例较高的样本结构中，仍具有更稳定的阳性识别能力和整体分类性能。AI-TIRADS的PPV达到95.59%，NPV为82.35%，均优于ACR-TIRADS和C-TIRADS。相比之下，C-TIRADS的PR-AUC与F1-score略高于ACR-TIRADS，且PPV（93.85%）及NPV（72.97%）亦优于ACR-TIRADS。ACR-TIRADS的PR-AUC与F1-score最低，NPV（68.42%）也相对较低（表5）。

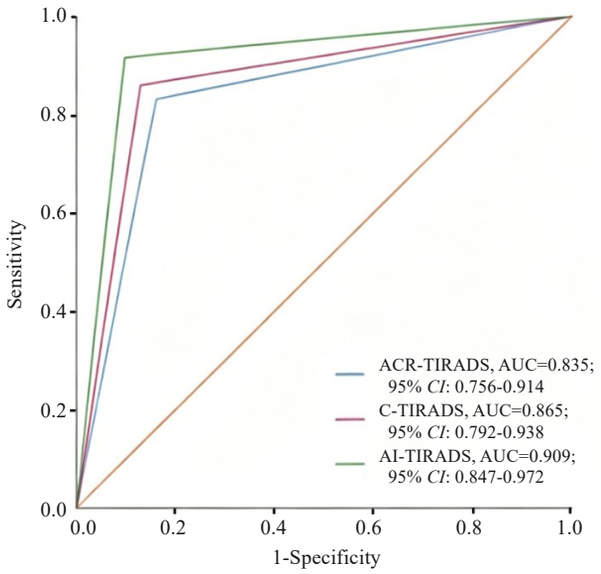

图1　ACR-TIRADS、C-TIRADS与AI-TIRADS诊断良恶性的ROC曲线

Fig. 1　ROC curves of ACR-TIRADS, C-TIRADS and AI-TIRADS for diagnosing benign and malignant.

2.5　甲状腺结节良恶性的超声检查影像学结果

甲状腺恶性结节、甲状腺良性结节最大横切面及最大纵切面超声声像图及AI截图(图2、3)。

3　讨论

甲状腺结节在临床中极为常见,其中大部分为良性,但少数具有恶性潜能。准确鉴别结节的良恶性对于指导临床决策、避免不必要的有创检查和治疗至关重要[10-12]。目前,多种影像学分类系统被用于评估甲状腺结节的恶性风险,包括美国放射学院提出的ACR-TIRADS、中华医学会超声医学分会专家委员会发布的C-TIRADS,以及近年来兴起的基于人工智能技术的AI-TIRADS。这些系统各有特点,但在实际应用中,它们对于甲状腺结节良恶性诊断的效能如何,特别是与其他系统相比是否存在差异,尚需更深入的比较研究[13, 14]。

本研究中,相较良性组,恶性组结节内部构成实性、回声强度低/极低回声、边界不规则或外扩、形态纵横比≥1占比更高,提示超声特征在甲状腺结节良恶性判断中具有重要意义,基于此可通过多种超声分类对甲状腺结节良、恶性进行评估。进一步分析各个系统检出情况,经病理确诊恶性甲状腺结节71例,ACR-TIRADS、C-TIRADS与AI-TIRADS正确确诊恶性甲状腺结节59例、61例、65例,提示3种系统均对甲状腺恶性结节具有一定检出能力。其中ACR-TIRADS作为国际公认的标准化系统,其诊断结果具有跨中心可比性,可作为多中心研究或国际合作项目的统一评估工具,尽管本研究中其效能较低,但对于需要与国际数据对标的场景,仍具有不可替代的参考意义[15, 16]。C-TIRADS的诊断效能略优于ACR-TIRADS,虽然无统计学意义,但C-TIRADS简化了部分评分标准,如对囊性成分的分级更直观,降低了临床操作难度,使其在基层医院的适用性更强[17-19]。

表5　3种TIRADS方法在类别不均衡条件下的诊断效能:PR-AUC、F1、PPV、NPV

Tab.5　Diagnostic efficacy of the three TIRADS methods under the condition of class imbalance: PR-AUC, F1, PPV, NPV

Method	PR-AUC	F1-score	PPV (%)	NPV (%)	TP	FP	TN	FN
ACR-TIRADS	0.900	0.874	92.19	68.42	59	5	26	12
C-TIRADS	0.919	0.919	93.85	72.97	61	4	27	10
AI-TIRADS	0.943	0.943	95.59	82.35	65	3	28	6

PR-AUC: Precision-recall area under the curve; PPV: Positive predictive value; NPV: Negative predictive value; TP: True positive; FP: False positive; TN: True negative; FN: False negative.

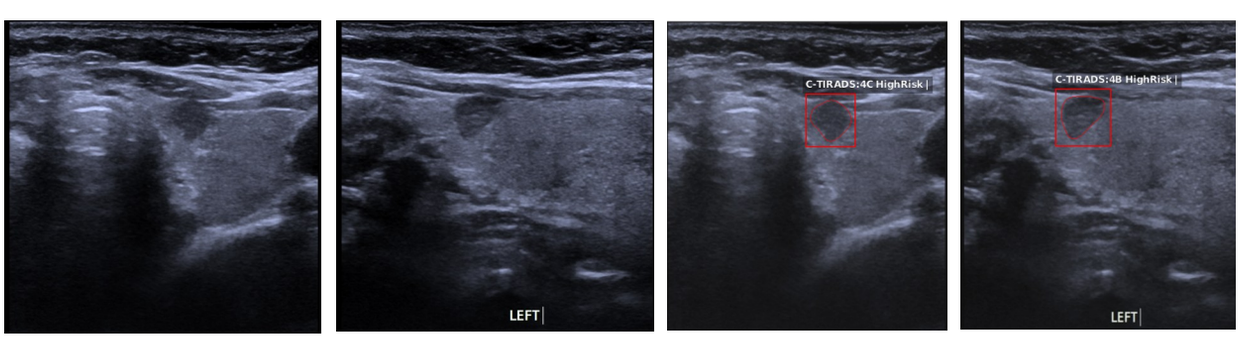

图2　甲状腺乳头状癌最大横切面及最大纵切面超声声像图及AI截图

Fig. 2　The maximum cross-sectional and longitudinal ultrasound images and AI screenshots of papillary thyroid carcinoma. C-TIRADS: 4C: Solid nodule (1), vertical position (1), irregular margin (1), hypoechoic (0), microcalcification (0), total score 3; ACR-TIRADS: 5: Solid nodules (2), vertical position (3), irregular margins (2), hypoechoic (2), microcalcification (0), total score 9; AI-TIRADS: Maximum transverse section 4C, maximum longitudinal section 4B; Pathology: Papillary thyroid carcinoma.

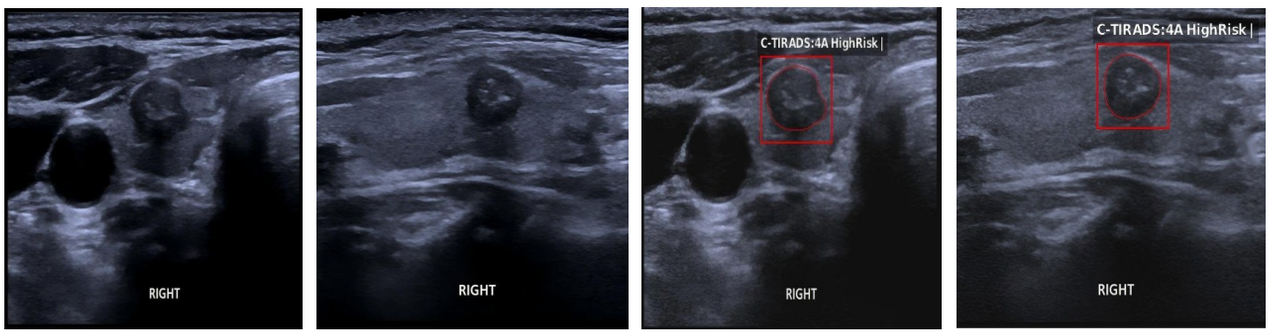

图3 甲状腺良性结节最大横切面及最大纵切面超声声像图及AI截图

Fig.3 The maximum cross-sectional and longitudinal ultrasound images and AI screenshots of benign thyroid nodules. C-TIRADS: 4B: Solid nodule (1), vertical position (0), irregular margin (0), hypoechoic (0), microcalcification (1), total score 2; ACR-TIRADS: 5: Solid nodules (2), vertical position (0), irregular margins (0), hypoechoic (2), microcalcification (3), total score 7; AI-TIRADS: Maximum transverse section 4A, maximum longitudinal section 4A; Pathology: Benign thyroid nodules.

研究显示，ACR-TIRADS和C-TIRADS对老年甲状腺结节均有较高的诊断价值，而C-TIRADS有更高的推荐穿刺率[20]，与本研究结果相互印证。AI-TIRADS的优势源于深度学习算法对超声图像的精准解析能力，可一定程度减少漏诊，尤其对特征不典型的早期恶性结节（如微小乳头状癌）识别能力突出，甲状腺恶性结节的典型特征（如低/极低回声、边界不规则、纵横比≥1等）在超声图像中可能表现为细微信号差异[21, 22]。

C-TIRADS、ACR-TIRADS均采用人工判读，易受主观经验、视觉疲劳等因素影响，导致对微小钙化、边缘模糊等特征的识别率降低[10, 23]。而AI-TIRADS通过训练大量标注数据，可自动提取百余个量化特征，如回声强度梯度、边缘曲率变化等，甚至捕捉人眼难以识别的亚毫米级信号，可一定程度升高恶性结节的敏感度；同时，其算法通过多维度特征融合，减少了单一特征误判，如将良性结节的粗大钙化误判为恶性征象，使特异度同步提高，有助于实现甲状腺癌的早诊早治[24-26]。同时，AI-TIRADS高特异度可降低对良性结节的过度诊断，减少不必要的细针穿刺或手术，减轻患者的负担。这一优势使其在三甲医院等具备AI技术支持的机构中，辅助医生制定更合理的诊疗方案[27, 28]。本研究中，ROC曲线分析结果显示，ACR-TIRADS、C-TIRADS与AI-TIRADS对甲状腺结节良恶性的诊断价值，AUC为0.835、0.865、0.909，其敏感度为83.10%、85.92%、91.55%和特异度分别为83.87%、87.10%、90.32%，AI-TIRADS的AUC高于ACR-TIRADS、C-TIRADS，提示AI-TIRADS对甲状腺结节良恶性的诊断效能相对较高。有研究亦显示，AI-TIRADS对甲状腺结节良恶性的诊断效能高于ACR-TIRADS[29]，与本研究结果一致。有研究分析了3499例具有手术病理和超声特征的甲状腺结节，并采用ACR-TIRADS和AI-TIRADS进行分类，ACR-TIRADS和AI-TIRADS均对半数以上未活检的恶性肿瘤进行了超声监测，AI TIRADS的漏癌率更低[30]，与本研究结果相互印证。

综上，AI-TIRADS对甲状腺结节良恶性的诊断效能相对较高，C-TIRADS临床可操作性强，可优化甲状腺结节的诊疗流程。本研究资料来源于手术队列，可导致的恶性占比过高，高风险结节过多，缺乏看似可疑但实为良性的病例资料，将本研究结果外推到普通筛查人群需谨慎，未来研究可从以下方面拓展：一是开展多中心大样本研究，纳入连续入组的、包含更多不确定结节的队列，进一步验证AI-TIRADS的泛化能力；二是探索多种诊断联合模式，利用AI的精准性与C-TIRADS的实用性，构建更高效的诊断路径。

参考文献：

[1] Joo L, Lee MK, Lee JY, et al. Diagnostic performance of ultrasound-based risk stratification systems for thyroid nodules: a systematic review and meta-analysis[J]. Endocrinol Metab, 2023, 38(1): 117-28.

[2] Kwon D, Kulich M, Mack WJ, et al. Malignancy risk of thyroid nodules that are not classifiable by the American thyroid association ultrasound risk stratification system: a systematic review and meta-analysis[J]. Thyroid, 2023, 33(5): 593-602.

[3] Yeh MW. The molecular landscape of thyroid nodules: diagnostic and prognostic Revelations from a large multicenter cohort[J]. J Clin Endocrinol Metab, 2023, 109(1): e422-3.

[4] Dong WL, Wu YY, Cai T, et al. Comparison of diagnostic performance and FNA management of the ACR-TIRADS and Chinese-TIRADS based on surgical histological evidence[J]. Quant Imaging Med Surg, 2023, 13(3): 1711-22.

[5] Ma Y, Huo XG, Kong SS, et al. A review about C-TIRADS, ACR-TIRADS, and K-TIRADS combined with real-time tissue elastography to diagnose thyroid nodules[J]. Discov Med, 2023, 35 (174): 1-10.

[6] 中华医学会内分泌学分会, 中华医学会外科学分会甲状腺及代谢外科学组, 中国抗癌协会头颈肿瘤专业委员会, 等. 甲状腺结节和分化型甲状腺癌诊治指南(第二版)[J]. 国际内分泌代谢杂志, 2023, 43 (2): 149-94.

[7] Dobruch-Sobczak K, Adamczewski Z, Dedecjus M, et al. Summary

of meta-analyses of studies involving TIRADS classifications (EU-TIRADS, ACR-TIRADS, and K-TIRADS) in evaluating the malignant potential of focal lesions of the thyroid gland[J]. J Ultrason, 2022, 22(89): 121-9.

[8] Li HJ, Yang YP, Liang X, et al. Comparison of the diagnostic performance of three ultrasound thyroid nodule risk stratification systems for follicular thyroid neoplasm: K-TIRADS, ACR-TIRADS and C-TIRADS[J]. Clin Hemorheol Microcirc, 2023, 85(4): 395-406.

[9] Si CF, Fu C, Cui YY, et al. Diagnostic and therapeutic performances of three score-based Thyroid Imaging Reporting and Data Systems after application of equal size thresholds[J]. Quant Imaging Med Surg, 2023, 13(4): 2109-18.

[10] Potipimpanon P, Charakorn N, Hirunwiwatkul P. A comparison of artificial intelligence versus radiologists in the diagnosis of thyroid nodules using ultrasonography: a systematic review and meta-analysis[J]. Eur Arch Otorhinolaryngol, 2022, 279(11): 5363-73.

[11] Kang YJ, Ahn HS, Stybayeva G, et al. Comparison of diagnostic performance of two ultrasound risk stratification systems for thyroid nodules: a systematic review and meta-analysis[J]. Radiol Med, 2023, 128(11): 1407-14.

[12] Yucel S, Balci IG, Tomak L. Diagnostic performance of thyroid nodule risk stratification systems: comparison of ACR-TIRADS, EU-TIRADS, K-TIRADS, and ATA guidelines[J]. Ultrasound Q, 2023, 39(4): 206-11.

[13] Cheong CWS, Au JXY, Lim MY, et al. The efficacy and safety of high-intensity focused ultrasound in the treatment of benign thyroid nodules: a systematic review and meta-analysis from 1990 to 2021 [J]. Ann Acad Med Singap, 2022, 51(2): 101-8.

[14] Nachum S, Tondi Resta I, Baloch Z, et al. Thyroid nodules with indeterminate cytology and negative molecular profile: prevalence of malignancy and practice paradigms for surveillance[J]. Thyroid, 2025, 35(3): 265-73.

[15] Bolland MJ, Grey A. Increased workload without clinical benefit: Results following implementation of the ACR-TIRADS system for thyroid nodules[J]. Clin Endocrinol, 2023, 99(3): 328-34.

[16] Cai MM, Chen LB, Shui LM, et al. Explore the diagnostic performance of 2020 Chinese Thyroid Imaging Reporting and Data Systems by comparing with the 2017 ACR-TIRADS guidelines: a single-center study[J]. Endocrine, 2023, 80(2): 399-407.

[17] Tian JL, Liang JL, Lin YH, et al. Diagnostic performance of ACR-TIRADS combined with superb microvascular imaging for differential diagnosis of mummified thyroid nodules and papillary thyroid carcinomas[J]. Endocr Connect, 2024, 13(3): e230388.

[18] Zhang WB, Deng WF, He BL, et al. Diagnostic value of CEUS combined with C-TIRADS for indeterminate FNA cytological thyroid nodules[J]. Clin Hemorheol Microcirc, 2024, 88(4): 475-83.

[19] Huang WJ, Chen DL, Zhong MY, et al. Correction: factors of central lymph node metastasis in papillary thyroid CancerBased on C-TIRADS analysis[J]. Horm Metab Res, 2023, 55(9): e5.

[20] 刘 牛, 吴意赟, 周正国, 等. ACR-TIRADS和C-TIRADS对老年甲状腺结节诊断价值的对比研究[J]. 实用老年医学, 2025, 39(3): 245-8.

[21] Peng Y, Wang TT, Wang JZ, et al. The application of artificial intelligence in thyroid nodules: a systematic review based on bibliometric analysis[J]. Endocr Metab Immune Disord Drug Targets, 2024, 24(11): 1280-90.

[22] Bojunga J, Trimboli P. Thyroid ultrasound and its ancillary techniques[J]. Rev Endocr Metab Disord, 2024, 25(1): 161-73.

[23] 耿祖仕, 樊玉霞, 卢秀波, 等. 人工智能辅助超声检查在甲状腺结节良恶性判别中的应用[J]. 郑州大学学报: 医学版, 2023, 58(3): 403-6.

[24] Gong ZJ, Xin J, Yin J, et al. Diagnostic value of artificial intelligence-assistant diagnostic system combined with contrast-enhanced ultrasound in thyroid TI-RADS 4 nodules[J]. J Ultrasound Med, 2023, 42(7): 1527-35.

[25] Yan XH, Chen Q, Xin YW, et al. A study on the diagnostic value of artificial intelligence combined with a contrast-enhanced ultrasound scoring system in partially cystic thyroid carcinoma[J]. Front Endocrinol, 2025, 16: 1514185.

[26] He LT, Chen FJ, Zhou DZ, et al. A comparison of the performances of artificial intelligence system and radiologists in the ultrasound diagnosis of thyroid nodules[J]. Curr Med Imaging, 2022, 18(13): 1369-77.

[27] 刘才广, 莫 屈, 肖 羿, 等. 医学影像人工智能在甲状腺癌诊疗中的应用: 现状与展望[J]. 中国普通外科杂志, 2024, 33(11): 1874-82.

[28] Shi XL, Liu RF, Xia Y, et al. Qualitative and quantitative superb vascular imaging in the diagnosis of thyroid nodules ≤10 mm based on the Chinese Thyroid Imaging Reporting and Data System 4 (C-TIRADS 4)[J]. Quant Imaging Med Surg, 2023, 13(5): 3213-21.

[29] 王玉春, 杨 斌, 黄鹏飞, 等. ACR TI-RADS与基于人工智能的报告系统对甲状腺结节的诊断效能及减少不必要穿刺能力的比较[J]. 中华超声影像学杂志, 2021, 30(5): 408-13.

[30] Cui YY, Fu C, Si CF, et al. Analysis and comparison of the malignant thyroid nodules not recommended for biopsy in ACR TIRADS and AI TIRADS with a large sample of surgical series[J]. J Ultrasound Med, 2023, 42(6): 1225-33.

（编辑：熊一凡）