

良恶性结节超声乳腺影像报告及数据系统量化评分分类标准模型建立的初步研究

李剑锋 赵旭晔 薛改琴

摘要 **目的** 初步建立良恶性结节超声乳腺影像报告和数据系统(US-BI-RADS)量化评分分类标准模型,探讨其应用价值。**方法** 收集在我院就诊的3589例女性乳腺肿瘤患者资料,以2018年1~12月的3623个结节为模型病例,分析乳腺结节的声像图特征,包括形态是否规则、边缘是否清晰、结节内部回声情况、后方回声是否衰减、纵横比是否大于1、病灶内是否伴微钙化及相邻导管是否扩张共7个超声征象。计算7个超声征象对乳腺良恶性结节的鉴别诊断价值,应用 Logistic 回归分析各征象诊断乳腺恶性结节的比值比(OR),根据各征象的权重给予相应的赋值评分,再结合患者年龄赋值评分,建立乳腺良恶性结节US-BI-RADS量化评分分类标准模型,绘制受试者工作特征(ROC)曲线分析该模型的诊断效能。以ACR BI-RADS为标准,提出US-BI-RADS分类为3类、4类(4a类、4b类、4c类)及5类的评分标准,以2019年1月的386个结节为测试病例,分析该模型评估结节恶性构成比情况。**结果** 乳腺良恶性结节形态、边缘、内部回声、后方回声、纵横比、是否存在微钙化和导管扩张比较差异均有统计学意义(均 $P<0.05$)。7个超声征象的OR值从大到小依次为:后方回声衰减、边缘模糊或微小分叶或成角或毛刺征、内部回声不均匀或混合回声、纵横比 >1 、形态不规则、相邻导管扩张、病灶内伴微钙化(OR=181.650、37.552、22.855、18.678、13.700、12.609、11.134),结合年龄赋值评分,建立US-BI-RADS量化评分分类标准模型:总评分=年龄评分+ $X_4*3+(X_2+X_3+X_5)*2+(X_1+X_6+X_7)*1$, X_1 、 X_2 、 X_3 、 X_4 、 X_5 、 X_6 、 X_7 分别为形态不规则、边缘模糊或微小分叶或成角或毛刺征、内部回声不均匀或混合回声、后方回声衰减、纵横比 >1 、病灶内伴微钙化、相邻导管扩张。ROC曲线分析显示,US-BI-RADS量化评分分类标准模型取5分为鉴别诊断乳腺良恶性结节的截断值,其曲线下面积0.985(95%可信区间0.981~0.989),敏感性94.76%,特异性99.18%,准确率97.29%,阳性预测值98.85%,阴性预测值96.21%。根据US-BI-RADS量化评分分类标准模型评估测试病例3类、4a类、4b类、4c类、5类结节恶性构成比依次为1.69%、6.06%、23.53%、90.91%、97.96%。**结论** 建立的US-BI-RADS量化评分分类标准模型可行性较高,能准确定性评估乳腺良恶性结节,具有较好的临床应用价值。

关键词 超声影像报告及数据系统;乳腺肿瘤,良恶性;量化评分

[中图分类号]R445.1;R737.9

[文献标识码]A

Preliminary study on establishing the ultrasound breast imaging reporting and data system quantitative score classification standard model for benign and malignant nodules

LI Jianfeng, ZHAO Xuye, XUE Gaiqin

School of Medical Imaging, Shanxi Medical University, Taiyuan 030001, China

ABSTRACT **Objective** To preliminarily establish the ultrasound breast imaging reporting and data system(US-BI-RADS) quantitative score classification standard model for benign and malignant nodules, and to explore its application value. **Methods** The data of 3589 cases of female breast tumor patients who were treated in our hospital were collected, and 3623 nodules from January to December 2018 were used as model cases to analyze the sonographic features of breast nodules, including whether the shape was regular, whether the edge was clear, the nodules internal echo, whether the posterior echo of tumor was attenuated, whether the aspect ratio was greater than 1, whether the microcalcification in the nodules, and whether the adjacent duct was dilated, a total of 7 ultrasonic signs. The differential diagnostic value of benign and malignant breast nodules by

基金项目:山西省科学技术厅重点研发社会发展项目(201803D32111161041009421748198871)

作者单位:030001 太原市,山西医科大学医学影像学院(李剑锋);山西省肿瘤医院乳腺二科(赵旭晔),超声科(薛改琴)

通讯作者:薛改琴,Email:xuegaiqin163@163.com

7 ultrasonic signs were calculated. Logistic regression was used to analyze the odds ratio (*OR*) of each sign in diagnosing breast malignant nodules, and the corresponding assigned score was given according to the weight of each sign, and then combined with the patient's age assignment score, the US-BI-RADS quantitative score classification standard model of benign and malignant breast nodules was established. Receiver operating characteristic (ROC) curve was drawn to analyze the diagnostic efficacy of the model. Taking ACR BI-RADS as the standard, the 3 category, 4 categories (4a, 4b, 4c) and category 5 scoring standards of US-BI-RADS quantitative score classification standard model were proposed, with 386 nodules in January 2019 as test cases, the composition ratio of malignant nodules evaluated by the model was analyzed. **Results** There were statistically significant differences in the morphology, edge, internal echo, posterior echo, aspect ratio, microcalcification and ductal dilatation of benign and malignant breast nodules (all $P < 0.05$). The *OR* values of 7 ultrasonic signs in descending order were: posterior echo attenuation, blurred edges or micro-lobed or angled or burr sign, internal echo inhomogeneous or mixed echo, aspect ratio > 1, irregular shape, adjacent duct dilatation and internal microcalcification ($OR = 181.650, 37.552, 22.855, 18.678, 13.700, 12.609, 11.134$), combined with age assignment score, US-BI-RADS quantitative score classification standard model was established, total score = age score + $X_4 * 3 + (X_2 + X_3 + X_5) * 2 + (X_1 + X_6 + X_7) * 1$, $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ were irregular shape, blurred edge or micro-lobed or angled or burr sign, internal echo inhomogeneous or mixed echo, posterior echo attenuation, aspect ratio > 1, internal microcalcification, adjacent duct dilatation, respectively. ROC curve analysis showed that the US-BI-RADS quantitative score classification standard model took 5 points as the cut-off value for the differential diagnosis of benign and malignant breast nodules, the area under the curve was 0.985 (95%CI 0.981~0.989), the sensitivity was 94.76%, the specificity was 99.18%, accuracy was 97.29%, the positive predictive value was 98.85%, and the negative predictive value was 96.21%. After classified according to the US-BI-RADS quantitative score classification standard model, the malignant nodules composition ratios of the test cases 3, 4a, 4b, 4c and 5 categories were 1.69%, 6.06%, 23.53%, 90.91%, 97.96%, respectively. **Conclusion** US-BI-RADS quantitative score classification standard model is highly feasible, which can accurately and quantitatively evaluate benign and malignant breast nodules, and has good clinical application value.

KEY WORDS Ultrasound breast imaging reporting and data system; Breast tumor, benign and malignant; Quantitative score

乳腺癌是女性主要恶性肿瘤之一,近年发病呈年轻化趋势,故早期检出并及时治疗至关重要,超声已成为乳腺肿瘤早期筛查的首选方法。美国放射学会超声乳腺影像报告和数据系统(ACR BI-RADS)描述了乳腺良恶性结节的声像图特征,以及各类结节理论上的恶性可能性和相应处理意见^[1-2],有利于临床医师对结节做出合适处理。但 ACR BI-RADS 分类中缺少客观的评分标准,超声医师在临床实践中对乳腺良恶性结节进行 BI-RADS 分类时具有主观性,因此建立良恶性结节超声乳腺影像报告和数据系统(US-BI-RADS)量化评分分类标准十分必要。本研究初步建立良恶性结节 US-BI-RADS 量化评分分类标准模型,并探讨其临床应用价值。

资料与方法

一、研究对象

选取 2018 年 1 月至 2019 年 1 月在我院经术后病理证实的乳腺结节患者 3589 例,均为女性。以 2018 年 1~12 月的 3314 例患者为模型病例,年龄 11~87 岁,平均(42.5±12.6)岁,共计 3623 个结节,位于左侧 1785 个,右侧 1838 个;以 2019 年 1 月的 275 例患者为测试病例,年龄 11~79 岁,平均(40.5±11.7)岁,共计 386 个结

节,位于左侧 189 个,右侧 197 个。纳入标准:①经常规超声检出乳腺结节,检查资料完整;②均行外科手术或穿刺活检术,病理结果完整。排除标准:①无完整病理结果者;②超声图像质量不符合要求者。本研究经山西省肿瘤医院医学伦理委员会批准,所有患者均知情同意。

二、仪器与方法

1. 超声检查:使用 Philips EPIQ 7、GE Logiq E 9 及西门子 Oxana 2 彩色多普勒超声诊断仪,线阵探头,频率 5~18 MHz。患者取平卧位,双手上举平放于头部下方,若患者乳腺较大,则取左侧卧、右侧卧或坐位进行扫查,探头垂直置于乳腺皮肤表面,尽量不加压。常规扫查并观察:①乳腺整体位置、对称性、内部回声、内部结构、脂肪层和腺体的比值、乳管情况、乳头和乳晕区、实质血管;②乳腺内部改变特征、病变数目、位置、形态、内部回声、病灶内部结构、轮廓;③乳腺结节与周边结构的关系。

2. 图像分析:观察乳腺结节形态是否规则、边缘是否清晰(模糊或微小分叶或成角或毛刺征)、结节内部回声情况(不均匀或混合回声)、后方回声是否衰减、纵横比是否大于 1、病灶内是否伴微钙化及相邻导管是否扩张共 7 个征象。

将 ACR BI-RADS 分类 3 类及以下结节判为病理良性, 4 类及以上结节判为病理恶性。以 ACR BI-RADS 为标准, 提出根据 US-BI-RADS 量化评分分类标准模型分类为 3 类、4 类(4a 类、4b 类、4c 类)及 5 类的评分标准, 计算分类后测试病例结节恶性构成比。

三、统计学处理

应用 SPSS 20.0 统计软件, 计数资料以频数或率表示, 行 χ^2 检验。计算 7 个超声征象对模型病例良恶性结节的鉴别诊断效能。应用 Logistic 回归分析各征象诊断乳腺恶性结节的比值比(OR), 根据各征象的权重给予相应赋值评分, 再结合患者年龄赋值评分, 建立良恶性结节 US-BI-RADS 量化评分分类标准模型, 绘制受试者工作特征(ROC)曲线分析其对乳腺良恶性结节的诊断效能。 $P < 0.05$ 为差异有统计学意义。

结 果

一、病理结果

模型病例对应的 3623 个结节中, 良性结节 2076 个, 包括腺病或纤维腺瘤 1656 个, 导管内乳头状瘤 193 个, 不典型增生 69 个, 炎症性病变 64 个, 良性叶状肿瘤 48 个, 硬化性腺病 46 个; 恶性结节 1547 个, 包括浸润性导管癌 1246 个, 浸润性小叶癌 133 个, 原位癌 80 个, 乳头状癌 31 个, 黏液癌 30 个, 导管内癌 11 个, Paget 病 10 个, 小管癌 5 个, 神经内分泌癌 1 个。测试病例对应的 386 个结节中, 良性结节 223 个, 包括腺病及或纤维腺瘤 115 个, 导管内乳头状瘤 48 个, 炎症性病变 24 个, 良性叶状肿瘤 12 个, 不典型增生 15 个, 硬化性

腺病 9 个; 恶性结节 163 个, 包括浸润性导管癌 86 个, 浸润性小叶癌 45 个, 原位癌 16 个, 乳头状癌 7 个, 导管内癌 5 个, 小管癌 4 个。

二、模型病例良恶性结节超声征象及诊断效能分析

模型病例良恶性结节形态、边缘、内部回声、后方回声、纵横比、是否存在微钙化和导管扩张比较差异均有统计学意义(均 $P < 0.05$)。见表 1。上述 7 个超声征象对乳腺结节良恶性的鉴别诊断效能见表 2。

三、Logistic 回归分析

7 个超声征象赋值情况见表 3。Logistic 回归分析结果见表 4。各超声征象诊断乳腺恶性结节的 OR 值从大到小依次为后方回声衰减、边缘模糊或微小分叶或成角或毛刺征、内部回声不均匀或混合回声、纵横比 > 1 、形态不规则、相邻导管扩张、病灶内伴微钙化。各超声征象的量化评分标准见表 5。将年龄 < 35 岁者赋值 0 分, 年龄 35~50 岁者赋值 1 分, 年龄 ≥ 50 岁者赋值 2 分, 建立 US-BI-RADS 量化评分分类标准模型: 总评分 = 年龄评分 + $X_4 * 3 + (X_2 + X_3 + X_5) * 2 + (X_1 + X_6 + X_7) * 1$ 。 X_1 、 X_2 、 X_3 、 X_4 、 X_5 、 X_6 、 X_7 分别为形态不规则、边缘模糊或微小分叶或成角或毛刺征、内部回声不均匀或混合回声、后方回声衰减、纵横比 > 1 、病灶内伴微钙化、相邻导管扩张。见图 1。

四、ROC 曲线分析

以 US-BI-RADS 量化评分 5 分为鉴别诊断乳腺良恶性结节的截断值, ROC 曲线下面积 0.985(95% 可信区间 0.981~0.989), 敏感性 94.76%, 特异性 99.18%, 准确率 97.29%, 阳性预测值 98.85%, 阴性预测值 96.21%。见图 2。

表 1 模型病例良恶性结节超声征象比较

病理结果	形态		边缘		内部回声		后方回声		纵横比		微钙化		相邻导管扩张	
	规则	不规则	清晰	模糊	均匀	不均匀	衰减	无衰减	> 1	≤ 1	是	否	有	无
良性(2076)	1968	108	1921	155	1784	292	2	2074	6	2070	140	1936	37	2039
恶性(1547)	236	1311	81	1466	80	1467	381	1166	163	1384	433	1114	43	1504
χ^2 值	2353.905		2732.406		2314.734		564.296		209.321		300.526		4.083	
P值	0.000		0.000		0.000		0.000		0.000		0.000		0.043	

表 2 各超声征象鉴别诊断模型病例良恶性结节的效能分析

超声征象	敏感性(%)	特异性(%)	准确率(%)	阳性预测值(%)	阴性预测值(%)	阳性似然比	阴性似然比
形态不规则	84.74	94.79	90.51	92.39	89.29	16.26	0.16
边缘模糊或微小分叶或成角或毛刺征	94.76	92.53	93.49	90.41	95.95	12.69	0.06
内部回声不均匀或混合回声	94.83	85.93	89.73	83.40	95.71	6.74	0.06
后方回声衰减	24.63	99.90	67.76	99.48	64.01	246.30	0.75
纵横比 > 1	10.54	99.71	61.63	96.45	59.93	36.34	0.89
病灶内伴微钙化	27.99	93.26	65.39	75.57	63.48	4.15	0.77
相邻导管扩张	2.59	97.74	57.11	45.98	57.38	1.15	0.99

表3 超声征象赋值情况

超声征象	变量	赋值说明
形态不规则	X_1	无=0, 有=1
边缘模糊或微小分叶或成角或毛刺征	X_2	无=0, 有=1
内部回声不均匀或混合回声	X_3	无=0, 有=1
后方回声衰减	X_4	无=0, 有=1
纵横比>1	X_5	无=0, 有=1
病灶内伴微钙化	X_6	无=0, 有=1
相邻导管扩张	X_7	无=0, 有=1

表4 超声征象的Logistic 回归分析结果

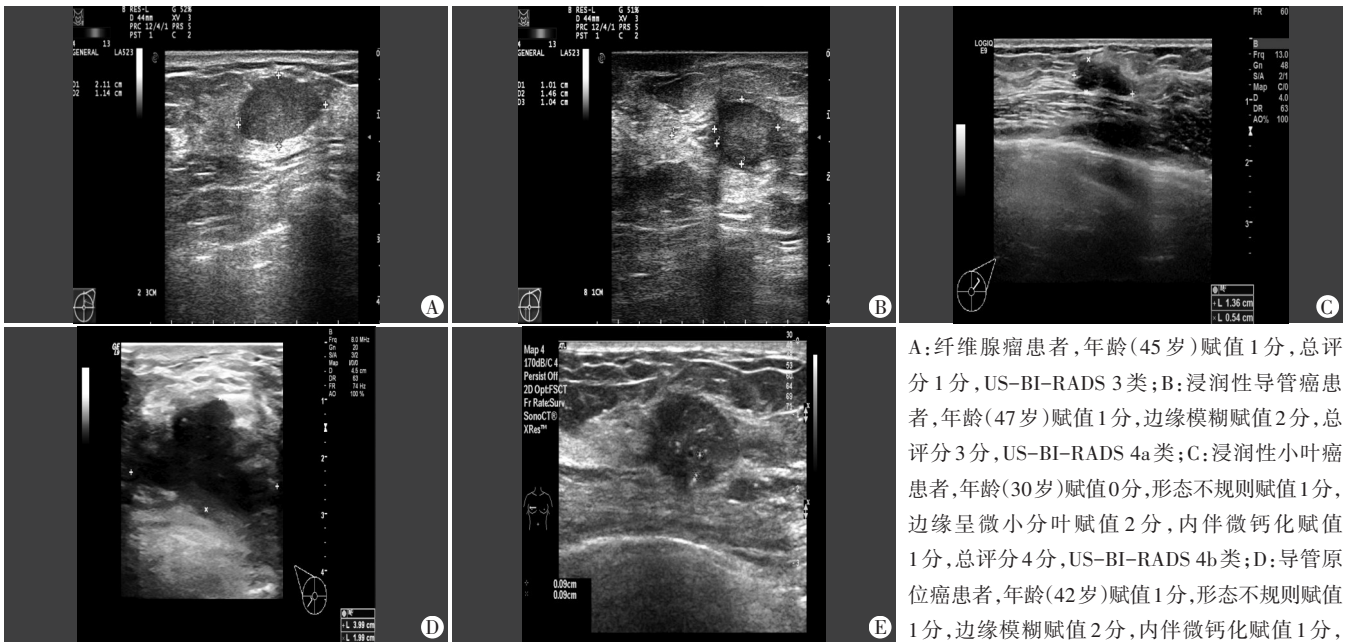
变量	回归系数	标准误	Wald χ^2 值	P值	OR值
形态不规则	2.617	0.238	120.640	0.000	13.700
边缘模糊或微小分叶或成角或毛刺征	3.626	0.205	313.803	0.000	37.552
内部回声不均匀或混合回声	3.129	0.201	242.041	0.000	22.855
后方回声衰减	5.202	0.947	30.179	0.000	181.650
纵横比>1	2.927	1.080	7.351	0.007	18.678
病灶内伴微钙化	2.410	0.298	65.340	0.000	11.134
相邻导管扩张	2.534	0.506	25.097	0.000	12.609

表5 各超声征象的量化评分标准

超声征象	评分
后方回声衰减	3
边缘模糊或微小分叶或成角或毛刺征	2
内部回声不均匀或混合回声	2
纵横比>1	2
形态不规则	1
病灶内伴微钙化	1
相邻导管扩张	1

五、US-BI-RADS 量化评分分类标准模型评估测试病例结节恶性构成比

设总评分 ≤ 2 分为US-BI-RADS 3类结节,总评分3分为4a类结节,总评分4分为4b类结节,总评分5分为4c类结节,总评分 ≥ 6 分为5类结节。根据US-BI-RADS 量化评分分类标准模型评估测试病例各类结节的恶性构成比见表6。



A: 纤维腺瘤患者,年龄(45岁)赋值1分,总评分1分,US-BI-RADS 3类;B: 浸润性导管癌患者,年龄(47岁)赋值1分,边缘模糊赋值2分,总评分3分,US-BI-RADS 4a类;C: 浸润性小叶癌患者,年龄(30岁)赋值0分,形态不规则赋值1分,边缘呈微小分叶赋值2分,内伴微钙化赋值1分,总评分4分,US-BI-RADS 4b类;D: 导管原位癌患者,年龄(42岁)赋值1分,形态不规则赋值1分,边缘模糊赋值2分,内伴微钙化赋值1分,总评分5分,US-BI-RADS 4c类;E: 浸润性导管癌(II级)患者,年龄(52岁)赋值2分,边缘呈微小分叶赋值2分,内部回声不均匀赋值2分,内伴微钙化赋值1分,总评分7分,US-BI-RADS 5类

图1 乳腺良恶性结节超声图像

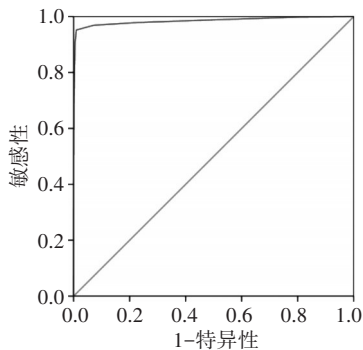


图2 US-BI-RADS 量化评分模型病例鉴别诊断分类标准模型良恶性结节的ROC 曲线图

表6 US-BI-RADS 量化评分分类标准模型评估测试病例结节恶性构成比

总评分	US-BI-RADS 分类	恶性 (个)	良性 (个)	合计 (个)	恶性构成比 (%)
≤ 2 分	3	3	175	178	1.69
3分	4a	2	31	33	6.06
4分	4b	4	13	17	23.53
5分	4c	10	1	11	90.91
≥ 6 分	5	144	3	147	97.96

讨 论

乳腺癌的早期检出和治疗对患者预后评估十分重要,超声在鉴别乳腺良恶性结节方面具有分辨率高、无创、操作简便等优势。为早期筛查乳腺肿瘤并规范随访监测,帮助临床制定合理的治疗方案,美国放射学会于2013年发布了第2版ACR BI-RADS,但因其未明确界定每个术语的良恶性含义及其在评估分类中的作用,导致不同医师在术语的理解上可能存在偏差,对BI-RADS 4类各亚类的诊断一致性较差,尤其是对于BI-RADS 4c类结节的诊断,反映了分类的多变性和主观性。本研究采用定量评分的方法初步建立良恶性结节US-BI-RADS量化评分分类标准模型,探讨该模型的临床应用价值。

本研究结果发现乳腺良恶性结节在形态、边缘、内部回声、后方回声、纵横比、是否存在微钙化和导管扩张等方面比较差异均有统计学意义(均 $P < 0.05$),原因是上述超声征象与病灶内部病理结构、病变是否侵犯周围组织密切相关。本研究进一步将上述差异有统计学意义的超声征象(形态不规则、边缘模糊或微小分叶或成角或毛刺征、内部回声不均匀或混合回声、后方回声衰减、纵横比 > 1 、病灶内伴微钙化、相邻导管扩张)行Logistic回归分析,发现其OR值分别为13.700、37.552、22.855、181.650、18.678、11.134、12.609。有文献^[3]发现 < 45 岁、 $45 \sim 59$ 岁和 > 59 岁组中BI-RADS 3类结节的阳性预测值比较差异有统计学意义($P = 0.048$),提示随年龄增加,BI-RADS 4a类和4b类乳腺结节的恶性可能有所提高。Patterson等^[4]研究显示,随着年龄增长,4a类结节的恶性率也会升高,特别是40岁以上的女性。Yoon等^[5]研究认为,若4a类乳腺结节患者年龄 > 40 岁,恶性的可能性更大。因此临床进行BI-RADS分类时需要考虑患者的年龄因素,对于年龄较大者(> 35 岁),适当提高分类级别有助于提高诊断准确率。本研究纳入患者年龄进行赋值评分,建立了US-BI-RADS量化评分分类标准模型:总评分=年龄评分 $+ X_4 * 3 + (X_2 + X_3 + X_5) * 2 + (X_1 + X_6 + X_7) * 1$ 。

国外学者^[5]依据乳腺结节超声征象与乳腺癌的OR值,建立US-BI-RADS量化评分分类方法,结果显示其评估乳腺良恶性结节的ROC曲线下面积为0.947~0.986。本研究ROC曲线分析显示,依据患者年龄及乳腺结节超声征象建立的US-BI-RADS量化评分分类标准模型评估乳腺结节良恶性的曲线下面积为0.985(95%可信区间0.981~0.989),以评分5分为截断值,其鉴别诊断乳腺良恶性病灶的敏感性、特异性、准确率、阳性预测值、阴性预测值分别为94.76%、

99.18%、97.29%、98.85%、96.21%。与上述文献结果相似。根据US-BI-RADS量化评分分类标准模型评估测试病例各类结节的恶性构成比均在ACR US-BI-RADS的理论范围之内,表明US-BI-RADS量化评分分类标准模型可以在一定程度上弥补ACR US-BI-RADS缺少量化评分分类标准的不足,减少超声医师评估乳腺结节时主观因素的影响,具有较好的应用价值。常规超声是基层医院检查乳腺结节的重要方法,部分早期乳腺癌患者在基层医院容易漏误诊,延误最佳治疗时机。本研究建立的US-BI-RADS量化评分分类标准模型依据患者的年龄及乳腺结节的常规超声征象进行评分,评价指标较客观,操作简便,能够在各级医院中开展,尤其适用于基层医院,可明显提高乳腺癌的诊断准确率。

本研究的不足:①为回顾性研究,纳入病例为手术切除或行活检术的患者,未涉及一些更可能是良性结节的3类回访病例;②纳入对象均为肿瘤医院病例;③仅加入了患者年龄这一临床指标。今后需进一步开展多中心、大样本的研究,加入家族史、病史、超声弹性成像等资料完善量化评分分类标准模型。

综上所述,随着超声医学新技术的发展,超声造影、超声弹性成像及自动乳腺全容积成像在BI-RADS分类中的应用越来越广泛,在此基础上尝试建立US-BI-RADS量化评分分类标准,量化各指标在乳腺结节良恶性风险评估中的重要性,有利于临床医师准确性评估乳腺良恶性结节,制定后续治疗方案,具有较好的临床应用价值。

参考文献

- [1] Rao AA, Feneis J, Lalonde C, et al. A pictorial review of changes in the BI-RADS Fifth Edition [J]. Radiographics, 2016, 36(3): 623-639.
- [2] Spak DA, Plaxco JS, Santiago L, et al. BI-RADS Fifth Edition: a summary of changes [J]. Diagn Interv Imaging, 2017, 98(3): 179-190.
- [3] Fu CY, Hsu HH, Yu JC, et al. Influence of age on PPV of sonographic BI-RADS categories 3, 4, and 5 [J]. Ultrason Med, 2011, 32(Suppl 1): 8-13.
- [4] Patterson SK, Neal CH, Jeffries DO, et al. Outcomes of solid palpable masses assessed as BI-RADS 3 or 4A: a retrospective review [J]. Breast Cancer Res Treat, 2014, 147(2): 311-316.
- [5] Yoon JH, Kim MJ, Moon HJ, et al. Subcategorization of ultrasonographic BI-RADS category 4: positive predictive value and clinical factors affecting it [J]. Ultrasound Med Biol, 2011, 37(5): 693-699.

(收稿日期:2021-02-22)