

DOI: 10.16781/j.CN31-2187/R.20230009

· 论 著 ·

超声人工智能辅助诊断系统在最大径 ≤ 2 cm的BI-RADS 4类乳腺结节诊断中的应用价值

陈蕊¹, 吴墅¹, 郭佳¹, 郭芳琪², 赵佳琦^{2*}

1. 上海中医药大学附属曙光医院超声医学科, 上海 201203

2. 同济大学附属上海市第四人民医院超声医学科, 上海 200434

[摘要] **目的** 探讨超声人工智能(AI)辅助诊断系统对最大径 ≤ 2 cm的乳腺影像报告与数据系统(BI-RADS) 4类乳腺结节的诊断价值。**方法** 回顾性分析2020年5月至2022年10月于上海中医药大学附属曙光医院进行超声检查并诊断为BI-RADS 4类乳腺结节的204例患者共210个最大径 ≤ 2 cm结节的二维超声图像。以术后病理结果为金标准,评价常规超声和AI系统(风险评分值阈值设为0.65、0.70)对最大径 ≤ 2 cm的BI-RADS 4类乳腺结节良恶性的诊断效能。**结果** 210个乳腺结节中良性结节94个,恶性结节116个。高年资超声医师常规超声检查诊断乳腺结节良恶性的灵敏度为92.24%,特异度为75.53%,准确度为84.76%;AI系统(阈值0.65)诊断乳腺结节良恶性的灵敏度为92.24%,特异度为71.28%,准确度为82.86%;AI系统(阈值0.70)诊断乳腺结节良恶性的灵敏度为90.52%,特异度为79.79%,准确度为85.71%。AI系统(阈值0.70)诊断BI-RADS 4a类结节的准确度高于常规超声和AI系统(阈值0.65)(79.41% vs 77.94%、75.00%)。高年资超声医师通过常规超声对最大径 ≤ 1 cm的结节诊断准确度最高,为86.36%,AI系统(阈值0.65)及AI系统(阈值0.70)准确度分别为81.82%、84.09%。**结论** 超声AI辅助诊断系统可辅助鉴别诊断最大径 ≤ 2 cm的BI-RADS 4类乳腺结节的良恶性。

[关键词] 超声检查;人工智能;计算机辅助诊断系统;乳腺肿瘤;乳腺影像报告与数据系统

[引用本文] 陈蕊,吴墅,郭佳,等.超声人工智能辅助诊断系统在最大径 ≤ 2 cm的BI-RADS 4类乳腺结节诊断中的应用价值[J].海军军医大学学报,2024,45(5):592-598. DOI: 10.16781/j.CN31-2187/R.20230009.

Ultrasonic artificial intelligence system in diagnosis of BI-RADS class 4 breast nodules with maximum diameter ≤ 2 cm

CHEN Rui¹, WU Shu¹, GUO Jia¹, GUO Fangqi², ZHAO Jiaqi^{2*}

1. Department of Ultrasound, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China

2. Department of Ultrasound, Shanghai Fourth People's Hospital Affiliated to Tongji University, Shanghai 200434, China

[Abstract] **Objective** To evaluate the diagnostic value of ultrasonic artificial intelligence (AI) system in breast imaging-reporting and data system (BI-RADS) class 4 breast nodules with maximum diameter ≤ 2 cm. **Methods** A total of 210 breast nodules were analyzed in 204 patients with pathological results obtained by surgery at Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine. The maximum diameter of the breast nodules was ≤ 2 cm. With the pathological results as the gold standard, all nodules were examined by routine ultrasound (US) and AI aided diagnosis system. The values of routine US by a senior physocian, AI (threshold value 0.65) and AI (threshold value 0.70) in diagnosing benign and malignant in BI-RADS class 4 breast nodules with maximum diameter ≤ 2 cm were evaluated. **Results** The pathological results showed that 210 breast nodules included 94 benign nodules and 116 malignant nodules. The sensitivity, specificity and accuracy of routine US in diagnosing benign and malignant breast nodules were 92.24%, 75.53% and 84.76%, respectively;

[收稿日期] 2023-01-12 [接受日期] 2023-11-09

[基金项目] 上海市虹口区卫生健康委员会医学科研课题(虹卫2302-26),上海市虹口区卫生健康委员会临床重点扶持专科项目(HKLCFC202404),海军军医大学(第二军医大学)第二附属医院人才建设三年行动计划——“金字塔人才工程”军事医学人才项目(1009),同济大学附属上海市第四人民医院科研启动专项(SYKYQD06101)。Supported by Medical Research Project of Health Commission of Shanghai Hongkou District (HW2302-26), Clinical Key Supporting Project of Health Commission of Shanghai Hongkou District (HKLCFC202404), Military Medical Talent Project of “Pyramid Talent Program” of Three-Year Action Plan for Talent Construction of The Second Affiliated Hospital of Naval Medical University (Second Military Medical University) (1009), Start-up Scientific Research Project of Shanghai Fourth People's Hospital Affiliated to Tongji University (SYKYQD06101).

[作者简介] 陈蕊,主治医师。E-mail: defuyf@163.com

*通信作者(Corresponding author). Tel: 021-55603999, E-mail: qiqiblue67@163.com

the sensitivity, specificity and accuracy of AI (threshold value 0.65) were 92.24%, 71.28% and 82.86%, respectively; and those of AI (threshold value 0.70) were 90.52%, 79.79% and 85.71%, respectively. The accuracy of diagnosing BI-RADS 4a nodules of AI (threshold value 0.70) was higher than those of routine US and AI (threshold value 0.65) (79.41% vs 77.94%, 75.00%). The senior physician had the highest diagnostic accuracy of 86.36% for nodules with maximum diameter ≤ 1 cm using routine US. The accuracies of the AI system with threshold value 0.65 and 0.70 were 81.82% and 84.09%, respectively.

Conclusion Ultrasonic AI diagnosis system can assist to improve the diagnostic efficiency of BI-RADS class 4 breast nodules with maximum diameter ≤ 2 cm.

[**Key words**] ultrasonography; artificial intelligence; computer aided diagnosis system; breast neoplasms; breast imaging-reporting and data system

[**Citation**] CHEN R, WU S, GUO J, et al. Ultrasonic artificial intelligence system in diagnosis of BI-RADS class 4 breast nodules with maximum diameter ≤ 2 cm[J]. Acad J Naval Med Univ, 2024, 45(5): 592-598. DOI: 10.16781/j.CN31-2187/R.20230009.

乳腺癌已成为全世界女性最常见的癌症并位列癌症相关死亡原因第2位^[1],在我国城市女性肿瘤发生率中居于首位^[2]。影像学检查是最常用的非侵入性乳腺结节诊断方法,国内指南推荐使用超声及X线用于乳腺结节的一线检查^[3]。乳腺位置表浅,使用高频探头可以获得高质量的超声图像,且乳腺良恶性病变声像图特征不同,因此乳腺影像报告与数据系统(breast imaging-reporting and data system, BI-RADS)为乳腺结节良恶性风险诊断提供了定性分类标准,但不可避免受图像质量、医师专业素质、主观判断等因素影响,存在操作者间差异^[4]。BI-RADS 4类乳腺结节诊断为恶性的符合率跨度较大(3%~94%),尤其对于最大径 < 2 cm的乳腺小结节、早期乳腺癌病灶较小且超声图像特征不典型的结节,常规超声检查因超声医师资历、经验不同,对其良恶性的鉴别诊断较为困难^[5]。

乳腺超声人工智能(artificial intelligence, AI)辅助诊断系统采用轻量化深度学习框架(DE-Light),根据乳腺二维BI-RADS和大数据分析,可自动识别乳腺结节,并对结节良恶性风险进行预估判读,为超声医师鉴别乳腺结节性质提供了依据,提高了乳腺结节超声诊断的工作效率。本研究探讨了超声AI辅助诊断系统对最大径 ≤ 2 cm的BI-RADS 4类乳腺结节良恶性的诊断价值。

1 资料和方法

1.1 研究对象 回顾性选择于2020年5月至2022年10月在上海中医药大学附属曙光医院进行超声检查并诊断为BI-RADS 4类乳腺结节的患者204例(共210个结节),均为女性,年龄为

28~89(55.52 \pm 3.47)岁,结节最大径为5~20(15.76 \pm 3.42)mm。纳入标准:(1)术前接受乳腺常规超声检查和乳腺超声AI辅助诊断系统检测;(2)结节经过手术病理检查证实;(3)观察结节有完整的横、纵标准超声切面图;(4)结节最大径 ≤ 2 cm,且经1名具备10年乳腺超声工作经验的主治医师诊断为BI-RADS 4类。排除标准:(1)观察结节体积偏大(最大径 > 2 cm)或曾诊断BI-RADS 3类及以下仅进行超声随访者;(2)超声检查横、纵标准切面显示不清晰;(3)曾诊断偏恶性结节仅进行细胞学穿刺者;(4)结节经过消融等处理者。本研究通过上海中医药大学附属曙光医院医学伦理委员会审批。

1.2 常规超声检查 常规超声检查采集的所有乳腺图像数据均来自上海中医药大学附属曙光医院超声医学科,超声仪器型号包括Toshiba Aplio 300、Hitachi Vision Preirus、Wisonic Clover60、Philips IU22、Esaote Mylab ClassC等,使用宽频线阵探头,频率为5.0~14.0 MHz。所有图像数据均不包含任何患者身份信息。超声检查医师按照上海市超声质量控制要求^[6]和《超声医学》第6版乳腺超声操作要点及诊断标准^[7],对术前患者的双侧乳腺及腋下淋巴结进行全面、细致检查。检查过程中若发现结节,则反复重点扫查,获得横向和纵向静态图像并存储。所有乳腺结节的超声图像均由1名具备20年以上乳腺超声工作经验的主任医师在对病理结果不知情的情况下进行回顾性分析,按照美国放射学会发布的BI-RADS分类标准^[4]对检出的乳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

5项特征作为偏恶性的判断条件,确定 BI-RADS 4a类(1个恶性超声表现)、4b类(2个恶性超声表现)和4c类(3或4个恶性超声表现),并做出偏良性或偏恶性诊断。

1.3 乳腺超声 AI 辅助诊断系统检测 采用乳腺超声 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) 进行检测。该系统经过 10 万余例乳腺结节穿刺及病理结果对照的深度学习,创新性地引入了 DE-Light,优化了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。将所有乳腺结节二维横、纵切面超声图像通过采集卡或 DICOM 方式传输至服务器, AI 系统算法自动探测病灶位置并进行标记,同时测定乳腺

结节的大小、边界、形态、内部回声、钙化灶 5 类声学特征。所有结节超声图像均包括纵、横切面,由 AI 系统自动判读结节的良恶性并给出风险评分值(图 1)。风险评分值范围为 0~1,其中 0~0.40 为偏良性(无恶性特征),0.41~0.60 为可疑恶性(存在恶性特征倾向),0.61~1 为偏恶性(存在恶性特征)。考虑设置不同阈值将影响 AI 系统检测结果的判断,一般情况下阈值越低 AI 系统诊断的灵敏度越高、特异度越低,阈值越高 AI 系统诊断的灵敏度越低、特异度越高,所以本研究将 AI 系统阈值设置为 0.65 和 0.70,分别对乳腺结节进行良恶性诊断,提示结果为“偏良性”或“偏恶性”。



图 1 乳腺超声 AI 辅助诊断系统自动识别乳腺结节并提示良恶性风险示例

Fig 1 An AI-assisted diagnostic system for breast ultrasound automatically identifying breast nodules and indicating the risk of malignancy

A: The AI assisted diagnostic system did not found nodules; B: The AI assisted diagnostic system diagnosed that the nodule may be benign with a risk probability value of 0.13 (indicated by the green box); C: The AI system diagnosed that the nodule may be malignant with a risk probability value of 0.74 (indicated by the red box). AI: Artificial intelligence.

1.4 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(个)数和百分数表示,组间比较采用 χ^2 检验。以术后病理结果为金标准,计算灵敏度、特异度、准确度。检验水准(α)为 0.05。

2 结果

2.1 术后病理结果 204 例患者共 210 个乳腺结节,所有患者均行外科手术治疗。根据第 5 版 WHO 乳腺肿瘤的分类^[8],术后病理提示良性结节 94 个,包括乳腺纤维腺瘤 31 个、乳腺腺病 36 个、导管内乳头状瘤 11 个、肉芽肿性乳腺炎 7 个、乳腺导管扩张伴囊肿形成 2 个、纤维上皮源性肿瘤 3 个、乳腺组织急慢性炎细胞浸润 2 个、纤维脂肪

组织伴充血 1 个、腺病伴局部导管乳头状增生 1 个;恶性结节 116 个,包括浸润性导管癌 98 个、浸润性小叶癌 3 个、导管原位癌 11 个、黏液癌 1 个、包裹性乳头状癌 3 个。

2.2 常规超声和 AI 系统对最大径 ≤ 2 cm 的 BI-RADS 4 类乳腺结节的良恶性鉴别结果 高年资超声医师通过常规超声诊断乳腺结节偏良性 80 个、偏恶性 130 个,与术后病理结果对照,正确诊断良性结节 71 个、恶性结节 107 个,灵敏度为 92.24%,特异度为 75.53%,准确度为 84.76%。AI 系统(阈值 0.65)检出乳腺结节为偏良性 76 个、偏恶性 134 个,与术后病理结果对照,正确诊断良性结节 67 个、恶性结节 107 个,灵敏度为 92.24%,特异度为 71.28%,准确度为 82.86%。AI 系统(阈

值 0.70) 检出乳腺结节为偏良性 86 个、偏恶性 124 个、恶性结节 105 个, 灵敏度为 90.52%, 特异度为 79.79%, 准确度为 85.71% (表 1)。

表 1 不同超声诊断方法对最大径 ≤ 2 cm 的 BI-RADS 4 类乳腺结节的良恶性鉴别诊断效能

Tab 1 Diagnosis efficacy of benign and malignant BI-RADS class 4 breast nodules with maximum diameter ≤ 2 cm by different ultrasound diagnostic methods

Diagnostic method	Benign, <i>n</i>	Malignant, <i>n</i>	Sensitivity/%	Specificity/%	Accuracy/%
Physician routine ultrasound			92.24	75.53	84.76
Partial benign	71	9			
Partial malignant	23	107			
AI (threshold value 0.65)			92.24	71.28	82.86
Partial benign	67	9			
Partial malignant	27	107			
AI (threshold value 0.70)			90.52	79.79	85.71
Partial benign	75	11			
Suspected malignant	19	105			

BI-RADS: Breast imaging-reporting and data system; AI: Artificial intelligence.

2.3 常规超声和 AI 系统对 BI-RADS 4a、4b、4c 类结节良恶性的诊断价值 常规超声、AI 系统 (阈值 0.65)、AI 系统 (阈值 0.70) 对最大径 ≤ 2 cm 的 BI-RADS 4b、4c 类乳腺结节的良恶性诊断准确度均高于 BI-RADS 4a 类结节 ($\chi^2=10.564$ 、9.937、

9.661, 均 $P<0.001$)。AI 系统 (阈值 0.70) 对 BI-RADS 4a 类结节的诊断准确度 (79.41%) 高于常规超声 (77.94%) 和 AI 系统 (阈值 0.65) (75.00%)。见表 2。

表 2 不同超声诊断方法对最大径 ≤ 2 cm 的 BI-RADS 4a、4b、4c 类乳腺结节的诊断准确度

Tab 2 Diagnostic accuracy of different ultrasound diagnostic methods for BI-RADS class 4a, 4b, and 4c nodules with maximum diameter ≤ 2 cm in the breast

Diagnostic method	BI-RADS class 4a <i>N</i> =136			BI-RADS class 4b <i>N</i> =57			BI-RADS class 4c <i>N</i> =17		
	Benign, <i>n</i>	Malignant, <i>n</i>	Accuracy, <i>n</i> (%)	Benign, <i>n</i>	Malignant, <i>n</i>	Accuracy, <i>n</i> (%)	Benign, <i>n</i>	Malignant, <i>n</i>	Accuracy, <i>n</i> (%)
Physician routine ultrasound			106 (77.94)			55 (96.49)			17 (100.00)
Partial benign	71	7		0	2		0	0	
Partial malignant	23	35		0	55		0	17	
AI (threshold value 0.65)			102 (75.00)			55 (96.49)			17 (100.00)
Partial benign	67	7		0	2		0	0	
Partial malignant	27	35		0	55		0	17	
AI (threshold value 0.70)			108 (79.41)			55 (96.49)			17 (100.00)
Partial benign	75	9		0	2		0	0	
Suspected malignant	19	33		0	55		0	17	

BI-RADS: Breast imaging-reporting and data system; AI: Artificial intelligence.

2.4 常规超声和 AI 系统对不同病理类型结节良恶性的诊断价值 将 210 个乳腺结节术后病理结果按照第 5 版 WHO 乳腺肿瘤分类进行分组, 分析 3 种方法对最大径 ≤ 2 cm 的 BI-RADS 4 类乳腺结节的诊断价值, 结果如表 3 所示。在恶性结节中, AI 系统 (阈值 0.65) 对浸润性导管癌的诊断准确度最高, 达到了 96.9%; AI 系统 (阈值 0.65) 和 AI 系统 (阈值 0.70) 对乳腺导管原位癌的诊断准确度较

低, 均仅为 54.5%; 3 种方法对恶性结节的诊断准确度差异无统计学意义 ($P>0.05$)。在良性结节中, 常规超声及 AI 系统 (阈值 0.70) 对导管内乳头状瘤的诊断准确度最高, 均达到 90.9%; AI 系统 (阈值 0.65) 对乳腺炎性病变等其他良性疾病的诊断准确度较低, 仅为 50.0%; 3 种方法对良性结节的诊断准确度差异无统计学意义 ($P>0.05$)。

表 3 不同超声诊断方法对不同病理类型最大径≤2 cm 的 BI-RADS 4 类乳腺结节的诊断准确度

Tab 3 Diagnostic accuracy of different ultrasound diagnostic methods for BI-RADS class 4 breast nodules with a maximum diameter ≤2 cm of different pathological types

Diagnostic method	Benign				Malignant			n (%)
	Adenosis	Fibroadenoma	IDP	Others	IDC	DCIS	Others	
	N=36	N=31	N=11	N=16	N=98	N=11	N=7	
Physician routine ultrasound								
Benign	28 (77.8)	22 (71.0)	10 (90.9)	11 (68.8)	5 (5.1)	4 (36.4)	0	
Malignant	8 (22.2)	9 (29.0)	1 (9.1)	5 (31.2)	93 (94.9)	7 (63.6)	7 (100.0)	
AI (threshold value 0.65)								
Benign	27 (75.0)	23 (74.2)	9 (81.8)	8 (50.0)	3 (3.1)	5 (45.5)	1 (14.3)	
Malignant	9 (25.0)	8 (25.8)	2 (18.2)	8 (50.0)	95 (96.9)	6 (54.5)	6 (85.7)	
AI (threshold value 0.70)								
Benign	28 (77.8)	26 (83.9)	10 (90.9)	11 (68.8)	4 (4.1)	5 (45.5)	2 (28.6)	
Malignant	8 (22.2)	5 (16.1)	1 (9.1)	5 (31.2)	94 (95.9)	6 (54.5)	5 (71.4)	

BI-RADS: Breast imaging-reporting and data system; AI: Artificial intelligence; IDP: Intraductal papilloma; IDC: Invasive ductal carcinoma; DCIS: Ductal carcinoma *in situ*.

2.5 常规超声和 AI 系统对不同大小结节良恶性的诊断价值 将 210 个乳腺结节按照结节最大径分为≤1 cm 组、>1~1.5 cm 组和>1.5~2 cm 组, 分析 3 种方法对不同大小的 BI-RADS 4 类乳腺结节的诊断价值, 结果如表 4 所示。常规超声对最大径≤1 cm 的 BI-RADS 4 类乳腺结节的诊断准确度最高, 为 86.36%, AI 系统 (阈值 0.65) 和 AI 系统 (阈值

0.70) 诊断准确度分别为 81.82%、84.09%。在最大径>1~1.5 cm 和最大径>1.5~2 cm 的 BI-RADS 4 类乳腺结节中, AI 系统 (阈值 0.70)、AI 系统 (阈值 0.65) 及常规超声的诊断准确度均较高, 分别为 84.85% 和 87.00%、83.33% 和 83.00%、83.33% 和 85.00%, 三者之间差异均无统计学意义 (均 $P>0.05$)。

表 4 不同超声诊断方法对不同大小 BI-RADS 4 类乳腺结节的诊断准确度

Tab 4 Diagnostic accuracy of different ultrasound diagnostic methods for BI-RADS class 4 nodules of different sizes in the breast

Diagnostic method	≤1 cm N=44			>1-1.5 cm N=66			>1.5-2 cm N=100		
	Benign,	Malignant,	Accuracy,	Benign,	Malignant,	Accuracy,	Benign,	Malignant,	Accuracy,
	n	n	n (%)	n	n	n (%)	n	n	n (%)
Physician routine ultrasound									
Benign	27	2	38 (86.36)	22	3	55 (83.33)	22	4	85 (85.00)
Malignant	4	11		8	33		11	63	
AI (threshold value 0.65)									
Benign	25	2	36 (81.82)	23	4	55 (83.33)	19	3	83 (83.00)
Malignant	6	11		7	32		14	64	
AI (threshold value 0.70)									
Benign	27	3	37 (84.09)	24	4	56 (84.85)	24	4	87 (87.00)
Malignant	4	10		6	32		9	63	

BI-RADS: Breast imaging-reporting and data system; AI: Artificial intelligence.

3 讨论

超声作为一种实时、无创、无辐射、可重复的影像学诊断技术, 目前已广泛应用于乳腺疾病的

诊断及治疗评估^[9-10]。常规二维超声是诊断乳腺结节的首选检查方法, 但对于部分乳腺癌病灶二维超声图像并不典型, 加之超声诊断结果易受超声医师操作、工作经验等主观因素的影响, 在一定程度

上不可避免地影响临床诊疗决策,因此需要一种可靠、稳定的辅助超声诊断的方法^[11]。随着医学影像学技术的蓬勃发展, AI驱动的临床超声诊断方法日趋成熟,其诊断结果也越来越贴近病理结论,特别是超声 AI 诊断工具的出现,为甲状腺结节、乳腺结节等疾病的早期筛查和良恶性评估提供了有益补充^[12]。

文献报道部分乳腺癌的早期影像学特征不典型,超声图像上存在同图异病的可能,常规超声对乳腺小结节(≤ 2 cm)的诊断效能偏低^[13]。美国放射学会指南建议对 BI-RADS 分类 4 类以上的乳腺结节行穿刺活检或手术病理检查^[4],然而单纯依靠超声医师主观判断 BI-RADS 分类进行临床干预可能会增加不必要的有创操作,甚至造成不可逆性损伤。在本研究中,超声医师通过常规超声检查为 BI-RADS 4 类的 210 个乳腺结节中只有 55.2% (116/210) 经术后病理证实为恶性结节,这提示当前急需一种新的诊断方法来辅助医师鉴别 BI-RADS 4 类乳腺结节的性质,以尽早发现恶性病变、减少不必要的外科处理。

乳腺超声 AI 辅助诊断系统能够客观地识别分析乳腺结节的边界、形态、回声等声像图特征,自动做出“可能良性”或“可能恶性”的诊断,有助于提高超声医师的工作效率及有效鉴别乳腺良恶性结节^[14-15]。但是关于超声 AI 辅助诊断系统在最大径 ≤ 2 cm 的 BI-RADS 4 类乳腺结节中的研究报道少见。本研究将超声 AI 辅助诊断系统应用在最大径 ≤ 2 cm 的 BI-RADS 4 类乳腺小结节中,通过高年资超声医师常规超声检查和应用不同阈值(0.65、0.70)的 AI 诊断系统检测,对比三者对该类乳腺结节的诊断效能,结果提示 AI 系统(阈值 0.70)诊断乳腺结节良恶性的特异度、准确度均优于高年资超声医师常规超声。这可能是因为 AI 系统通过深度学习,能够精准提取乳腺结节的细微形态特征,如边缘毛刺、微钙化等,较超声医师肉眼识别结节更为灵敏和客观。

本研究对 BI-RADS 4a、4b、4c 类乳腺结节进行分析,发现 AI 系统(阈值 0.70)对 BI-RADS 4a 类结节的诊断准确度高于高年资超声医师常规超声和 AI 系统(阈值 0.65),高年资超声医师常规超声、AI 系统(阈值 0.65)及 AI 系统(阈值 0.70)对最大径 ≤ 2 cm 的 BI-RADS 4b 类乳腺及 BI-RADS

4c 类结节中恶性结节均有较高诊断价值,提示日后可通过合理设置诊断阈值来提高对不同类别 BI-RADS 4 类结节的诊断准确度。本研究还比较了高年资超声医师常规超声、AI 系统(阈值 0.65)、AI 系统(阈值 0.70) 3 种检查方法对不同病理类型乳腺小结节的诊断价值,结果表明,在恶性结节中,3 种方法对浸润性导管癌的诊断准确度较高、对乳腺导管原位癌的准确度较低,3 种方法之间差异无统计学意义;在良性结节中,3 种方法对导管内乳头状瘤的诊断准确度较高、对乳腺炎性病变等其他良性病变的诊断准确度较低,3 种方法之间差异无统计学意义。这一结果提示无论是超声医师检查还是 AI 系统检测,诊断的准确度均与病理类型相关,与检查方法无关。另外,本研究对不同大小的 BI-RADS 4 类乳腺结节进行了分组分析,结果显示高年资超声医师常规超声对判断最大径 ≤ 1 cm 的乳腺 BI-RADS 4 类结节有较高价值,提示对于小结节的诊断不能仅依靠 AI 辅助诊断系统,还要结合高年资超声医师的经验以提高对小结节的诊断准确度。

本研究中, AI 系统正确诊断了 1 例最大径为 0.6 cm 的乳腺恶性结节,而超声医师常规检查后考虑其为良性,分析原因为乳腺结节较小,二维灰阶图像上肉眼所见的恶性征象并不明显,超声医师诊断结果偏保守,造成假阴性误诊;而 AI 系统通过计算机视觉技术提取了该结节的边缘及内部细微特征,确定其二维图像中存在可疑恶性特征,不受主观认知的影响,因此做出了符合病理判断的判断,证明 AI 系统检测可提高对小体积乳腺癌的检出率。

AI 系统诊断 BI-RADS 4 类乳腺结节也存在不足之处。例如, AI 系统将 1 例最大径为 1.6 cm 的乳腺结节误诊为偏可疑恶性结节,而高年资超声医师诊断该结节为良性结节,术后病理结果为乳腺组织急慢性炎。分析原因可能为该患者为年轻女性,因乳腺肿痛就诊,超声图像虽表现为边界模糊、形态不规则等恶性结节征象,但超声医师结合患者病史和临床体征考虑该病灶为良性病变的可能性大;而 AI 系统仅依靠当时提取的乳腺二维图像特征,直接误判为可疑恶性结节。因此, AI 系统的诊断主要依赖于图像信息,却不能结合患者的临床资料及病史特点,在临床应用时仅可作为超声诊断的有效补充和辅助工具。

本研究存在以下局限性: (1) 本研究的病例数偏少。(2) 纳入研究的超声图像选择会受到超声医师对图像认知程度的影响, 所以图像的留存与分析不可避免存在可变性。(3) 本研究中乳腺超声 AI 辅助诊断系统对于乳腺肿块周边组织的特点、腋窝淋巴结是否增大及其他特殊情况的分析能力尚有欠缺。

综上所述, 乳腺超声 AI 辅助诊断系统对最大径 ≤ 2 cm 的 BI-RADS 4 类乳腺结节的诊断结果与病理结果一致性较好, 其诊断效能与高年资超声医师接近。乳腺超声 AI 辅助诊断系统在一定程度上实现了超声诊断的量化及标准化, 可辅助超声医师快速、无创地判断乳腺结节的良恶性, 减少因经验技能不足导致的误诊及漏诊, 具有一定的临床实用价值。

[参 考 文 献]

- [1] SIEGEL R L, MILLER K D, JEMAL A. Cancer statistics, 2018[J]. *CA Cancer J Clin*, 2018, 68(1): 7-30. DOI: 10.3322/caac.21442.
- [2] 郑荣寿, 孙可欣, 张思维, 等. 2015 年中国恶性肿瘤流行情况分析[J]. *中华肿瘤杂志*, 2019, 41(1): 19-28. DOI: 10.3760/cma.j.issn.0253-3766.2019.01.008.
- [3] 中国抗癌协会乳腺癌专业委员会. 中国抗癌协会乳腺癌诊治指南与规范(2021年版)[J]. *中国癌症杂志*, 2021, 31(10): 954-1040. DOI: 10.19401/j.cnki.1007-3639.2021.10.013.
- [4] LEE J. Practical and illustrated summary of updated BI-RADS for ultrasonography[J]. *Ultrasonography*, 2017, 36(1): 71-81. DOI: 10.14366/usg.16034.
- [5] UEHIRO N, HORII R, IWASE T, et al. Validation study of the UICC TNM classification of malignant tumors, seventh edition, in breast cancer[J]. *Breast Cancer*, 2014, 21(6): 748-753. DOI: 10.1007/s12282-013-0453-7.
- [6] 孙德胜, 刘俐, 石宇, 等. “全流程超声质控管理”软件系统(V1.0版)应用与实践[J]. *中华医学超声杂志(电子版)*, 2019, 16(5): 329-332. DOI: 10.3877/cma.j.issn.1672-6448.2019.05.003.
- [7] 郭万学. 超声医学[M]. 6版. 北京: 人民军医出版社, 2011: 263-278.
- [8] WHO Classification of Tumours Editorial Board. Breast tumours[M]. 5th ed. Lyon (France): International Agency for Research on Cancer, 2019.
- [9] GUO R, LU G, QIN B, et al. Ultrasound imaging technologies for breast cancer detection and management: a review[J]. *Ultrasound Med Biol*, 2018, 44(1): 37-70. DOI: 10.1016/j.ultrasmedbio.2017.09.012.
- [10] 马骥, 张艳, 陈青青, 等. 常规超声联合 ABVS 在乳腺 BI-RADS 4 类小结节诊断中的价值[J]. *同济大学学报(医学版)*, 2021, 42(6): 825-830. DOI: 10.12289/j.issn.1008-0932.21378.
- [11] 赵佳琦, 刁宗平, 徐琪, 等. 人工智能时代超声医学新发展[J]. *第二军医大学学报*, 2019, 40(5): 478-482. DOI: 10.16781/j.0258-879x.2019.05.0478.
- [12] ZHAO J Q, DIAO Z P, XU Q, et al. New development of ultrasound medicine in the era of artificial intelligence[J]. *Acad J Sec Mil Med Univ*, 2019, 40(5): 478-482. DOI: 10.16781/j.0258-879x.2019.05.0478.
- [13] AKKUS Z, CAI J, BOONROD 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. DOI: 10.1016/j.jacr.2019.06.004.
- [14] 刘洁, 王荣, 王兴田, 等. 声触诊组织成像及声触诊组织成像量化技术在乳腺实性小结节鉴别诊断中的应用[J]. *中国临床医学影像杂志*, 2019, 30(7): 463-467. DOI: 10.12117/jccmi.2019.07.002.
- [15] 杜建文, 葛雪, 刘欢颜, 等. S-Detect 技术与 BI-RADS 分类对乳腺肿物良恶性评估的比较[J]. *中国超声医学杂志*, 2019, 35(7): 580-583. DOI: 10.3969/j.issn.1002-0101.2019.07.003.
- [16] XING B, CHEN X, WANG Y, et al. Evaluating breast ultrasound S-detect image analysis for small focal breast lesions[J]. *Front Oncol*, 2022, 12: 1030624. DOI: 10.3389/fonc.2022.1030624.

[本文编辑] 杨亚红