

DOI:10.16781/j.0258-879x.2020.10.1077

· 论 著 ·

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

郭芳琪¹, 赵佳琦^{1*}, 陈蕊¹, 刘晟²

1. 海军军医大学(第二军医大学)长征医院超声诊疗科, 上海 200003

2. 海军军医大学(第二军医大学)长征医院普外三科, 上海 200003

[摘要] **目的** 应用2.0版人工智能(AI)自动检测系统 AI-SONIC™ Thyroid 对甲状腺结节进行术前超声评估, 并与不同年资超声医师应用常规超声检查主观诊断结论进行比较, 探讨2.0版AI自动检测系统在甲状腺结节良恶性鉴别诊断中的应用价值。**方法** 选择2019年8月至2020年1月于我院普通外科接受手术治疗的247例甲状腺结节患者(325枚甲状腺结节)。所有患者术前均由1名有13年甲状腺超声诊断工作经验的高级职称超声医师和1名有4年工作经验的初级职称超声医师分别进行常规超声检查, 同时由另1名有20年工作经验的超声医师在不知前2名医师判读结果的条件下利用2.0版AI自动检测系统进行超声检查。采用Kappa检验评价不同年资医师常规超声检查及2.0版AI自动检测系统的诊断结果与术后病理结果的一致性。**结果** 术后病理确诊恶性结节229枚, 良性结节96枚。低年资医师应用常规超声检查诊断甲状腺结节良恶性的灵敏度、特异度、准确度分别为85.15%(195/229)、66.67%(64/96)、79.69%(259/325), 高年资超声医师应用常规超声检查诊断甲状腺结节良恶性的灵敏度、特异度、准确度分别为93.45%(214/229)、79.17%(76/96)、89.23%(290/325), 2.0版AI自动检测系统诊断甲状腺结节良恶性的灵敏度、特异度、准确度分别为92.58%(212/229)、71.88%(69/96)、86.46%(281/325)。Kappa一致性检验结果显示, 高年资医师应用常规超声检查与病理诊断结果一致性较高(Kappa值为0.78, $P < 0.01$), 低年资医师应用常规超声检查、2.0版AI自动检测系统与病理诊断结果一致性一般(Kappa值为0.55、0.74, P 均 < 0.01)。**结论** 2.0版AI自动检测系统 AI-SONIC™ Thyroid 诊断甲状腺结节良恶性的灵敏度、准确度、特异度与高年资医师应用常规超声的检查结果相近, 有望成为术前评估甲状腺结节良恶性的可靠辅助手段。

[关键词] 人工智能; 高频超声; 甲状腺结节; 甲状腺影像报告和数据系统

[中图分类号] R 581

[文献标志码] A

[文章编号] 0258-879X(2020)10-1077-07

Efficacy of preoperative ultrasound evaluation of thyroid nodules by artificial intelligence automatic detection system version 2.0: a preliminary study

GUO Fang-qi¹, ZHAO Jia-qi^{1*}, CHEN Rui¹, LIU Sheng²

1. Department of Ultrasound, Changzheng Hospital, Naval Medical University (Second Military Medical University), Shanghai 200003, China

2. Department of General Surgery (III), Changzheng Hospital, Naval Medical University (Second Military Medical University), Shanghai 200003, China

[Abstract] **Objective** To investigate the efficacy of AI-SONIC™ Thyroid system, a version 2.0 artificial intelligence (AI) automatic detection system, in the preoperative ultrasound diagnosis of thyroid nodules, and to evaluate the application value of AI automatic detection system version 2.0 in the differential diagnosis of benign and malignant thyroid nodules by comparing with the subjective diagnosis conclusions of sonographers with different seniorities. **Methods** A total of 247 patients (325 thyroid nodules) admitted to the Department of General Surgery in our hospital from Aug. 2019 to Jan. 2020 were selected for this study. All patients underwent routine ultrasound examinations by a senior sonographer with 13 years of experience in thyroid ultrasound diagnosis and a junior sonographer with 4 years of work experience. At the same time, the patients were also examined by another sonographer with 20 years of work experience using AI automatic detection system version 2.0, without knowing the diagnosis conclusions of the above two sonographers. Kappa test was used to evaluate

[收稿日期] 2020-03-04 **[接受日期]** 2020-05-19

[基金项目] 国家自然科学基金(81501492). Supported by National Natural Science Foundation of China (81501492).

[作者简介] 郭芳琪, 硕士生, 住院医师. E-mail: 1002801624@qq.com

*通信作者(Corresponding author). Tel: 021-81886656, E-mail: qiqiblue67@163.com

the consistency of the results of routine ultrasound examination of sonographers with different seniorities and AI automatic detection system version 2.0 and the postoperative pathological results. **Results** The postoperative pathology confirmed 229 malignant nodules and 96 benign nodules. The sensitivity, specificity and accuracy in the diagnosis of benign and malignant thyroid nodules were 85.15% (195/229), 66.67% (64/96) and 79.69% (259/325), 93.45% (214/229), 79.17% (76/96) and 89.23% (290/325), and 92.58% (212/229), 71.88% (69/96) and 86.46% (281/325) for junior sonographer, senior sonographer and AI automatic detection system version 2.0, respectively. The *Kappa* consistency test results showed that the diagnosis result of senior sonographer was highly consistent with the pathological diagnosis result (*Kappa* value 0.78, $P < 0.01$), while the diagnosis results of junior sonographer and AI automatic detection system version 2.0 were generally consistent with the pathological diagnosis result (*Kappa* values 0.55 and 0.74, both $P < 0.01$). **Conclusion** The sensitivity, accuracy and specificity of the AI automatic detection system version 2.0 AI-SONIC™ Thyroid in diagnosing benign and malignant thyroid nodules are similar to those of routine ultrasound examination by senior sonographers, and the system might be a reliable auxiliary means for preoperative evaluation of benign and malignant thyroid nodules.

[Key words] artificial intelligence; high-frequency ultrasound; thyroid nodules; thyroid imaging reporting and data system
[Acad J Sec Mil Med Univ, 2020, 41(10): 1077-1083]

我国甲状腺癌患者以每年 20% 的速度持续增长, 2013 年全国甲状腺癌发病患者约有 14.39 万例^[1]。准确鉴别甲状腺结节的良恶性对于甲状腺癌的早诊断、早治疗十分必要。高频超声因无创、便捷、准确等优势被广泛用于甲状腺结节检查, 但是由于良恶性甲状腺结节的特征存在重叠, 图像特征判读也易受医师主观经验的影响^[2], 这可能导致误诊甚至过度诊断, 造成不可逆损伤^[3]。2017 年 5 月美国放射协会 (American College of Radiology, ACR) 发布了甲状腺影像报告和数据库系统 (thyroid imaging reporting and data system, TI-RADS) 分类诊断标准指南^[4], 着重评估甲状腺结节的内部结构、回声、纵横比、边界、强回声点 5 类特征, 并采用积分法进行风险分级^[5], 但仍然因医学成像设备成像条件、医师经验等而造成误诊或漏诊。前期研究表明, 人工智能 (artificial intelligence, AI) 自动检测系统 AI-SONIC™ Throid (DEMETICS® 超声诊断机器人) 对甲状腺结节良恶性诊断的灵敏度、准确度接近但稍逊于经验丰富的超声医师, 可作为术前评估甲状腺结节良恶性的有效补充, 但由于该研究样本量偏小, 特异度仅为 68.29%^[6]。2.0 版 AI 自动检测系统 DEMETICS® 超声诊断机器人通过扩大样本量进行深度学习, 除继续对甲状腺结节的声学细节特征进行自动识别、精细分析外, 也对微小结节、不典型增生结节及细胞生长活跃结节等的良恶性风险诊断概率值划分进行了优化调整, 其灵敏度、特异度、准确度较初始版 AI 自动检测系统显著提高。本研究回顾了 247 例手术患者共 325 枚甲状腺结节的超声影像学资料, 以术后病理诊断为金标准, 比较 2.0 版 AI 自动检测

系统与不同年资超声医师常规检查诊断良恶性甲状腺结节的差异, 探讨升级后 AI 自动检测系统术前评估甲状腺结节良恶性的价值。

1 资料和方法

1.1 研究资料 选择 2019 年 8 月至 2020 年 1 月 247 例于我院普通外科接受手术治疗甲状腺结节患者的 325 枚甲状腺结节。纳入标准: (1) 术前在我院接受高年资医师 (具有 13 年工作经验的高级职称医师) 常规超声检查、低年资医师 (具有 4 年工作经验的初级职称医师) 常规超声检查和 2.0 版 AI 自动检测系统检查; (2) 观察结节经过手术病理证实; (3) 观察结节有完整的横、纵标准切面图; (4) 观察结节的超声图像均无测量线或点的标注。排除标准: (1) 观察结节体积偏大, 超出切面的观察范围; (2) 超声检查横、纵标准切面显示不清晰; (3) 曾诊断偏良性结节仅进行超声随访者; (4) 曾诊断偏恶性结节仅进行细胞学穿刺者。本研究通过我院伦理委员会审批, 受检者检查前均签署了知情同意书。

1.2 仪器与方法

1.2.1 常规超声检查 采用日本日立公司 Hitachi Preirus 超声诊断仪进行甲状腺常规超声检查, 探头型号 L74M, 频率为 5~13 MHz。系统参数恒定, 增益为 30 dB, 时间增益补偿为中间位 (零补偿), 图像聚焦区深度为 2.00~3.75 cm, 组织热指数 < 0.4 , 机械指数为 1.2。

由 2 名多年从事甲状腺超声诊断工作经验的超声医师 (具有 13 年工作经验的高级职称医师 1 名, 具有 4 年工作经验的初级职称医师 1 名),

在检查前未获知患者信息及临床资料的情况下,按照上海市超声质量控制要求^[7]和《超声医学》第6版^[8]甲状腺超声操作要点及诊断标准,对术前患者的甲状腺组织(两侧叶、峡部)进行全面细致检查。检查过程中若发现结节,则需反复重点扫查,并按照ACR发布的TI-RADS分类标准^[9]对检出甲状腺结节进行声学特征评估,以实性、低回声或极低回声、边界不规则或边缘小分叶、纵横比 >1 、微钙化这5项特征作为偏恶性的判断条件,确立TI-RADS 3类(无恶性超声表现)、4a类(1个恶性超声表现)、4b类(2个恶性超声表现)、4c类(3个或4个恶性超声表现)和5类(5个恶性超声表现)。将TI-RADS 1~3类归为偏良性, TI-RADS 4~5类归为偏恶性^[10-11]。

1.2.2 2.0版AI自动检测系统检查 采用2.0版AI-SONIC™ Thyroid甲状腺结节AI辅助诊断系统(DEMETICS®超声诊断机器人,浙江德尚韵兴医疗科技有限公司)进行检查。该系统经过6万余例甲状腺结节穿刺及病理结果对照的深度学习,创新性地引入旋转不变网络层、split dropout等新概念^[12],持续优化深度学习理论和变分能量泛函结合的图像处理与分析方法。

检查操作步骤同常规超声检查,由1名从事甲状腺超声诊断工作20年的超声医师在不知患者任何临床信息和前2名医师判读结果的条件下进行操作,并由1名AI自动检测系统公司的应用工程师对该操作医师进行全程规范的监督与指导,便于获取最佳的标准化超声切面图像。医师将扫查的图像通过采集卡或DICOM方式传输至服务器,利用算法自动探测病灶位置,并进行自动标记、处理、分析。通过AI算法自动定量甲状腺结节的大小、边界、形态、内部回声、钙化灶5类特征,每个结节分别经过4次纵、横切面显示,实时冻结存储传输至自动判读系统,直至TI-RADS分类及风险概率值趋于稳定时,记录系统自动判读结节的最高1次TI-RADS分类及风险概率值^[6]。风险概率值为0~1,其中0~0.40为偏良性(无恶性特征),建议随访;0.41~0.60为可疑恶性(存在恶性特征倾向),建议进一步检查;0.61~1为偏恶性(存在恶性特征),建议进行病理检查。当结节风险概率值为0.41时说明该结节具有一定的恶性特征倾向,存在一定的恶性风险,因此将风险概率值0.41归为诊断恶性临界值,由安装在客户端软件的超声工作站自动呈现出良恶性超声诊断结论及处理建议(图1)。

1.3 统计学处理 应用SPSS 21.0软件进行统计学分析。计量资料以 $\bar{x} \pm s$ 表示,计数资料以例(枚)数和百分数表示。以术后病理结果为金标准,采用基于Kappa系数的一致性检验^[13]评价2名不同年资医师常规超声检查及2.0版AI自动检测系统诊断结果与术后病理结果的一致性,计算Kappa系数、灵敏度、特异度、准确度。Kappa系数越大说明一致性越高,Kappa值 ≥ 0.75 说明一致性好,Kappa值为0.40~ <0.75 说明一致性一般,Kappa值 <0.40 说明一致性差。检验水准(α)为0.05。

2 结果

2.1 甲状腺结节特征 共纳入247例接受手术治疗的甲状腺结节患者(325枚甲状腺结节),其中男40例(69枚甲状腺结节),女207例(256枚甲状腺结节),年龄为13~84岁,平均年龄为(49.1±10.8)岁。术后病理确诊恶性结节229枚,平均最大径为(11.40±6.17)mm,其中乳头状癌202枚、滤泡上皮肿瘤6枚、滤泡上皮不典型增生17枚、髓样癌1枚、恶性肿瘤3枚;良性结节96枚,平均最大径为(19.45±9.16)mm,其中腺瘤25枚、结节性甲状腺肿62枚、滤泡性结节增生2枚、胶质潴留1枚、甲状旁腺增生2枚、良性结节1枚、纤维性钙化3枚。

2.2 不同年资超声医师与2.0版AI自动检测系统对甲状腺结节良恶性的诊断效能 325枚结节中,病理诊断良性结节96枚,恶性结节229枚;低年资超声医师检出98枚结节偏良性,227枚偏恶性;高年资超声医师检出91枚偏良性,234枚偏恶性;2.0版AI自动检测系统检出86枚偏良性(建议随访),31枚可疑恶性(建议进一步检查),208枚偏恶性(建议手术)。2.0版AI自动检测系统与术后病理结果一致的结节有281枚,其中69枚偏良性,21枚可疑恶性,191枚偏恶性。2.0版AI自动检测系统误诊44枚结节,其中27枚良性结节中误诊为可疑恶性10枚、偏恶性17枚,17枚恶性结节误诊为偏良性。在2.0版AI自动检测系统误诊的44枚结节中,31枚结节高年资超声医师也同时误诊,其中12枚恶性结节误诊为偏良性,19枚良性结节误诊为偏恶性,包括9枚滤泡上皮不典型增生、2枚乳头状癌、1枚髓样癌、5枚滤泡腺瘤、1枚嗜酸性腺瘤、11枚结节性甲状腺肿、1枚局灶纤维结节伴钙化、1枚囊肿吸收后改变;其余13枚结节为2.0版AI自动检测系统单独误诊,其中5枚恶性结节误诊为偏良性,8枚良性结节误诊

为偏恶性,包括4枚乳头状癌、1枚滤泡上皮不典型增生、5枚结节性甲状腺肿、1枚肉芽肿性甲状腺炎、1枚纤维性钙化、1枚甲状旁腺增生。2.0版AI自动检测系统诊断甲状腺结节良恶性的灵敏度(92.58%, 212/229)、特异度(71.88%, 69/96)和准确度(86.46%, 281/325)均与高年资医师[分别为93.45%(214/229)、79.17%(76/96)和89.23%(290/325)]较接近,且均高于低年资医师[分别为85.15%(195/229)、66.67%(64/96)和79.69%(259/325)]。2.0版AI自动检测系统、低年资医师应用常规超声检查、高年

资医师应用常规超声检查诊断甲状腺结节良恶性的阳性预测值和阴性预测值分别为88.70%、85.90%、91.45%和71.88%、65.31%、83.52%。与术后病理诊断结果进行一致性检验结果显示,低年资医师应用常规超声、2.0版AI自动检测系统诊断甲状腺结节良恶性的结果与术后病理诊断结果一致性一般($Kappa$ 值为0.55、0.74, P 均 <0.01),高年资医师应用常规超声诊断甲状腺结节良恶性的结果与术后病理诊断结果一致性较高($Kappa$ 值为0.78, $P<0.01$)。见表1。

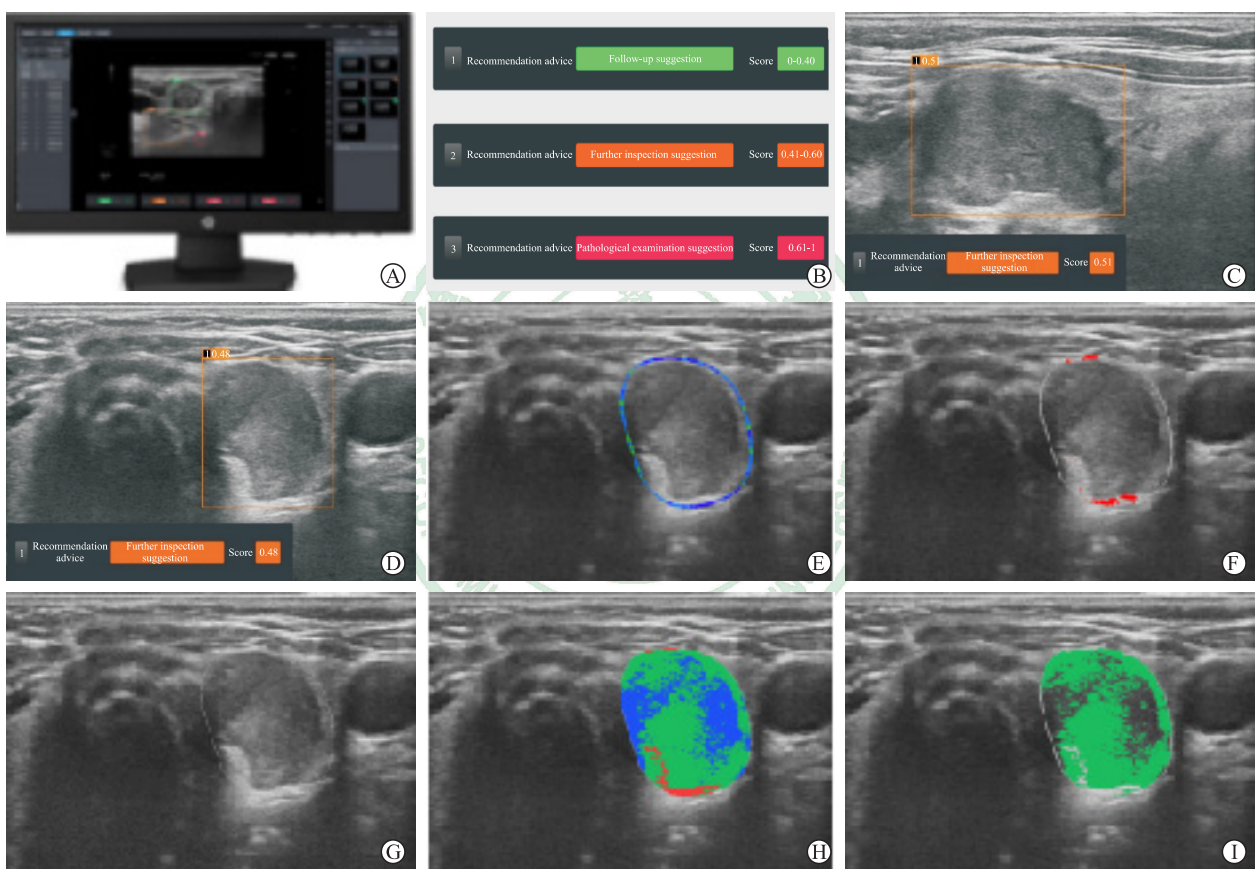


图1 2.0版AI自动检测系统工作界面

Fig 1 Working interface of AI automatic detection system version 2.0

The risk probability scores issued for the lesion were 0.51 (vertical) and 0.48 (horizontal), indicating suspicious malignancy, and further testing was recommended. The pathological results showed thyroid follicular tumor. A: The ultrasonic workstation of the software; B: The automatic presentation of ultrasonic diagnosis recommendations and risk probability scores; C, D: Long-axis (C) and short-axis (D) standard section figures of left thyroid nodule, and AI-SONIC™ Throid system image analysis showing the suspected malignant lesion, with risk probability scores being 0.51 and 0.48 in vertical and horizontal sections, respectively; E: The edge feature analysis showed that the color changed from clear to blurred from blue, green, yellow and red, and the edge of the lesion was still clear by AI-SONIC™ Throid system; F: The edge of the lesion was automatically identified and completely wrapped, and a few hyperechoes (red dots) inside the boundary were analyzed; G: AI-SONIC™ Throid system showed that the internal components of the lesion were mainly solid; H: The internal echo of lesion changed from uniform to non-uniform was showed in turn by blue, green, yellow and red, and AI-SONIC™ Throid system showed the internal echo non-uniformity of the lesion; I: The red, green, blue, dark blue and purple echoes inside the nodule represented high echo, medium echo, low echo, very low echo and no echo, respectively, and AI-SONIC™ Throid system detected the lesion as medium echo. AI: Artificial intelligence

表 1 术前不同年资医师常规超声检查与 2.0 版 AI 自动检测系统对甲状腺结节良恶性的诊断效能比较

Tab 1 Comparison of preoperative diagnosis efficacy of benign and malignant thyroid nodules by sonographers with different seniorities and AI automatic detection system version 2.0

Diagnosis method	Pathology <i>n</i>		Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	<i>Kappa</i> value	<i>P</i> value
	Benign	Malignant							
Junior sonographer			85.15	66.67	79.69	85.90	65.31	0.55	<0.01
Partial benign (TI-RADS 1-3)	64	34							
Partial malignant (TI-RADS 4-5)	32	195							
Senior sonographer			93.45	79.17	89.23	91.45	83.52	0.78	<0.01
Partial benign (TI-RADS 1-3)	76	15							
Partial malignant (TI-RADS 4-5)	20	214							
AI automatic detection system			92.58	71.88	86.46	88.70	71.88	0.74	<0.01
Partial benign (risk probability score 0-0.40)	69	17							
Suspected malignant (risk probability score 0.41-0.60)	10	21							
Partial malignant (risk probability score 0.61-1)	17	191							

AI: Artificial intelligence; TI-RADS: Thyroid Imaging Reporting and Data System; PPV: Positive predictive value; NPV: Negative predictive value

2.3 2.0 版 AI 自动检测系统对可疑恶性甲状腺结节的诊断结果 对 AI 自动检测系统诊断的 31 枚可疑恶性结节与术后病理结果进行对比发现, 21 枚为恶性, 10 枚为良性。10 枚良性结节中 7 枚为结节性甲状腺肿, 1 枚为肉芽肿性甲状腺炎, 1 枚为滤泡性腺瘤, 1 枚为滤泡性结节增生 (图 2)。

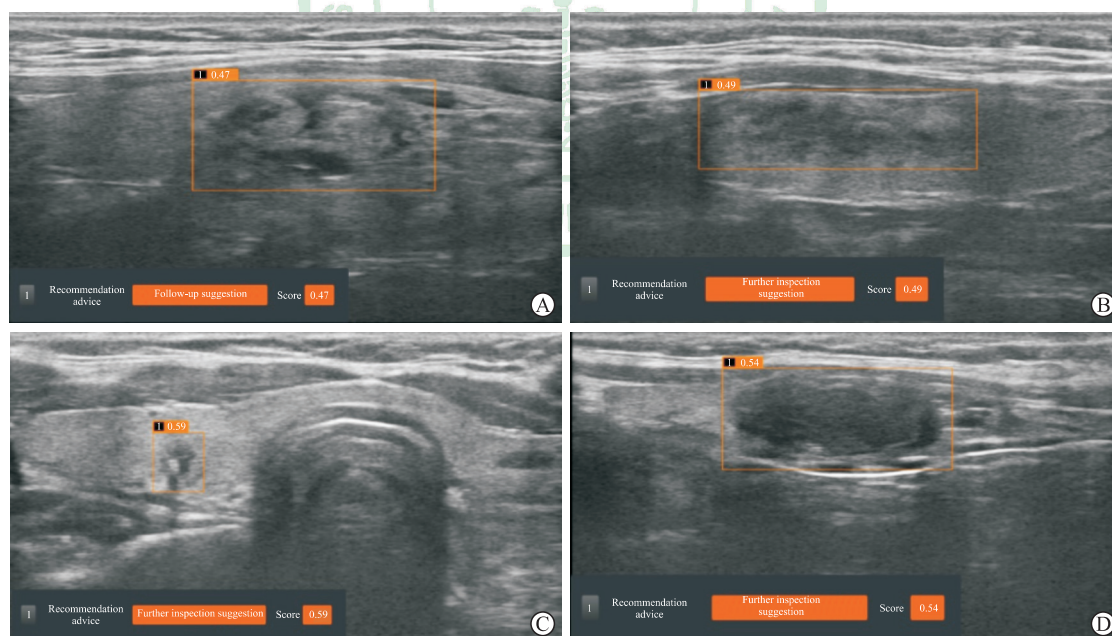


图 2 2.0 版 AI 自动检测系统误诊为可疑恶性的良性甲状腺结节

Fig 2 Benign thyroid nodules misdiagnosed as suspected malignancy by AI automatic detection system version 2.0

A: In the longitudinal section of this nodule, the risk probability score by AI-SONIC™ Throid system was 0.47, indicating suspected malignancy, and pathological results diagnosed as nodular goiter; B: In the longitudinal section of this nodule, the risk probability score by AI-SONIC™ Throid system was 0.49, indicating suspected malignancy, and the pathological results diagnosed as granulomatous thyroiditis; C: In the transverse section of this nodule, the risk probability score by AI-SONIC™ Throid system was 0.59, indicating suspected malignancy, and pathological results diagnosed as follicular adenoma; D: In the longitudinal section of this nodule, the risk probability score by AI-SONIC™ Throid system was 0.54, indicating suspected malignancy, and pathological results diagnosed as follicular nodular hyperplasia. AI: Artificial intelligence

3 讨论

随着 AI 辅助诊断技术的不断更新与改进, 运用 AI 辅助诊断技术可对组织内异常区域进行自动分割、快速分析, 定量病灶的影像学特征, 并可对检测区域进行精准评估, 降低人为因素所致的医疗失误^[14-15]。近年来, AI 驱动的超声技术也越来越成熟, 诊断结果越来越贴近病理学诊断, 特别是超声 AI 诊断工具的出现, 为甲状腺结节、乳腺结节等高发疾病的早期筛查和良恶性评估提供了有益补充^[16]。

AI-SONIC™ Throid 系统 2.0 版经过后期不断输入病例数据进行特征学习、病灶分割、多层次特征提取, 较 1.0 版不仅在甲状腺结节超声灰阶像提供了客观一致的视觉影像, 对系统分析、界面显示、诊断结果划分进行了改善, 降低了不同超声医师检查的差异及同一超声医师多次判读的差异, 还能提供可修正的诊断结果, 快速生成数字化报告供医师参考。本研究纳入 247 例患者共 325 枚甲状腺结节作为研究对象, 比较了 2.0 版 AI 自动检测系统与不同年资医师应用常规超声检查对甲状腺结节良恶性的诊断效能, 结果显示 2.0 版 AI 自动检测系统的灵敏度、特异度、准确度均低于高年资医师, 但高于低年资医师。分析原因可能是低年资医师临床经验相对不足, 对结节的认知、观察及专注能力等均不及高年资医师, 在短时间内做出诊断不可避免会存在漏诊和误诊。

此外, 2.0 版 AI 自动检测系统较 1.0 版在用户管理方面优化了预览界面, 新增了良恶性诊断详情, 并且对需进一步跟踪随访的病例增加了收藏或标注设置, 在 AI 分析算法方面也对目标区域进行了降噪、增强、细化图像特征等预处理, 提升了 AI 诊断的稳定性, 灵敏度、特异度、准确度也有所提升。通过高效智能整合, 2.0 版 AI 自动检测系统对甲状腺结节的风险概率值划分更精细化, 尤其针对一些无法用语言描述, 甚至精细结构肉眼较难观察的病变, 如微小结节、木乃伊结节及超声特征不明显的恶性结节, 系统界面做出了 0.41~0.60 可疑区域风险概率值的黄色框优化提示, 为超声医师做出最终诊断提供了重要参考。本研究中, 2.0 版 AI 自动检测系统诊断甲状腺结节良恶性的灵敏度、特异度和准确度与 1.0 版 AI 自动检测系统^[6]相比均更高 [分别为 92.58% (212/229) vs 89.58%

(86/96)、71.88% (69/96) vs 68.29% (28/41)、86.46% (281/325) vs 83.21% (114/137)], 与病理诊断结果一致性也相对较高 (Kappa 值为 0.74, 1.0 版 AI 自动检测系统与病理诊断结果一致性的 Kappa 值为 0.59^[6])。

本研究中, 2.0 版 AI 自动检测系统共误诊 44 枚甲状腺结节, 其中 31 枚结节在高年资医师利用常规超声检查时同样误诊。其原因可能是多数良性结节与腺体增生或炎症形成有关, 组织质地较软, 而恶性结节多数由细胞异常增殖和分化所致, 组织质地较硬, 经验丰富的超声医师可通过借助弹性成像分析结节组织的软硬程度或硬度属性对良恶性进行评估, 硬度越大甲状腺结节恶性程度越高^[17]。对于良恶性超声特征存在重叠的结节, 如胶质潴留干涸后局部晶体析出形成的类似钙化的成分、炎症细胞分布不规则的肉芽肿性甲状腺炎等, 高年资超声医师可借助常规超声实时、动态、可回放等优势, 反复多切面观察结节的物理特性及彩色血流分布特征, 并结合患者的甲状腺功能及相关抗体的实验室检查结果进行针对性综合判断分析, 从而确保降低误诊率。

研究发现, 甲状腺癌中以乳头状癌多见, 髓样癌的发病率仅占甲状腺癌的 2%, 约 70% 的甲状腺髓样癌患者在触诊发现甲状腺结节时就已经发生淋巴结转移^[18], 但在超声图像上缺乏特异性。本研究中, 2.0 版 AI 自动检测系统与常规超声检查同时误诊的 1 枚甲状腺髓样癌结节, 高、低年资超声医师均诊断为 TI-RADS 3 类, 考虑腺瘤囊性变可能; 2.0 版 AI 自动检测系统风险概率值为 0.24 (纵)、0.32 (横), 亦诊断为偏良性、建议随访。由于该结节体积较大, 患者接受手术切除治疗, 术后病理提示甲状腺髓样癌。分析该例患者误诊原因, 甲状腺髓样癌较乳头状癌少见, 尽管在前期工作中不断完善加载 AI 自动检测系统病例数据的学习, 但还是相对缺乏该类少见病理类型甲状腺结节的大数据输入, 大数据样本的准确度、可靠度仍需要一定的时间进行搜集和积累, 同时超声图像上甲状腺实质可显示质地纹理不均匀, 与周围结节分界不明显, 使精准全自动分割甲状腺结节具有挑战。

AI 自动检测系统基于甲状腺结节的超声图像特征, 在前期已加载大量数据进行了较深入的研究, 能够辅助术前甲状腺结节的良恶性评估^[6], 但

目前的版本都只针对静态灰阶超声图像进行处理和分析,尚不能对甲状腺结节弹性成像、彩色多普勒血流显像等多模态进行智能诊断。今后可以研发有效分析动态视频、彩色多普勒图像及弹性图像的多模态超声AI诊断工具,运用数字图像处理技术对超声图像进行相应的标注,以期为临床提供更全面、更直观、更快速的信息,提高诊断准确性。

[参考文献]

- [1] 杨雷,郑荣寿,王宁,曾红梅,袁延楠,张思维,等. 2013年中国甲状腺癌发病与死亡情况[J]. 中华肿瘤杂志, 2017,39:862-867.
- [2] CHOI S H, KIM E K, KWAK J Y, KIM M J, SON E J. Interobserver and intraobserver variations in ultrasound assessment of thyroid nodules[J]. *Thyroid*, 2010, 20: 167-172.
- [3] LA VECCHIA C, MALVEZZI M, BOSETTI C, GARAVELLO W, BERTUCCIO P, LEVI F, et al. Thyroid cancer mortality and incidence: a global overview[J]. *Int J Cancer*, 2015, 136: 2187-2195.
- [4] TESSLER F N, MIDDLETON W D, GRANT E G, HOANG J K, BERLAND L L, TEEFEY S A, 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: 587-595.
- [5] WANG Y, LEI K R, HE Y P, LI X L, REN W W, ZHAO C K, et al. Malignancy risk stratification of thyroid nodules: comparisons of four ultrasound Thyroid Imaging Reporting and Data Systems in surgically resected nodules[J/OL]. *Sci Rep*, 2017, 7: 11560. doi: 10.1038/s41598-017-11863-0.
- [6] 郭芳琪,赵佳琦,刘晟. 人工智能自动检测系统在甲状腺结节术前超声诊断中的应用[J]. 第二军医大学学报, 2019,40: 1183-1189.
GUO F Q, ZHAO J Q, LIU S. Application of artificial intelligence automatic detection system in preoperative ultrasonic diagnosis of thyroid nodules[J]. *Acad J Sec Mil Med Univ*, 2019, 40: 1183-1189.
- [7] 孙德胜,刘俐,石宇,胡会英,董理聪,胡正明,等. “全流程超声质控管理”软件系统(V1.0版)应用与实践[J/CD]. 中华医学超声杂志(电子版), 2019, 16: 329-332.
- [8] 郭万学,燕山,杨浣宜,伍于添,王志刚. 超声医学[M]. 6版. 北京:人民军医出版社, 2011:263-278.
- [9] WILDMAN-TOBRINER B, BUDA M, HOANG J K, MIDDLETON W D, THAYER D, SHORT R G, et al. Using artificial intelligence to revise ACR TI-RADS risk stratification of thyroid nodules: diagnostic accuracy and utility[J]. *Radiology*, 2019, 292: 112-119.
- [10] 宋琳琳,周贵明,朱莉,孟召伟,仇晓菲,陈冬玥. 甲状腺影像报告和数据系统联合甲状腺细胞病理学 Bethesda 报告系统对甲状腺结节的应用价值[J/CD]. 中华医学超声杂志(电子版), 2018, 15: 713-718.
- [11] 候雪琴,范雪,杨凌菲,董建,宋涛. 甲状腺影像报告和数据系统分类诊断甲状腺结节良恶性的准确性[J]. 中国影像医学杂志, 2015, 23: 489-493.
- [12] PEDERSEN D H, MANSOURVAR M, SORTSØ C, SCHMIDT T. Predicting dropouts from an electronic health platform for lifestyle interventions: analysis of methods and predictors[J/OL]. *J Med Internet Res*, 2019, 21: e13617. doi: 10.2196/13617.
- [13] 郭轶斌,郭威,秦宇辰,何倩,张新信,吴骋. 基于 Kappa 系数的一致性检验及其软件实现[J]. 中国卫生统计, 2016, 33: 169-170, 174.
- [14] LEHMAN C D, WELLMAN R D, BUIST D S, KERLIKOWSKA K, TOSTESON A N, MIGLIORETTI D L; Breast Cancer Surveillance Consortium. Diagnostic accuracy of digital screening mammography with and without computer-aided detection[J]. *JAMA Intern Med*, 2015, 175: 1828-1837.
- [15] SHIMAUCHI A, GIGER M L, BHOOSHAN N, LAN L, PESCE L L, LEE J K, et al. Evaluation of clinical breast MR imaging performed with prototype computer-aided diagnosis breast MR imaging workstation: reader study[J]. *Radiology*, 2011, 258: 696-704.
- [16] AKKUS Z, CAI J, BOONROD A, ZEINODDINI A, WESTON A D, PHILBRICK K A, et al. A survey of deep-learning applications in ultrasound: artificial intelligence-powered ultrasound for improving clinical workflow[J]. *J Am Coll Radiol*, 2019, 16(9 Pt B): 1318-1328.
- [17] RAGO T, SCUTARI M, LOIACONO V, SANTINI F, TONACCHERA M, TORREGROSSA L, et al. Low elasticity of thyroid nodules on ultrasound elastography is correlated with malignancy, degree of fibrosis, and high expression of galectin-3 and fibronectin-1[J]. *Thyroid*, 2017, 27: 103-110.
- [18] SIEGEL R L, MILLER K D, JEMAL A. Cancer statistics, 2018[J]. *CA Cancer J Clin*, 2018, 68: 7-30.

[本文编辑] 杨亚红