

doi: 10.3978/j.issn.2095-6959.2021.07.012

View this article at: <https://dx.doi.org/10.3978/j.issn.2095-6959.2021.07.012>

ATA 指南、Kwak 与 ACR TI-RADS 分类对 甲状腺结节的诊断效能

唐海玲, 刘千琪, 王建新, 邢萍, 巴黎, 吴长君

(哈尔滨医科大学附属第一医院超声科, 哈尔滨 150001)

[摘要] 目的: 对比2011 Kwak版甲状腺影像报告和数据系统(Thyroid Imaging Reporting and Data System by Kwak, Kwak TI-RADS)、2015 ATA《成人甲状腺结节与分化型甲状腺癌指南》(ATA指南)及2017 ACR版甲状腺影像报告和数据系统(American College of Radiology Thyroid Imaging Reporting and Data System, ACR TI-RADS)3种方法对甲状腺良恶性结节的诊断效能。方法: 回顾性分析265个甲状腺结节的超声特征, 应用以上3种方法对结节进行诊断分级, 以病理结果为金标准, 构建受试者工作特征曲线(receiver operating characteristic, ROC)。结果: Kwak TI-RADS、ATA指南与ACR TI-RADS的ROC曲线下面积(the area under the ROC curve, AUC)分别为0.914、0.829、0.910, 三者对甲状腺结节均具有较高的诊断效能, 其中Kwak TI-RADS与ACR TI-RADS诊断效能更高于ATA指南方法($P < 0.05$), Kwak TI-RADS与ACR TI-RADS分级方法的诊断效能间无显著差异($P > 0.05$); 三者的最佳诊断值分别为Kwak TI-RADS 4c, 高度可疑恶性, ACR TI-RADS 5。Kwak TI-RADS诊断恶性结节的准确率高于另外2种方法($P < 0.05$), ACR TI-RADS的特异度及阳性预测值较高、敏感度及阴性预测值较低, ATA指南对部分甲状腺结节无法进行明确分类。结论: Kwak TI-RADS、ATA指南及ACR TI-RADS对甲状腺结节的诊断均具有较好的诊断效能, 其中Kwak TI-RADS与ACR TI-RADS诊断效能更高于ATA指南。

[关键词] 甲状腺结节; 超声检查; 诊断效能

Diagnostic efficacy of the ATA guidelines, the Kwak TI-RADS and the ACR TI-RADS for thyroid nodules

TANG Hailing, LIU Qianqi, WANG Jianxin, XING Ping, BA Li, WU Changjun

(Department of Ultrasound, First Affiliated Hospital of Harbin Medical University, Harbin 150001, China)

Abstract **Objective:** To compare the diagnostic efficacy of the 2011 thyroid imaging reporting and data system by Kwak (Kwak TI-RADS), 2015 American Thyroid Association (ATA) management guidelines and 2017 American College of Radiology thyroid imaging reporting and data system (ACR TI-RADS) for benign and malignant thyroid nodules. **Methods:** The conventional ultrasound features of 265 thyroid nodules

收稿日期 (Date of reception): 2020-05-29

通信作者 (Corresponding author): 吴长君, Email: uschangjunwu@sina.com

were retrospectively analyzed and the above three methods were applied to diagnose and grade the nodules. Based on the gold standard of pathological results, the receiver operating characteristic curve (ROC) was constructed. **Results:** The area under the ROC curve of Kwak TI-RADS, ATA guidelines grading and ACR TI-RADS classification was respectively 0.914, 0.829, and 0.910, and they all had high diagnostic efficacy for thyroid nodules, among which the diagnostic efficacy of Kwak TI-RADS and ACR TI-RADS was higher than that of ATA guidelines ($P < 0.05$), with no significant difference between Kwak TI-RADS and ACR TI-RADS ($P > 0.05$). The optimal diagnostic value of these three methods was respectively Kwak TI-RADS 4c, highly suspected malignancy of ATA guidelines and ACR TI-RADS 5. The accuracy of Kwak TI-RADS in the diagnosis of malignant nodules was higher than that of the other two ($P < 0.05$). The specificity and positive predictive value of ACR TI-RADS were higher, and its sensitivity and negative predictive value were lower than those of the other two. Therefore, some thyroid nodules could not be clearly classified by ATA guidelines. **Conclusion:** The three methods have good efficacy in the diagnosis of thyroid nodule, and the diagnostic efficacy of Kwak TI-RADS and ACR TI-RADS was higher than that of ATA guidelines.

Keywords thyroid nodule; ultrasonography; diagnostic efficacy

目前, 较为常用的甲状腺结节超声检查的临床指南主要为2011 Kwak版甲状腺影像报告和数据系统(Thyroid Imaging Reporting and Data System by Kwak, Kwak TI-RADS), 2015 ATA《成人甲状腺结节与分化型甲状腺癌指南》(ATA指南)及2017 ACR版甲状腺影像报告和数据系统(American College of Radiology Thyroid Imaging Reporting and Data System, ACR TI-RADS), 本研究旨在对比3种临床指南对甲状腺良恶性结节的诊断效能, 并确定各个临床指南中鉴别良恶性结节的最佳诊断截点。

1 对象与方法

1.1 对象

回顾性分析2017年12月至2019年5月于哈尔滨医科大学附属第一医院就诊的236例患者, 共265个甲状腺结节, 男64例, 女172例, 年龄20~69(45.8±10.3)岁。所有结节均经手术或甲状腺细针抽吸并通过细胞学检查获得病理结果。纳入标准: 1)术前行超声检查, 有完整的静态及动态图像。2)有病理结果。排除标准: 纯囊性结节。

1.2 仪器与方法

应用西门子ACUSON S3000彩色多普勒超声诊断仪, 探头频率4~9 MHz。回顾性分析265个甲状腺结节二维横切面、纵切面及动态图像, 观察并记录结节各项超声特征, 分别应用Kwak TI-RADS、ATA指南及ACR TI-RADS方法进行诊断分析。

Kwak TI-RADS分级标准^[1]: 5个可疑超声征象包括实性、低回声或极低回声、边缘不规则、纵横比>1及微钙化。TI-RADS 1级: 正常甲状腺; TI-RADS 2级: 良性, 胶质结节(囊性结节, 海绵样结节, 囊实性等回声结节, 合并点状强回声); TI-RADS 3级: 无可疑超声特征, 桥本结节(桥本背景下, 高回声、等回声或低回声, 部分有包膜, 周围血流); TI-RADS 4a级: 1个可疑超声特征; TI-RADS 4b级: 2个可疑超声特征; TI-RADS 4c级: 3个或4个可疑超声特征; TI-RADS 5级: 5个可疑超声特征。

ATA指南分级标准^[2]: 1)高度可疑型: 实性低回声或实性部分为低回声的囊实性结节, 伴有以下一项或多项可疑征象, 包括不规则边界、微钙化、纵横比>1、环形钙化伴局部软组织破坏、腺体外侵犯。2)中度可疑型: 边界清晰的实性低回声结节, 不伴有微钙化、纵横比>1及腺体外侵犯等可疑特征。3)低度可疑型: 边界清晰的等回声或高回声结节, 不伴有微钙化、纵横比>1及腺体外侵犯等可疑特征。4)极低度可疑型: 海绵状结节或不具有上述可疑特征的囊实性结节。5)良性: 纯囊性结节。

ACR TI-RADS分类标准^[3]: 根据结节的所有的超声特征进行评分。成分: 囊性或几乎囊性0分, 海绵样0分, 囊实混合性1分, 实性或几乎实性2分。回声: 无回声0分, 高回声或等回声1分, 低回声2分, 极低回声3分。形态: 纵横比>1为3分, 纵横比≤1为0分。边缘: 光滑0分, 模糊0分, 分叶或不规则2分, 向甲状腺外侵犯3分。强回声:

无或大彗尾0分, 粗钙化1分, 周边型钙化2分, 点状强回声3分。所有超声特征评分的总和确定ACR TI-RADS分级, 包括TI-RADS 5 \geq 7分, TI-RADS 4为4~6分, TI-RADS 3为3分, TI-RADS 2为2分, TI-RADS 1为1分。

1.3 统计学处理

应用SPSS 19.0统计软件进行数据分析, 年龄采用均数 \pm 标准差($\bar{x}\pm s$)表示; 以病理结果为金标准, 构建受试者工作特征(receiver operating characteristic, ROC)曲线, 计算曲线下面积(the area under the ROC curve, AUC), 采用Z检验比较各AUC值间有无差异; 确定3种临床指南的最佳诊断界值点, 采用Z检验比较各个最佳界值点诊断的敏感度、特异度、阳性预测值、阴性预测值及准确率; $P<0.05$ 为差异有统计学意义。

2 结果

2.1 病理结果

265个甲状腺结节中, 132个为恶性, 病理类型分别为127个乳头状癌、3个滤泡状癌、2个髓样癌; 133个为良性, 病理类型分别为118个结节性

甲状腺肿、15个甲状腺腺瘤。

2.2 Kwak TI-RADS、ATA 指南与 ACR TI-RADS 分级诊断效能的比较

应用3种临床指南对265个结节进行诊断分析, 其分布情况及各个级别中恶性率情况见表1。Kwak TI-RADS、ATA指南与ACR TI-RADS的AUC分别为0.914、0.829、0.910, 三者对甲状腺良性结节的诊断均具有较好的诊断效能。其中Kwak TI-RADS与ACR TI-RADS对甲状腺结节诊断效能高于ATA指南分级, 差异有统计学意义($P<0.05$), Kwak TI-RADS与ACR TI-RADS的诊断效能无显著差异($P>0.05$)。以约登指数最大时对应的点为临界点, 3种指南的最佳诊断值分别为TI-RADS 4c, ATA指南的高度可疑恶性及TI-RADS 5, 比较得出, Kwak TI-RADS诊断恶性结节的准确率高于另外2种指南方法, ACR TI-RADS的特异性和阳性预测值高于前2者, 但敏感性和阴性预测值则低于前两者, 差异均有统计学意义($P<0.05$, 表2), ATA指南方法对21个甲状腺结节无法明确分类(表3), 应用Kwak TI-RADS与ACR TI-RADS方法进行诊断分析, 计算诊断的准确率分别为66.7%和71.4%。

表1 甲状腺良恶性结节在三种临床指南分类中的分布情况

Table 1 Distribution of benign and malignant thyroid nodules in three clinical guidelines

指南	分类	肿瘤/个		恶性率/%	参考恶性率/%
		良性	恶性		
Kwak	2	1	0	0.00	0.0
	3	12	0	0.00	1.7
	4a	67	4	5.63	3.3
	4b	36	8	18.18	9.2
	4c	17	105	86.07	44.4~72.4
	5	0	15	100.00	87.5
ATA	极低度可疑	4	0	0.00	<3.0
	低度可疑	76	3	3.80	5.0~10.0
	中度可疑	28	3	12.50	10.0~20.0
	高度可疑	17	113	86.80	70.0~90.0
	无法分类	8	13	61.90	
ACR	TR2	12	0	0.00	<2.0
	TR3	67	2	2.90	<5.0
	TR4	49	40	44.94	5.0~20.0
	TR5	5	90	94.73	\geq 20.0

表2 3种临床指南诊断甲状腺良性恶性结节效能对比

Table 2 Comparison of efficacy of three clinical guidelines in the diagnosis of benign and malignant thyroid nodules

	AUC	敏感度	特异度	阳性预测值	阴性预测值	准确率
分类						
4c ^a	0.914	0.909	0.872	0.876	0.906	0.891
5R ^b	0.910	0.682	0.962	0.947	0.753	0.823
高度可疑 ^c	0.829	0.848	0.812	0.861	0.947	0.830
<i>P</i>						
<i>P</i> ^{ab}	0.875	<0.001	0.008	0.068	0.001	0.026
<i>P</i> ^{ac}	0.014	0.131	0.179	0.728	0.224	0.045
<i>P</i> ^{bc}	0.016	0.001	<0.001	0.036	<0.001	0.819

P<0.05为差异有统计学意义。

P<0.05 was considered statistically significant.

表3 21个ATA分级无法明确分类的结节在Kwak及ACR分级中的分布情况

Table 3 Distribution of 21 nodes in Kwak TI-RADS and ACR TI-RADS that could not be clearly classified by the ATA guidelines

病例号	成分	回声	纵横比	边缘	微钙化	粗钙化及 环形钙化	腺体外侵犯	Kwak	ACR	
									分值	分类
1	实性	等	≤1	不规则	无	无	无	4b	5	TR4
2	囊实性	混合	≤1	不规则	有	无	有侵及被膜	4b	8	TR5
3	实性	等	>1	不规则	无	无	有侵及被膜	4c	9	TR5
4	实性	等	>1	不规则	无	无	有侵及被膜	4c	9	TR5
5	实性	等	≤1	光整	有	无	无	4b	6	TR4
6	实性	等	≤1	不规则	有	无	无	4c	8	TR5
7	囊实性	混合	≤1	光整	有	无	无	4a	5	TR4
8	实性	等	≤1	模糊	有	有	无	4c	6	TR4
9	囊实性	混合	≤1	不规则	有	无	有淋巴结转移	4b	8	TR5
10	实性	等	≤1	小分叶	有	无	无	4c	8	TR5
11	实性	等	≤1	不规则	有	无	无	4c	8	TR5
12	实性	等	≤1	不规则	无	无	无	4b	5	TR4
13	实性	等	≤1	不规则	有	无	无	4c	8	TR5
14	囊实性	混合	≤1	不规则	有	无	无	4b	5	TR4
15	实性	等	>1	不规则	无	无	无	4c	8	TR5
16	实性	等	≤1	不规则	无	无	无	4b	6	TR4
17	实性	等	>1	光整	无	无	无	4b	6	TR4
18	实性	等	≤1	不规则	无	无	无	4b	6	TR4
19	实性	等	≤1	不规则	无	无	无	4b	6	TR4
20	实性	等	≤1	不规则	无	无	无	4b	5	TR4
21	实性	等	≤1	不规则	无	无	无	4b	5	TR4

3 讨论

本研究得出: 3种临床指南对甲状腺结节均具有较高的诊断效能(AUC分别为Kwak 0.914, ATA 0.829, ACR 0.910), 其中Kwak TI-RADS及ACR TI-RADS分级方法的诊断效能更高于ATA指南, 本结论与以往的大部分研究结果一致^[4-7], 但与史宜鑫等^[8]的研究结果不同, 该研究得出ACR TI-RADS诊断恶性结节的AUC为0.761, 且最佳诊断值为TR4类。分析2组实验得出: 该研究的样本量大于本研究, 且被分为TR5的良性结节所占占比显著高于本研究(144/600 vs 5/95), 推测该研究样本中具有可疑恶性征象的良性样本比例大于本研究, 致使二者结果的不同。研究^[5]将4b类作为恶性结节的诊断标准, 本研究得出4b类的恶性风险约为18.18%, 若将4b作为恶性诊断标准, 虽漏诊率减低, 但误诊率则会显著增高(敏感度和特异度分别为96.9%和60.0%)。3种指南方法各具优缺点, ATA指南因对部分结节无法明确分类, 使其总体的诊断效能略低于Kwak TI-RADS与ACR TI-RADS; Kwak TI-RADS虽只是将可疑恶性征象简单叠加, 没有赋予不同的权重, 但其诊断恶性结节的准确率高与另外二者; ACR TI-RADS将其各个可疑征象赋予不同的分值, 最终得出的诊断敏感性和阴性预测值则相抵较低, 但特异性和阳性预测值较高。

本研究中, Kwak TI-RADS各个级别的结节恶性率均略高于初始研究^[1], ACR TI-RADS与ATA分类中, 除ACR TI-RADS 4类结节恶性率显著高于初始研究外, 其余分级的恶性率与初始研究基本符合^[2-3]。分析部分分级恶性率较高的原因, 可能是研究样本的良恶性结节所占占比不同。Kwak及ACR TI-RADS方法的初始研究中, 良性结节所占占比均大于恶性结节^[1,3](前者1 383 vs 275, 后者3 070 vs 352), 而本研究良恶性结节所占比例相当(133 vs 132)。ACR TI-RADS 4R差异最为显著, 在本研究中其恶性率为44.9%, 研究^[6]得出ACR TI-RADS 4类结节恶性风险为35.26%, 均高于初始研究^[3]中5%~20%的参考范围。本研究中TR4类结节共89个, 其中评分为6分的结节所占比例为55.0%(49/89), 其中恶性结节占67.3%(33/49); 评为4分的结节占比为36.0%(32/89), 其中恶性结节仅占6.3%(2/32)。因此, 样本量中良性结节所占占比对TR4类结节恶性率有一定的影响, 并且评为TR4类的结节具有良和恶2种可能性, 相比之下评为4分或6分对结节性质的判断更具提示作用。

本研究中有21个甲状腺结节应用ATA指南无法

进行分类, 其中包括17个边界清晰或不清晰的实性等回声结节伴微钙化、纵横比>1等恶性特征及4个囊实性结节伴以上可疑超声特征; 病理结果为恶性的占比为61.9%(13/21), 故对于此类结节, ATA指南分级的应用受到一定限制, 而应用其他2种诊断方法诊断的准确率为分别为66.7%(14/21)和71.4%(15/21), 二者虽对结节性质的判定有一定的提示作用, 但其诊断准确率也有所降低(低于89.1%和82.3%)。因此, 此类结节性质的判断存在一定难度, 在临床工作中, 不应局限或完全依赖上述3种方法。

另外, 本实验样本量较小及单一观察者的主观性可能会对研究结果造成一定的影响, 希望今后工作中开展更大样本量的、多位观察者参与的研究, 使实验结果更具说服力。本研究纳入的结节均为行手术或细针抽吸细胞学检查的结节, 良性结节部分是因有可疑征象或结节体积过大而切除, 部分则因为与恶性结节同时并发而切除, 其良恶性结节所占比例与临床实际发病率存在偏倚, 所以本研究结果更适用于需入院诊治的甲状腺结节的评估, 故存在一定的偏倚及局限性。

综上所述, 3种临床指南对甲状腺结节的诊断均具有较好的诊断效能, 其中Kwak TI-RADS及ACR TI-RADS分级方法的诊断效能更高于ATA指南, 3种方法各具优缺点, 临床应用时应根据情况酌情选用。

参考文献

1. Kwak JY, Han KH, Yoon JH, et al. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. *Radiology*, 2011, 260(3): 892-899.
 2. Haugen BR, Alexander EK, Bible KC, 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. *Thyroid*, 2016, 26(1): 1-133.
 3. 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 Am Coll Radiol*, 2017, 14(5): 587-595.
 4. 顾诗瑶, 黄瑛. 应用超声引导下甲状腺细针穿刺技术比较四种临床指南的诊断效果[J]. *中国医学影像学杂志*, 2019, 27(11): 853-856,861.
- GU Shiyao, HUANG Ying. Comparison of diagnostic efficacy of four clinical guidelines by ultrasound-guided fine needle aspiration biopsy

- of the thyroid[J]. Chinese Journal of Medical Imaging, 2019, 27(11): 853-856,861.
5. 姚建锋, 张煜华, 王全江, 等. ACR TI-RADS与Kwak TI-RADS对比在甲状腺结节定性诊断中的效能[J]. 中国临床医学影像杂志, 2019, 30(8): 537-539,547.
- YAO Jianfeng, ZHANG Yuhua, WANG Quanjiang, et al. Comparison of diagnostic efficacy between ACR TI-RADS and Kwak TI-RADS in qualitative diagnosis of thyroid nodules[J]. Journal of China Clinic Medical Imaging, 2019, 30(8): 537-539,547.
6. 张于芝, 徐婷, 顾经宇, 等. 2017美国放射学会甲状腺影像报告和数据系统(ACR-TIRADS)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, 2018, 27(6): 505-509.
- ZHANG Yuzhi, XU Ting, GU Jingyu, et al. Effectiveness evaluation of the thyroid imaging report and data system proposed by american radiological society (2017) (ACR-TIRADS) for differential diagnosis in thyroid nodules[J]. Chinese Journal of Ultrasonography, 2018, 27(6): 505-509.
7. Peng JY, Pan FS, Wang W, et al. Malignancy risk stratification and FNA recommendations for thyroid nodules: a comparison of ACR TI-RADS, AACE/ACE/AME and ATA guidelines. Am J Otolaryngol, 2020, 41(6): 102625.
8. 史宜鑫, 夏蜀珺, 陈林, 等. ACR2017版甲状腺超声影像与数据报告系统在中国人中的应用价值[J]. 中国超声医学杂志, 2020, 36(5): 394-397.
- SHI Yixin, XIA Shujun, CHEN Lin, et al. Values of thyroid imaging reporting and data system of American college of radiology 2017 edition in Chinese population[J]. Chinese Journal of Ultrasound in Medicine, 2020, 36(5): 394-397.

本文引用: 唐海玲, 刘千琪, 王建新, 邢萍, 巴黎, 吴长君. ATA指南、Kwak与ACR TI-RADS分类对甲状腺结节的诊断效能[J]. 临床与病理杂志, 2021, 41(7): 1541-1546. doi: 10.3978/j.issn.2095-6959.2021.07.012

Cite this article as: TANG Hailing, LIU Qianqi, WANG Jianxin, XING Ping, BA Li, WU Changjun. Diagnostic efficacy of the ATA guidelines, the Kwak TI-RADS and the ACR TI-RADS for thyroid nodules[J]. Journal of Clinical and Pathological Research, 2021, 41(7): 1541-1546. doi: 10.3978/j.issn.2095-6959.2021.07.012