

C-TIRADS在甲状腺结节诊断中的应用研究：与ACR-TIRADS和K-TIRADS的性能对比

魏敬怡¹, 车红霞², 尹鸿涛^{1,3}, 刘丽娟^{1,3}, 贺雯¹, 王梦宇¹, 曾孺宗¹, 李彩英¹, 甄东户^{1,3*}

¹兰州大学第一临床医学院, 甘肃兰州 730000; ²甘肃省第三人民医院内分泌科, 甘肃兰州 730000; ³兰州大学第一医院内分泌科, 甘肃兰州 730000

[中图分类号] R736.1 [文献标志码] A [DOI] 10.11855/j.issn.0577-7402.0662.2025.1013

[声明] 本文所有作者声明无利益冲突

[引用本文] 魏敬怡, 车红霞, 尹鸿涛, 等. C-TIRADS在甲状腺结节诊断中的应用研究:与ACR-TIRADS和K-TIRADS的性能对比[J]. 解放军医学杂志, 2025, 50(12): 1491-1498.

[收稿日期] 2025-04-13 [录用日期] 2025-06-16 [上线日期] 2025-10-13

[摘要] 目的 系统比较中国、美国、韩国不同版本的甲状腺影像报告与数据系统(TIRADS), 即2020版C-TIRADS、2017版ACR-TIRADS及2021版K-TIRADS在中国人群中的诊断效能, 并探讨3种系统在不同大小甲状腺结节亚组中的性能, 为优化临床分层管理策略提供依据。方法 回顾性分析2019年12月—2023年11月于兰州大学第一医院接受甲状腺结节常规手术治疗的1001例患者(共1279个甲状腺结节)的临床资料, 基于患者的术前超声图像, 依据3种系统对结节进行分类。以术后病理结果为金标准, 采用受试者操作特征(ROC)曲线分析计算曲线下面积(AUC)、敏感度及特异度等指标, 并比较不同大小结节亚组[小结节组(≤ 10 mm)、中等结节组(10~20 mm)、大结节组(> 20 mm)]的诊断性能差异。结果 C-TIRADS的总体AUC最大(0.870, 95%CI 0.851~0.888), 与ACR-TIRADS(0.868)差异无统计学意义($P=0.664$), 但均优于K-TIRADS(0.837, $P<0.001$)。亚组分析显示, 小结节组中C-TIRADS的AUC最大(0.782, $P=0.002$), 其敏感度为91.9%; 中等结节亚组中, ACR-TIRADS的AUC较优(0.832, $P=0.003$), 特异度达80.6%; 大结节组中3种系统的差异无统计学意义($P>0.05$)。结论 C-TIRADS与ACR-TIRADS的总体诊断性能更优。C-TIRADS适用于高风险人群的小结节筛查, ACR-TIRADS在中等结节中可减少误诊, 提高诊断准确性, 而大结节的诊断性能需进一步验证。

[关键词] 甲状腺结节; 超声检查; 甲状腺影像报告与数据系统

Diagnostic performance of C-TIRADS in thyroid nodules: a comparative study with ACR-TIRADS and K-TIRADS

Wei Jing-Yi¹, Che Hong-Xia², Yin Hong-Tao^{1,3}, Liu Li-Juan^{1,3}, He Wen¹, Wang Meng-Yu¹, Zeng Ru-Zong¹, Li Cai-Ying¹, Zhen Dong-Hu^{1,3*}

¹First Clinical Medical College, Lanzhou University, Lanzhou, Gansu 730000, China

²Department of Endocrinology, the Third People's Hospital of Gansu Province, Lanzhou, Gansu 730000, China

³Department of Endocrinology, the First Hospital of Lanzhou University, Lanzhou, Gansu 730000, China

*Corresponding author, E-mail: zhdh8279@163.com

This work was supported by the Natural Science Foundation of Gansu Province (25JRRA570, 23JRRA0940, 23JRRA1594), the Gansu Provincial Science and Technology Program (23YFFA0032), the Major Project for Key Chronic Diseases (2023ZD0508402), and the Gansu Provincial Health Industry Research Program (GSWSKY2021-078)

[Abstract] Objective To systematically compare the diagnostic performance of the 2020 version of Chinese Thyroid Imaging Reporting and Data System (C-TIRADS), 2017 version of American College of Radiology Thyroid Imaging Reporting and

[基金项目] 甘肃省自然科学基金(25JRRA570, 23JRRA0940, 23JRRA1594); 甘肃省科技计划项目(23YFFA0032); 四大慢病重大专项(2023ZD0508402); 甘肃省卫生健康行业研究项目(GSWSKY2021-078)

[作者简介] 魏敬怡, 硕士研究生, 主要从事内分泌代谢疾病方面的研究

[通信作者] 甄东户, E-mail: zhdh8279@163.com

Data System (ACR-TIRADS), and 2021 version of Korean Thyroid Imaging Reporting and Data System (K-TIRADS) in a Chinese population and to evaluate their efficacy among different thyroid nodule size subgroups, thereby providing evidence for optimizing clinical risk stratification management. **Methods** Clinical data of 1001 patients (with 1279 thyroid nodules) who underwent conventional thyroid surgery at the First Hospital of Lanzhou University between December 2019 and November 2023 were retrospectively analyzed. Based on preoperative ultrasound images, all nodules were classified according to the three TIRADS (risk stratification systems). Using postoperative histopathology as the gold standard, receiver operating characteristic (ROC) curve analysis was performed to calculate the area under the curve (AUC), sensitivity, specificity, and other metrics. Diagnostic performance was compared among nodule size subgroups: small nodule subgroup (≤ 10 mm), medium nodule subgroup (10-20 mm), and large nodule subgroup (> 20 mm). **Results** C-TIRADS showed the highest overall AUC (0.870, 95%CI 0.851-0.888), which was not significantly different from that of ACR-TIRADS (AUC=0.868, $P=0.664$), but both systems outperformed K-TIRADS (AUC=0.837, $P<0.001$). Subgroup analysis demonstrated that for small nodule subgroup, C-TIRADS achieved the highest AUC (0.782, $P=0.002$) with a sensitivity of 91.9%. For medium nodule subgroup, ACR-TIRADS showed better performance (AUC=0.832, $P=0.003$) with a specificity of 80.6%. No significant differences were observed among the three systems for large nodule subgroup ($P>0.05$). **Conclusion** Both C-TIRADS and ACR-TIRADS demonstrate superior overall diagnostic performance compared to K-TIRADS. C-TIRADS is particularly suitable for screening small nodules in high-risk populations, while ACR-TIRADS helps reduce misdiagnosis and improves diagnostic accuracy for medium-sized nodules. The diagnostic performance for large nodules requires further validation.

[Key words] thyroid nodules; ultrasonography; thyroid imaging reporting and data system (TIRADS)

随着高分辨率超声技术的快速发展和广泛应用及公众健康筛查意识的增强,甲状腺结节的检出率明显增高^[1]。一项Meta分析显示,普通人群中甲状腺结节的患病率达24.8%^[2]。尽管其中约90%为良性结节且无需临床干预,但鉴于甲状腺癌的侵袭性特征仍严重威胁患者的生命健康^[3-4],建立精准的甲状腺结节恶性风险分层体系,对优化诊疗决策并减少过度医疗具有重要的临床价值。超声影像学凭借其无创性、便捷性及高敏感度的特点,已成为甲状腺结节的首选筛查手段^[5]。甲状腺影像报告与数据系统(thyroid imaging reporting and data system, TIRADS)通过分析结节的超声特征及大小,可定量评估其恶性风险,为超声引导下细针穿刺活检(fine needle aspiration, FNA)指征及随访管理提供循证依据^[6-8]。目前主流的TIRADS系统包括中国甲状腺影像报告和数据库系统(Chinese thyroid imaging reporting and data system, C-TIRADS)^[9]、美国放射学会甲状腺影像报告和数据库系统(American College of Radiology thyroid imaging reporting and data system, ACR-TIRADS)^[10]和韩国甲状腺影像报告和数据库系统(Korean thyroid imaging reporting and data system, K-TIRADS)^[11]。然而,在不同地区及人群中的验证研究表明,TIRADS系统的诊断性能和适用性可能因医疗资源、种族特征及研究方法的差异而存在分歧^[12-15]。例如,ACR-TIRADS在人群中的特异度可达58.4%,而K-TIRADS的敏感度最高(94.2%)^[16]。此外,一项比较6种不同系统的真实世界研究显示,C-TIRADS具有最大的受试者操作特征(receiver operating characteristic, ROC)曲线下面积(area under the curve, AUC)(0.891)和最低

的漏诊率(1.2%),但该研究未纳入结节大小这一指标进行分层分析^[17]。基于此,本研究系统比较上述3种TIRADS系统在人群中的诊断效能差异,并进一步探讨其在不同大小结节亚组中的潜在局限性,旨在为甲状腺结节患者提供更精准、个性化的分层管理策略。

1 资料与方法

1.1 研究对象 根据既往文献^[17],假设C-TIRADS、ACR-TIRADS和K-TIRADS的预期AUC分别为0.85、0.80和0.83,临床有意义AUC差异阈值为0.05。预设恶性结节占比40%,通过MedCalc软件估算,需至少600例(恶性240例,良性360例)以检测AUC两两比较的差异($\alpha=0.0167$,效能=80%)。纳入标准:(1)诊断为甲状腺结节且经评估后行手术切除治疗;(2)术前甲状腺灰阶/彩超图像完整且质量合格;(3)术后病理诊断明确。排除标准:(1)妊娠/哺乳期;(2)年龄 <18 岁。最终共纳入2019年12月—2023年11月在兰州大学第一医院行甲状腺结节常规手术治疗的1001例患者(共1279个甲状腺结节)。本研究获兰州大学第一医院伦理委员会审批(LDYLL2025-2015)。

1.2 超声检查与质量控制 所有超声检查由两名具有10年以上经验通过中华医学会超声医学分会甲状腺超声认证的放射科医师完成,设备为Philips EPIQ 5或Mindray Resona R7(高频线阵探头5~15 MHz)。记录的特征包括最大直径、组成(囊性、囊实混合或实性)、回声(无回声、低回声、等回声或高回声)、形态(水平位或垂直位)、边缘(光滑、模糊、分叶、不规则或甲状腺外侵犯)及钙化(彗星尾伪像、粗钙化、

边缘钙化、微钙化)。然后由1名放射科医师及1名内分泌科医师协同依据C-TIRADS、ACR-TIRADS和K-TIRADS对结节进行分类,为控制回顾性数据录入偏倚与重读误差,随机抽取10%原始影像由另一名高级医师复评,若发现分类不一致,由3名医师共同回看图像后协商一致确定。

1.3 指标分析 分析纳入患者的一般临床资料、病理结果分类及超声特征,根据C-TIRADS、ACR-TIRADS和K-TIRADS系统分别对所有1279个甲状腺结节进行分类,以病理结果为金标准,绘制ROC曲线,并通过约登指数(Youden index)(敏感度+特异度-1)最大化原则确定各系统的最佳诊断阈值,进而评估其整体诊断效能。计算指标包括AUC、敏感度、特异度、阳性预测值、阴性预测值和似然比。此外,评估各TIRADS系统分类结果与病理结果之间的一致性(Kappa值)。然后根据结节最大直径(d)将其分为3个亚组:小结节组(≤ 10 mm)、中等结节组(10~20 mm)、大结节组(> 20 mm),并分别比较3个TIRADS系统在各亚组中的诊断性能。

1.4 统计学处理 采用SPSS 26.0及MedCalc 20.0软件进行数据分析。分类变量以例(%)表示,组间比较采用 χ^2 检验或Fisher精确检验。符合正态分布的连续变量以 $\bar{x} \pm s$ 表示,非正态分布变量则以 $M(Q_1, Q_3)$ 表示。采用De-long检验对不同诊断系统的AUC进行两两比较。采用McNemar检验比较敏感度及特异度。 $P < 0.05$ 为差异有统计学意义,亚组分析组内采用Bonferroni法校正 P 值(显著性水平设为 $0.05/9=0.006$)。

2 结果

2.1 甲状腺结节手术患者的基线特征比较 共纳入1001例甲状腺结节患者(男261例,女740例),中位年龄51岁,其中良性病例526例(男129例,女397例),恶性病例475例(男132例,女343例)。恶性结节患者年龄明显低于良性组($P < 0.001$)(表1)。共纳入1279个结节(恶性527个,良性752个),恶性结节中位最大径明显小于良性结节($P < 0.001$)。小结节组(≤ 10 mm)共456个,其中良性147个,恶性309个;中等大小结节组(10~20 mm)共288个,其中良性134个,恶性154个;大结节组(> 20 mm)共535个,其中良性471个,恶性64个(表1)。在恶性结节中,乳头状癌511个(96.9%),滤泡状癌9个(1.7%),髓样癌3个(0.6%),以及低分化/未分化癌4个(0.8%);在良性结节中,甲状腺滤泡性结节病398个(52.9%),滤泡性甲状腺腺瘤260个(34.6%),嗜酸细胞腺瘤1个(0.1%),交界性滤泡性肿瘤29个(3.9%),甲状腺炎57个(7.6%),包括26个慢性淋巴细胞性甲状腺炎、28个桥本甲状腺炎、2个慢性肉芽肿性炎及1个亚急

性甲状腺炎);其他不能分类的病变包括大片纤维钙化、纤维组织增生、不典型腺瘤等7个(0.9%)。

表1 甲状腺结节手术患者的基线特征比较

Tab. 1 Baseline characteristics of patients undergoing thyroid nodule surgery

特征	合计 (n=1001)	良性组 (n=526)	恶性组 (n=475)	χ^2/Z	P
性别[例(%)]				1.380	0.240
女	740(73.9)	397(75.5)	343(72.2)		
男	261(26.1)	129(24.5)	132(27.8)		
年龄[岁, M(Q ₁ , Q ₃)]	51(41, 58)	54(48, 61)	47(36, 53)	-10.695	<0.001
结节特征					
个数	1279	752	527		
直径[mm, M(Q ₁ , Q ₃)]	15(8, 34)	28(12.38, 39)	9(6.2, 13.65)	-17.773	<0.001
最大直径(d) [例(%)]				339.304	<0.001
≤ 10 mm	456(35.6)	147(19.6)	309(58.6)		
10~20 mm	288(22.5)	134(17.8)	154(29.2)		
> 20 mm	535(41.8)	471(62.6)	64(12.2)		

2.2 超声特征与结节组织病理学的相关性 甲状腺良恶性结节在超声形态学特征上呈现出明显的异质性。在成分方面,96.8%的恶性结节为实性,而良性结节中仅51.6%为实性,差异有统计学意义($P < 0.001$);在回声方面,91.1%的恶性结节中为低回声或极低回声,而良性结节中该比例为40.6%,差异有统计学意义($P < 0.001$);在形态方面,42.3%的恶性结节呈垂直位,而良性结节仅为5.2%,差异有统计学意义($P < 0.001$)。此外,53.9%的恶性结节形状不规则或呈分叶状,53.3%伴有微钙化,而良性结节中这两类特征的比例分别为11.7%和9.6%,差异均有统计学意义($P < 0.001$)(表2)。

2.3 三种超声分类系统推荐恶性率与实际病理结果的比较 C-TIRADS中3类、4a类和4b类实际恶性率(分别为3.2%、15.6%和51.5%)高于推荐恶性率,其中仅4a类差异有统计学意义($P < 0.001$)。ACR-TIRADS中2类、3类和4类实际恶性率(分别为3.4%、13.9%和29.3%)高于推荐恶性率,其中3类和4类差异有统计学意义($P < 0.001$),2类差异无统计学意义($P = 0.138$)。K-TIRADS中2类实际恶性率(4.8%)高于推荐恶性率,差异无统计学意义($P = 0.252$)(表3)。

2.4 三种系统的整体诊断效果 ROC曲线分析确定各系统约登指数取得最大值时所对应的分类等级为最佳诊断阈值:C-TIRADS为4b类,ACR-TIRADS为5类,K-TIRADS为5类,对应的约登指数分别为0.626、0.627和0.630(图1)。C-TIRADS、ACR-TIRADS与K-TIRADS的敏感度分别为87.9%、73.2%及83.1%,

表2 甲状腺良恶性结节的超声与组织病理学结果相关特征

Tab.2 Correlated features between ultrasound and histopathological results in benign and malignant thyroid nodules

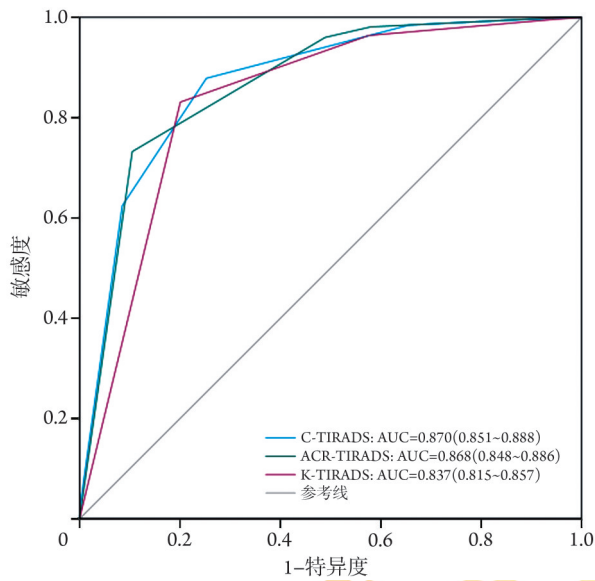
特征	结节总数(n=1279)	良性结节(n=752)	恶性结节(n=527)	Z/ χ^2	P
成分[例(%)]					
囊性或海绵状	41(3.2)	40(5.3)	1(0.2)	26.274	<0.001
囊实混合性	341(26.7)	325(43.2)	16(3.0)	255.862	<0.001
实性	898(70.2)	388(51.6)	510(96.8)	302.385	<0.001
回声[例(%)]					
无回声	40(3.1)	39(5.2)	1(0.2)	25.532	<0.001
等或高回声	454(35.5)	408(54.3)	46(8.7)	280.492	<0.001
低回声	730(57.1)	296(39.4)	434(82.4)	233.757	<0.001
极低回声	55(4.3)	9(1.2)	46(8.7)	42.713	<0.001
形态[例(%)]					
水平位	1017(79.5)	713(94.8)	304(57.7)	262.240	<0.001
垂直位	262(20.5)	39(5.2)	223(42.3)	262.240	<0.001
形状[例(%)]					
光滑	643(50.3)	493(65.6)	150(28.5)	170.559	<0.001
边缘模糊	427(33.4)	213(28.3)	214(40.6)	21.019	<0.001
分叶状或不规则	372(29.1)	88(11.7)	284(53.9)	267.376	<0.001
甲状腺外侵犯	9(0.7)	1(0.1)	8(1.5)	6.640	0.010
钙化[例(%)]					
彗星尾伪像	27(2.1)	25(3.3)	2(0.4)	13.004	<0.001
粗钙化	68(5.3)	37(4.9)	31(5.9)	0.570	0.450
边缘钙化	35(2.7)	24(3.2)	11(2.1)	1.419	0.234
微钙化	353(27.6)	72(9.6)	281(53.3)	296.752	<0.001

表3 三种超声分类系统判别结果与甲状腺结节病理结果的比较

Tab.3 Comparison of three ultrasound risk stratification systems and thyroid nodule pathological outcomes

超声分类	总结节数(个)	病理良性(个)	病理恶性(个)	实际恶性率(%)	推荐恶性率(%)	Z	P
C-TIRADS							
2	19	19	0	0	0	-	-
3	247	239	8	3.2	<2	1.39	0.246
4a	360	304	56	15.6	2~10	3.52	<0.001
4b	260	126	134	51.5	10~50	0.50	0.930
4c	378	63	315	83.3	50~90	-	-
5	15	1	14	93.3	>90	-	-
ACR-TIRADS							
1	29	29	0	0	≤2	-	-
2	297	287	10	3.4	≤2	1.69	0.138
3	79	68	11	13.9	5	3.64	<0.001
4	409	289	120	29.3	5~20	4.72	<0.001
5	465	79	386	83.0	>20	-	-
K-TIRADS							
2	42	40	2	4.8	<3	0.67	0.252
3	296	279	17	5.7	3~10	-	-
4	352	282	70	19.9	10~40	-	-
5	589	151	438	74.4	>60	-	-

C-TIRADS. 中国甲状腺影像报告和数据系统; ACR-TIRADS. 美国放射学会甲状腺影像报告和数据系统; K-TIRADS. 韩国甲状腺影像报告和数据系统; “-”表示未进行此项比较(实际恶性率符合推荐恶性率范围)



C-TIRADS. 中国甲状腺影像报告和数据系统; ACR-TIRADS. 美国放射学会甲状腺影像报告和数据系统; K-TIRADS. 韩国甲状腺影像报告和数据系统

图1 三种系统的ROC曲线

Fig.1 ROC curves of the three risk stratification systems

特异度分别为 74.7%、89.5% 及 79.9%。C-TIRADS 的 AUC 最大(0.870, 95%CI 0.851~0.888), 与 ACR-TIRADS (0.868, 95%CI 0.848~0.886) 差异无统计学意义 ($P>0.05$); 二者均优于 K-TIRADS (AUC=0.837, 95%CI 0.815~0.857, $P<0.001$) (表 4)。

2.5 三种系统在不同大小结节亚组的诊断效果 根据结节最大直径(d)划分为 3 组, 其中, 小结节组共 456 个、中等结节组共 288 个, 大结节组共 535 个。比较 3 种系统对不同大小结节的诊断效能, 在小结节组中, C-TIRADS 的 AUC 值最大 (0.782, 95%CI 0.741~0.819), 与 ACR-TIRADS (0.768, 95%CI 0.726~0.806) 差异无统计学意义 ($Z=0.978, P=0.328$), 优于

K-TIRADS ($Z=2.397, P<0.006$)。在中等结节组中, ACR-TIRADS 的 AUC 最大 (0.832, 95%CI 0.784~0.873), 与 C-TIRADS (AUC=0.824, 95%CI 0.775~0.867) 差异无统计学意义 ($Z=2.516, P=0.012$), 优于 K-TIRADS (AUC=0.782, 95%CI 0.729~0.828) ($Z=2.998, P<0.006$)。在大结节组中, C-TIRADS、ACR-TIRADS 与 K-TIRADS 间 AUC 差异均无统计学意义 ($P>0.006$) (表 5)。

3 讨论

目前, TIRADS 分类系统作为甲状腺结节风险评估的重要工具, 在不同国家和地区的应用中表现出一定的差异。C-TIRADS 作为基于中国临床实践和人群特征制定的分级系统, 与美国和韩国的 TIRADS 在影像特征的量化标准和诊断策略上存在差异, 这些差异可能对甲状腺结节的良恶性判断及分层管理产生影响^[6]。因此, 本研究基于结节的超声特征和不同 TIRADS 系统的诊断表现, 探讨其临床应用价值及优化策略, 旨在为甲状腺结节的精准诊疗提供科学依据。

本研究 752 个良性结节中, 甲状腺滤泡性结节病 (52.9%) 与滤泡性腺瘤 (34.6%) 共同构成主要的病理类型, 与 2022 年 WHO 甲状腺肿瘤分类中滤泡源性病变占主导地位的特征相符^[18]。值得注意的是, 交界性滤泡性肿瘤占比达 3.9%, 此类病变的分子特征与临床处理策略仍需多中心研究进行验证。本研究纳入的 527 个甲状腺恶性病变中, 乳头状癌占比达 96.9%, 略高于全球甲状腺癌流行病学趋势中该病理类型占比 (90.0%), 可能与地域流行病学特征有关^[19]。

甲状腺结节的超声特征分析为良恶性鉴别提供了重要依据。本研究发现, 恶性结节呈现出特征性的超声表型, 绝大多数表现为实性低回声病灶, 其

表 4 三种系统的诊断性能

Tab.4 Diagnostic performance of the three risk stratification systems

项目	C-TIRADS	ACR-TIRADS	K-TIRADS
最佳截距值	4b	5	5
敏感度 [% (95%CI)]	87.9 (84.8~90.5)	73.2 (69.2~77.0)	83.1 (79.6~86.2)
特异度 [% (95%CI)]	74.7 (71.5~77.8)	89.5 (87.1~91.6)	79.9 (76.9~82.7)
阳性预测值 [% (95%CI)]	70.9 (68.2~73.5)	83.0 (79.8~85.8)	74.4 (71.4~77.1)
阴性预测值 [% (95%CI)]	89.8 (87.4~91.7)	82.7 (80.5~84.6)	87.2 (84.8~89.1)
阳性似然比 [% (95%CI)]	3.48 (3.06~3.95)	6.97 (5.62~8.64)	4.14 (3.57~4.80)
阴性似然比 [% (95%CI)]	0.16 (0.13~0.21)	0.30 (0.26~0.35)	0.21 (0.17~0.26)
Kappa 值 (95%CI)	0.60 (0.55~0.66)	0.64 (0.58~0.69)	0.62 (0.56~0.67)
约登指数	0.626	0.627	0.630
AUC (95%CI)	0.870 (0.851~0.888) ⁽¹⁾	0.868 (0.848~0.886) ⁽¹⁾	0.837 (0.815~0.857)

C-TIRADS. 中国甲状腺影像报告和数据系统; ACR-TIRADS. 美国放射学会甲状腺影像报告和数据系统; K-TIRADS. 韩国甲状腺影像报告和数据系统; 95%CI. 95% 置信区间; AUC. 曲线下面积; 与 K-TIRADS 比较, (1) $P<0.001$

表5 三种系统对不同大小甲状腺结节的诊断性能

项目	小结节组 (n=456)						中等结节组 (n=288)						大结节组 (n=535)					
	C-TIRADS		ACR-TIRADS		K-TIRADS		C-TIRADS		ACR-TIRADS		K-TIRADS		C-TIRADS		ACR-TIRADS		K-TIRADS	
	4b	5	5	5	5	5	4b	5	5	5	5	5	4b	5	4	5	5	
最佳截距值	91.9(88.3~94.7)	76.4(71.2~81.0)	88.7(84.6~92.0)	89.0(82.9~93.4)	80.5(73.4~86.5)	84.4(77.7~89.8)	85.6(82.7~88.2)	80.6(72.9~86.9)	82.7(77.0~87.2)	76.0(70.9~80.5)	79.5(72.5~85.1)	79.5(72.1~83.4)	84.2(82.1~86.3)	80.6(72.9~86.9)	82.7(77.0~87.2)	76.0(70.9~80.5)	79.5(72.5~85.1)	79.5(72.1~83.4)
敏感度 [% (95%CI)]	51.7(43.3~60.0)	80.0(77.1~82.6)	75.2(66.9~82.0)	1.90(1.60~2.26)	0.16(0.10~0.24)	0.48(0.39~0.56)	0.436	88.7(84.6~92.0)	57.8(49.4~65.9)	81.5(78.5~84.3)	70.8(63.3~77.4)	2.10(1.73~2.55)	0.20(0.14~0.28)	80.5(73.4~86.5)	84.4(77.7~89.8)	76.0(70.9~80.5)	79.5(72.5~85.1)	79.5(72.1~83.4)
特异性 [% (95%CI)]	80.0(77.1~82.6)	75.2(66.9~82.0)	1.90(1.60~2.26)	0.16(0.10~0.24)	0.48(0.39~0.56)	0.436	0.492	57.8(49.4~65.9)	81.5(78.5~84.3)	70.8(63.3~77.4)	2.10(1.73~2.55)	0.20(0.14~0.28)	80.5(73.4~86.5)	84.4(77.7~89.8)	76.0(70.9~80.5)	79.5(72.5~85.1)	79.5(72.1~83.4)	79.5(72.1~83.4)
阳性预测值 [% (95%CI)]	75.2(66.9~82.0)	1.90(1.60~2.26)	0.16(0.10~0.24)	0.48(0.39~0.56)	0.436	0.492	0.492	73.7(69.0~77.9)	83.3(75.8~88.9)	2.43(1.93~3.06)	0.17(0.11~0.28)	0.53(0.42~0.64)	80.6(72.9~86.9)	82.7(77.0~87.2)	78.3(72.1~83.4)	4.15(2.91~5.91)	0.24(0.17~0.34)	0.61(0.49~0.73)
阴性预测值 [% (95%CI)]	1.90(1.60~2.26)	0.16(0.10~0.24)	0.48(0.39~0.56)	0.436	0.492	0.492	0.492	83.3(75.8~88.9)	2.43(1.93~3.06)	0.17(0.11~0.28)	0.53(0.42~0.64)	0.61(0.49~0.73)	82.7(77.0~87.2)	78.3(72.1~83.4)	4.15(2.91~5.91)	0.24(0.17~0.34)	0.61(0.49~0.73)	0.54(0.43~0.66)
阳性似然比 [% (95%CI)]	0.48(0.39~0.56)	0.436	0.492	0.492	0.492	0.492	0.492	0.53(0.42~0.64)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)
阴性似然比 [% (95%CI)]	0.436	0.492	0.492	0.492	0.492	0.492	0.492	0.53(0.42~0.64)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)
Kappa值 (95%CI)	0.436	0.492	0.492	0.492	0.492	0.492	0.492	0.53(0.42~0.64)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)
约登指数	0.436	0.492	0.492	0.492	0.492	0.492	0.492	0.53(0.42~0.64)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)	0.61(0.49~0.73)	0.54(0.43~0.66)
AUC (95%CI)	0.782(0.741~0.819) ^{(1)*}	0.768(0.726~0.806)	0.741(0.698~0.781)	0.824(0.775~0.867)	0.832(0.784~0.873) ^{(1)*}	0.782(0.729~0.828)	0.806(0.770~0.839)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)	0.801(0.765~0.834)

C-TIRADS. 甲状腺影像报告和数据系统; ACR-TIRADS. 美国放射学会甲状腺影像报告和数据系统; K-TIRADS. 韩国甲状腺影像报告和数据系统; 95%CI. 95%置信区间; AUC. ROC曲线下面积; 与K-TIRADS比较, (1)P<0.006; *表示经Bonferroni校正(校正后 $\alpha=0.006$)仍然显著

中超过90%的病例同时具备低/极低回声和实性成分, 超过50%伴有微钙化及不规则边界等典型恶性特征, 这一特征谱与既往研究报道高度吻合^[20-22], 进一步证实了超声检查在甲状腺结节风险评估中不可替代的作用。值得注意的是, 在形态学特征方面, 恶性结节表现出明显差异, 近50%呈垂直位生长, 而良性结节中这一比例仅为5.2%, 这种差异为临床鉴别诊断提供了重要依据。

本研究最佳截断值均通过约登指数最大化确定, 该方法以敏感度与特异度的均衡优化为目标, 假设漏诊与误诊的临床代价相等, 适用于无预设临床偏好的普适性分析, 可提高诊断效率。而实际场景中二者权重不同需结合成本效应分析来调整阈值。若筛查优先, 可选择牺牲特异度以最大限度减少漏诊; 若以精准诊断为目标, 可选择高特异度以降低假阳性导致的过度活检。ROC曲线分析显示, C-TIRADS在截断值4b时诊断性能最优, 一项Meta分析进一步证实了该结论, 并表明C-TIRADS 4b的诊断性能明显优于4a和4c^[15]。对C-TIRADS、ACR-TIRADS及K-TIRADS 3种系统的两两比较显示, C-TIRADS的AUC值最大, 与ACR-TIRADS无明显差异, 但两者均优于K-TIRADS(P<0.001)。因此, C-TIRADS与ACR-TIRADS在整体诊断性能上接近, 但临床适用场景存在差异, C-TIRADS的敏感度更高, 适用于筛查优先策略, 而ACR-TIRADS特异度高, 则可有效降低误诊风险。一项纳入4238个甲状腺结节的单中心研究对4种指南进行比较, 发现ACR-TIRADS的特异度最高, 与本研究结果一致^[22]。针对活检决策优化, Cai等^[23]比较2020版C-TIRADS、2017版ACR-TIRADS及2015版ATA指南, 证实C-TIRADS可降低不必要的细针穿刺活检(FNA)率, 尤其在直径>20 mm结节中的优势较明显。此外, 一项未纳入C-TIRADS的比较研究显示, ACR-TIRADS的不必要活检率最低^[24], 而其他包含C-TIRADS与ACR-TIRADS比较的研究均报道, C-TIRADS的不必要FNA率较低, 且与ACR-TIRADS差异无统计学意义^[13,25], 提示两者在避免过度活检方面具有可比性。

大部分指南认为, 对<10 mm的结节通常无需积极行FNA, 除非具备明确的高风险特征, 而>20 mm的结节多需行FNA, 因此, TIRADS系统主要在10~20 mm范围内影响结节的临床决策及随访计划^[9-11,26]。本研究通过亚组分析, 比较了3种分类系统在不同大小甲状腺结节中的诊断性能。在小结节组中, C-TIRADS的AUC值最大, 其高敏感度特性尤其适用于高风险人群(如甲状腺癌家族史或辐射暴露史)的筛查, 建议将其作为首选策略以最大程度降低漏诊风险, 但仍需结合患者意愿及随访资源的可及性,

谨慎评估过度诊断的潜在风险。若以避免过度干预为首要目标,可优先采用ACR-TIRADS,但需辅以3~6个月短期超声随访动态监测结节变化。K-TIRADS因AUC值较小且综合诊断性能不足,不推荐用于小结节的分层管理。在中等结节组中,ACR-TIRADS的AUC值最高,其高特异度与可接受的敏感度能够有效平衡漏诊与误诊风险,适合作为初筛工具以优化良性结节的诊断准确性。对于高风险人群,C-TIRADS的高敏感度可作为补充策略,以为恶性病变的识别提供额外的保障,但其特异度较低,可能增加误诊风险,因此需结合其他辅助诊断指标进行评估。有研究显示,与直径 ≥ 20 mm的结节相比,直径10~20 mm的结节FNA推荐可能更多依赖于回声性这一单一的超声特征^[27],在ACR-TIRADS中,回声作为独立赋值指标,可能是其在中等大小结节中整体诊断性能表现较优的原因。在大结节组的3种评分系统AUC无明显差异。综上,C-TIRADS对高风险人群的小结节筛查更具针对性,而ACR-TIRADS在评估中等结节时表现更佳,能有效提升诊断准确性,但其对大结节的评估价值尚需更多证据支持。

本研究存在以下局限性:(1)作为一项单中心回顾性研究,纳入对象均为接受手术的甲状腺结节患者,导致恶性结节比例显著高于普通人群,这种选择偏倚可能高估TIRADS评分对恶性病变的识别能力;(2)受限于病理类型分布特征,纳入的恶性肿瘤以乳头状癌为主,而滤泡癌、髓样癌等亚型样本量不足,未能充分评估各系统对不同病理亚型的鉴别效能;(3)由于仅纳入接受手术的患者,而多数良性结节因体积过大引发压迫症状需手术干预,从而导致大结节组中恶性病例样本量不足(64个),因此,未来仍需联合多中心非手术随访队列进一步验证。

综上所述,本研究系统揭示了C-TIRADS、ACR-TIRADS和K-TIRADS在不同直径甲状腺结节中的差异化诊断价值。对于小结节,C-TIRADS在截断值4b时可优先满足高风险人群的筛查需求;在中等结节的管理中,ACR-TIRADS凭借高特异度可显著优化活检决策的精准性,为降低不必要的医疗干预提供了循证依据;而大结节的诊断效能则受限于当前手术人群的分布特征,需扩大样本的前瞻性队列来避免选择偏倚的影响。这一分层管理框架的提出,不仅为“高危筛查-精准干预”的双路径决策提供了理论支持,也提示K-TIRADS系统在中国人群中的诊断效能欠佳。已有研究表明,结合超声造影、超微血管成像及三维超声成像等多模态参数构建融合诊断体系,可有效弥补单一分类系统的不足^[28-30]。此外,开发整合分子标志物(如BRAF V600E突变)与影像组学特征的动态风险预测模型,将是未来实现甲

状腺结节精准化分层管理的重要突破方向。

【参考文献】

- [1] Grani G, Sponziello M, Filetti S, *et al.* Thyroid nodules: diagnosis and management[J]. *Nat Rev Endocrinol*, 2024, 20(12): 715-728.
- [2] Mu C, Ming X, Tian Y, *et al.* Mapping global epidemiology of thyroid nodules among general population: a systematic review and meta-analysis[J]. *Front Oncol*, 2022, 12: 1029926.
- [3] Erdoğan MF, Bahçecioglu AB, Avcı Merdin F, *et al.* Changes in the volume and diameter of benign thyroid nodules: a 10-year follow-up study[J]. *Thyroid*, 2025, 35(1): 31-40.
- [4] Studen KB, Domagala B, Gabersček S, *et al.* Diagnosing and management of thyroid nodules and goiter—current perspectives[J]. *Endocrine*, 2025, 87(1): 39-47.
- [5] Bekker J, Wreesmann VB. The growing precision of diagnostic pathways in thyroid nodule evaluation[J]. *Eur J Surg Oncol*, 2025, 51(1): 109374.
- [6] Sipos JA. The history of thyroid ultrasound: past, present, and future directions[J]. *Endocr Pract*, 2024, 30(12): 1220-1226.
- [7] Mann H, Arroyo N, Hsiao V, *et al.* Temporal trends in thyroid nodule size on ultrasonography: a systematic review and meta-analysis[J]. *JAMA Otolaryngol Head Neck Surg*, 2025, 151(1): 47-55.
- [8] Kobaly K, Kim CS, Mandel SJ. Contemporary management of thyroid nodules[J]. *Annu Rev Med*, 2022, 73: 517-528.
- [9] 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.
- [10] Tessler FN, Middleton WD, Grant EG, *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.
- [11] Ha EJ, Chung SR, Na DG, *et al.* 2021 Korean thyroid imaging reporting and data system and imaging-based management of thyroid nodules: Korean society of thyroid radiology consensus statement and recommendations[J]. *Korean J Radiol*, 2021, 22(12): 2094-2123.
- [12] Piticchio T, Russ G, Radzina M, *et al.* Head-to-head comparison of American, European, and Asian TIRADSs in thyroid nodule assessment: systematic review and meta-analysis[J]. *Eur Thyroid J*, 2024, 13(2): e230242.
- [13] Jin Z, Pei S, Shen H, *et al.* Comparative study of C-TIRADS, ACR-TIRADS, and EU-TIRADS for diagnosis and management of thyroid nodules[J]. *Acad Radiol*, 2023, 30(10): 2181-2191.
- [14] 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): 1407-1414.
- [15] Yang L, Li C, Chen Z, *et al.* Diagnostic efficiency among Eu-/C-/ACR-TIRADS and S-detect for thyroid nodules: a systematic review and network meta-analysis[J]. *Front Endocrinol (Lausanne)*, 2023, 14: 1227339.
- [16] Tobcu E, Karavaş E, Yılmaz GT, *et al.* Comparison of K-TIRADS, EU-TIRADS and ACR-TIRADS guidelines for malignancy risk determination of thyroid nodules[J]. *Diagnostics (Basel)*, 2025, 15(8): 1015.

- [17] Topcuoglu OM, Uzunoglu B, Orhan T, *et al.* A real-world comparison of the diagnostic performances of six different TIRADS guidelines, including ACR-/Kwak-/K-/EU-/ATA-/C-TIRADS[J]. *Clin Imaging*, 2025, 117: 110366.
- [18] Baloch ZW, Asa SL, Barletta JA, *et al.* Overview of the 2022 WHO classification of thyroid neoplasms[J]. *Endocr Pathol*, 2022, 33(1): 27-63.
- [19] Kitahara CM, Schneider AB. Epidemiology of thyroid cancer[J]. *Cancer Epidemiol Biomarkers Prev*, 2022, 31(7): 1284-1297.
- [20] Chen Q, Lin M, Wu S. Validating and comparing C-TIRADS, K-TIRADS and ACR-TIRADS in stratifying the malignancy risk of thyroid nodules[J]. *Front Endocrinol(Lausanne)*, 2022, 13: 899575.
- [21] Lee JY, Lee CY, Hwang I, *et al.* Malignancy risk stratification of thyroid nodules according to echotexture and degree of hypoechogenicity: a retrospective multicenter validation study[J]. *Sci Rep*, 2022, 12(1): 16587.
- [22] Özdemir M, Türk G, Bilgili M, *et al.* Comparison of diagnostic performances of ATA guidelines, ACR-TIRADS, and EU-TIRADS and modified K-TIRADS: a single center study of 4238 thyroid nodules[J]. *Exp Clin Endocrinol Diabetes*, 2025, 133(2): 98-104.
- [23] Cai Y, Yang R, Yang S, *et al.* Comparison of the C-TIRADS, ACR-TIRADS, and ATA guidelines in malignancy risk stratification of thyroid nodules[J]. *Quant Imaging Med Surg*, 2023, 13(7): 4514-4525.
- [24] 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): 206-211.
- [25] Dong W, Wu Y, 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): 1711-1722.
- [26] Russ G, Bonnema SJ, Erdogan MF, *et al.* European thyroid association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS[J]. *Eur Thyroid J*, 2017, 6(5): 225-237.
- [27] Rucz K, Hegedús L, Bonnema SJ, *et al.* Echogenicity as a standalone nodule characteristic is not inferior to the TIRADS systems in the 10-20 mm nodule diameter range in patient selection for fine needle aspiration: a pilot study[J]. *Eur Thyroid J*, 2024, 13(6): e240149.
- [28] Wang Z, Huang SS, Zhu YF, *et al.* A new thyroid imaging reporting and data system for nodules: Based on grayscale and color Doppler ultrasonography[J]. *Eur J Radiol*, 2025, 183: 111866.
- [29] Wang X, Niu Y, Liu H, *et al.* ThyroNet-X4 genesis: an advanced deep learning model for auxiliary diagnosis of thyroid nodules' malignancy[J]. *Sci Rep*, 2025, 15(1): 4214.
- [30] Liu R, Jiang Y, Lai X, *et al.* Omniview of three-dimensional ultrasound for prospective evaluation of extrathyroidal extension of differentiated thyroid cancer[J]. *BMC Med Imaging*, 2025, 25(1): 42.

(责任编辑: 张小利)

解放军医学杂志®