·临床研究。

早孕期胎儿颜面部超声筛查切面图像质量评价系统的 构建及验证

余卫峰 薛 浩 刘中华 柳培忠 郭 旭 吴秀明 吕国荣

摘 要 目的 构建一种基于YOLOv4模型的早孕期胎儿颜面部超声筛查切面自动识别及图像质量评价系统,并验证该模型的识别效能。方法 收集 2021年1月至 2022年12月我院妊娠11~14周胎儿颜面部超声筛查切面图像为研究对象,分为模型构建集(1001张图像)和临床验证集(610张图像),其中模型构建集按8:1:1分为训练集、测试集和验证集,用于训练、测试和验证人工智能(AI)模型对正中矢状面(MSP)、鼻后三角切面(RNT)的分类与识别,临床验证集用于评价 AI模型的临床泛化能力。采用 Kappa 检验及 Kendall 协同系数分析 AI模型与专家级超声医师识别 MSP 和 RNT 及评价图像质量等级的一致性。结果 成功构建了自动识别早孕期胎儿颜面部超声筛查切面及图像质量评价的 AI模型,该模型对测试集图像中各关键解剖结构的查准率、查全率和 F1分数均>80%。AI模型识别验证集图像中 MSP 和 RNT 的准确率分别为 94.9%、92.9%,与专家级超声医师分类的一致性均好(Kappa=0.888、0.810);AI模型识别临床验证集图像中MSP 和 RNT 的准确率分别为 94.0%、84.4%,与专家级超声医师分类的一致性均良好(Kappa=0.751、0.690)。AI模型对临床验证集图像中 MSP 和 RNT 图像质量评价结果与专家级超声医师评价结果的总体一致性均良好(Kendall's W=0.760、0.789)。结论 应用 AI模型识别早孕期颜面部超声筛查切面具有较高的准确性,可作为初级医师及基层医师培训和图像质量评价的辅助方法。

关键词 超声检查;早孕期;筛查切面,颜面部;人工智能;质量控制 [中图法分类号]R445.1;R714.5 [文献标识码]A

Construction and validation of the quality evaluation system for fetal facial ultrasound screening in early pregnancy

YU Weifeng, XUE Hao, LIU Zhonghua, LIU Peizhong, GUO Xu, WU Xiuming, LV Guorong Department of Ultrasound, Quanzhou First Hospital Affiliated to Fujian Medical University, Fujian 362000, China

ABSTRACT Objective To construct an automatic recognition and quality evaluation system for fetal facial ultrasound screening sections in early pregnancy based on the YOLOv4 model, and to verify the recognition performance of the model. Methods Fetal facial ultrasound screening sections of early pregnancy at 11~14 weeks of gestation in our hospital from January 2021 to December 2022 were collected as research objects, including 1001 images of model construction set and 610 images of clinical validation set. The model construction set was divided into training set, test set and validation set according to the ratio of 8:1:1, which was used to train and verify the classification and recognition of the midstagittal plane (MSP) and retronasal triangle (RNT) of the artificial intelligence (AI) model, and the clinical validation set was used to test the clinical generalization ability of the AI model. Kappa test and Kendall coefficient of concordance were used to analyze the consistency of the AI model and expert—level sonographers in identifying MSP and RNT in facial ultrasound screening sections of early pregnancy and evaluating the image quality level. Results The AI model of automatic recognition and quality evaluation system for fetal facial ultrasound screening sections in early pregnancy was successfully constructed. The precision, recall and F1 score of the model for key anatomical structures in the test set were all higher than 80%. The accuracy of MSP and RNT of AI model recognition

基金项目:福建省自然科学基金项目(2021J011404、2023J011773);泉州市科技计划项目(2022NS057)

作者单位:362000 福建省泉州市,福建医科大学附属泉州第一医院超声科(余卫峰、刘中华、郭旭、吴秀明);华侨大学工学院(薛浩、柳培忠); 泉州医学高等专科学校母婴健康服务应用技术协同创新中心(吕国荣)

validation set images were 94.9% and 92.9%, respectively, which were highly consistent with the classification of expert–level sonographers (Kappa=0.888, 0.810). The accuracy of MSP and RNT of AI model recognition clinical validation set images were 94.0% and 84.4%, respectively, which were consistent with the classification of expert–level sonographers (Kappa=0.751, 0.690). The overall consistency between the evaluation results of MSP and RNT in clinical validation set images by AI model and expert–level sonographers were good (Kendall's W=0.760, 0.789). **Conclusion** AI model has high accuracy in recognizing fetal facial ultrasound screening sections in early pregnancy, and can be used as an auxiliary method for the training of primary and basic physicians and the evaluation of image quality.

KEY WORDS Ultrasonography; Early pregnancy; Screening section, facial; Artificial intelligence; Quality control

获得标准的早孕期胎儿颜面部筛查切面是产前超声早期诊断胎儿颜面部结构畸形的基础。近年来已有研究们证实应用人工智能(artificial intelligence, AI)可实现对中、晚孕胎儿颜面部超声标准切面的自动识别及图像质量控制,但其在早孕期胎儿颜面部超声筛查切面的自动识别及图像质量评价方面的应用鲜见报道。本研究旨在构建一种基于YOLOv4模型的早孕期胎儿颜面部超声筛查切面自动识别及图像质量评价系统,并验证其识别效能。

资料与方法

一、研究对象

收集2021年1月至2022年12月我院妊娠11~14周 胎儿颜面部超声筛查切面图像为研究对象,分为模型 构建集(1001张图像)和临床验证集(610张图像)。所 有图像均来自我院PACS系统,仪器使用Philips EPIQ7、 EPIQ 5、iU 22和GE Voluson E 8彩色多普勒超声诊断 仪(C5-1、C9-2或C1-5-D探头,频率1.0~5.0 MHz)。 纳入标准:①单胎,产后随访证实为健康胎儿,无颜面 部及其他结构畸形:②妊娠11~14周接受多次超声评 估者仅纳入研究1次;③纳入图像包括正中矢状面 (midsagittal plane, MSP)、鼻后三角切面(retronasal triangle, RNT)图像。排除标准:①由于孕妇肥胖、探头 抖动等原因导致图像模糊、拖尾,目标结构显示欠清 晰;②超声检查或产后证实胎儿异常。模型构建集图 像由具有10年以上产前超声筛查经验的超声医师依 据早孕期胎儿颜面部超声筛查切面标准进行筛选及 手工标注关键解剖结构,再由另一专家级产科超声 医师(获得产前诊断证书且从事产前超声诊断工作 15年以上的副主任医师)对标注的图像进行审核修 改。临床验证集图像由不同资历的超声医师在临床 工作中保存。将模型构建集图像按8:1:1分为训练 集、测试集和验证集,用于训练、测试及验证AI模型, 临床验证集用于评价AI模型的临床泛化能力。各数 据集分布见表1。

表1 早孕期胎儿颜面部超声筛查切面数据集分布 引

数据集	MSP	RNT	合计
模型构建集			
训练集	470	339	809
测试集	59	42	101
验证集	53	38	91
临床验证集	386	224	610

MSP:正中矢状面;RNT:鼻后三角切面

图像标准:①MSP,显示呈带状平滑强回声的额骨、鼻尖、鼻骨,呈带状不规则弧形强回声的硬腭,呈圆点状高回声的下颌骨,不能显示上颌突及眼眶;②RNT,显示鼻骨,双侧对称的上颌突,呈圆点状高回声的双侧上牙槽骨,双侧对称的下颌骨断面。见图1。

二、方法

1.参考临床手动评估早孕期胎儿颜面部超声筛查



A:MSP观;B:RNT观

图1 早孕期胎儿颜面部超声筛查标准切面示意图

切面的质量控制准则^[2-4],由3名具有20年以上检查经验的超声专家制定早孕期胎儿颜面部超声筛查切面质量控制协议,以及是否为标准切面判定标准和图像质量的评价标准。

2. 分析早孕期胎儿颜面部超声筛查切面图像的特 点及规律。本研究基于深度学习的自动识别算法模 型,对标准切面自动识别中使用的特征提取、特征分类 方法进行系统性研究和分析,然后通过YOLOv4模型[5] 从训练集图像中学习提取有识别性的特征,并通过联 合训练和交替优化,构建一个完全端到端的卷积神经 网络(CNN)对象检测模型,以此来检测关键解剖结构。 该AI模型根据早孕期胎儿颜面部超声筛查切面质量 控制协议对检测到的关键解剖结构进行赋值,根据评 分结果自动判断是否为标准切面,从而实现早孕期胎 儿颜面部超声筛查切面的识别分类。其中MSP选取4项 关键解剖结构进行赋值,分别为鼻骨、硬腭、第三脑 室、下颌骨,其中显示第三脑室计2分,其余每项计1分; RNT选取4项关键解剖结构进行赋值,分别为上颌突、 上牙槽骨、鼻骨、下颌骨,其中上颌突、下颌骨在该切面 中为对称结构,每侧计0.5分,两侧共1分,上牙槽骨计 2分,鼻骨计1分。见图2,3。标准切面判定标准:每 个切面总分达到4分及以上。图像质量评价标准:分 为优秀、合格、不合格,其中RNT优秀、合格、不合格对 应的评分分别为5分、4.5分、≤4分;MSP优秀、合格、不 合格对应的评分分别为5分、4分、≤3分。本研究在验 证集中进行是否为标准切面判定,在临床验证集中同 时进行是否为标准切面判定及图像质量评价。其中 AI模型识别关键解剖结构的评价指标为香准率、查全 率和F1分数。AI模型工作流程见图4。

3. 由专家级超声医师对临床验证集图像进行分

类,依据早孕期胎儿颜面部超声筛查切面的标准程度^[6]评价其等级,分为优秀、合格、不合格。

将带有标签的早孕期胎儿颜面部超声筛查切面 图像输入AI模型中自动检测关键解剖结构,然后根据 制定的早孕期胎儿颜面部超声筛查切面质量控制协 议评分准则对超声图像进行评分,并根据评分结果判 断是否为标准平面,同时对图像质量等级进行评价。

三、统计学处理

应用 SPSS 22.0 统计软件,计数资料以例或率表示,组间比较采用 χ 检验。AI模型与专家级超声医师识别 MSP 和 RNT的一致性分析采用Kappa 检验;AI模型与专家级超声医师评价图像质量等级的一致性分析采用Kendall 协同系数。P<0.05 为差异有统计学意义。

标准	分数	
鼻骨清晰显示	1分	
硬腭清晰显示	1分 —	
第三脑室清晰显示	2分	
下颌骨显示	1分 —	

图 2 早孕期胎儿颜面部超声筛查切面质量控制协议(图左)及对应的解剖结构在胎儿MSP图像(图右)中的位置

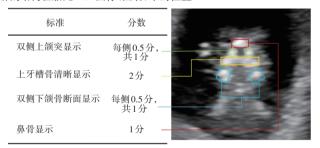


图 3 早孕期胎儿颜面部超声筛查切面质量控制协议(图左)及对应的解剖结构在胎儿RNT图像(图右)中的位置

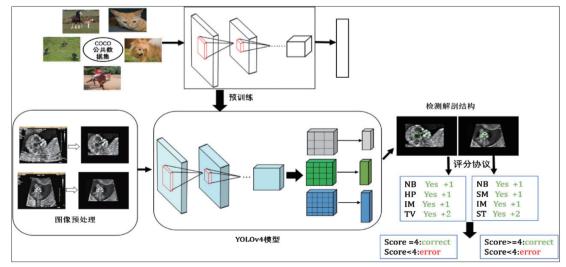


图4 AI模型工作流程图

结 果

一、AI模型识别测试集图像中各关键解剖结构的 性能评估

成功构建的AI模型识别测试集图像中各关键解剖结构的查准率、查全率和F1分数均>80%。见表2。

表 2 AI模型识别测试集图像中各关键解剖结构的性能 %

解剖结构	查准率	查全率	F1 分数
鼻骨	85.56	80.63	83
硬腭	99.12	99.12	99
下颌骨	90.87	87.40	89
第三脑室	99.08	95.58	97
上颌突	86.32	80.16	83
上牙槽骨	93.75	96.15	95

二、AI模型对验证集图像中MSP和RNT的识别效能及其与专家级超声医师分类的一致性分析

以专家级超声医师分类为标准,AI模型识别验证集图像中MSP和RNT的灵敏度分别为94.9%、96.8%,特异度分别为95.0%、81.8%,准确率分别为94.9%、92.9%。见表3。该模型识别验证集图像中MSP、RNT与专家级超声医师分类的一致性均好(Kappa=0.888、0.810)。

表3 AI模型对验证集图像中MSP和RNT的识别效能 %

切面	灵敏度	特异度	阳性预测值	阴性预测值	准确率
MSP	94.9	95.0	97.4	90.5	94.9
RNT	96.8	81.8	93.8	90.0	92.9

三、AI模型对临床验证集图像中MSP和RNT的识别效能及其与专家级超声医师分类的一致性分析

以专家级超声医师分类为标准,AI模型识别临床验证集图像中 MSP和 RNT的灵敏度分别为 95.7%、75.9%,特异度分别为 85.0%、98.8%,准确率分别为 94.0%、84.4%。见表4。该模型识别临床验证集图像中 MSP、RNT 与专家级超声医师分类的一致性均良好 (*Kappa*=0.751、0.690)。

四、AI模型与专家级超声医师在早孕期胎儿颜面 部超声筛查切面图像质量评价中的一致性分析

AI模型对临床验证集图像中 MSP 和 RNT 图像质量评价结果与专家级超声医师的总体一致性均良好

表4 AI模型对临床验证集图像中MSP和

%

RNT的识别效能

切面	灵敏度	特异度	阳性预测值	阴性预测值	准确率
MSP	95.7	85.0	97.2	78.5	94.0
RNT	75.9	98.8	99.1	70.7	84.4

(Kendall's W=0.760、0.789),但AI模型与专家级超声 医师在优秀、合格两个等级评价中分歧较大。见表5。

表 5 AI模型与专家级超声医师对早孕期胎儿颜面部 超声筛查切面图像质量评价结果

张 方法 优秀 合格 不合格 AI模型 MSP 225 95 66 108 75 RNT 41 专家级超声医师 35 307 MSP 44 RNT 135

讨 论

目前我国胎儿结构筛查仍主要集中在中孕期(妊 娠18~24周),随着超声技术、仪器性能的不断提高及 临床对早孕期胎儿结构研究的深入,早孕期胎儿超声 结构筛查逐渐受到重视。与其他部位结构比较,早孕 期胎儿颜面部结构具有器官多、体积小、准确识别难 度大等问题。超声检查时对早孕期胎儿颜面部标准 切面的准确获取、评价及其内部关键解剖结构的精准 识别是诊断颜面部结构畸形的基础。研究[7]显示初级 医师及社区、乡镇等基层医师对 MSP、眼眶切面等早 孕期胎儿颜面部超声筛查切面的重要性认识不足(认 为该切面对早孕期胎儿颜面部结构筛查有效者仅 75.1%、58.7%)。究其原因为这些医师对图像中关键 解剖结构识别困难。由于孕妇腹部脂肪厚度、胎儿 体位、羊水量、检查机器性能的影响,导致部分初级 及基层医师即使在认真学习了标准切面及模式图 后,临床工作中对超声图像中关键解剖结构的准确识 别仍存在较大困难。本研究结果显示, AI 模型对硬 腭、第三脑室和上牙槽骨的识别效果最好,查准率、查 全率和F1分数均>93%,识别下颌骨的查准率、查全率 和 F1 分数均>87%, 识别鼻骨的香准率、查全率和 F1 分数分别为85.56%、80.63%和83%;识别上颌突的查 准率、查全率和 F1 分数分别为 86.32%、80.16% 和 83%,表明AI模型在早孕期胎儿颜面部超声筛查切面 关键解剖结构的识别中具有优异的性能。目前AI识 别胎儿关键解剖结构的研究多集中在中、晚孕期[8-9], 而早孕期则侧重于对颈项透明层厚度的研究[10],在识 别早孕期关键解剖结构方面鲜见报道。鉴于AI模型 的优异性能及自动标注的高效性,经专家审核并通过 模型标注的海量超声图像能很好地辅助初级及基层 医师对早孕期胎儿颜面部超声筛查切面中关键解剖 结构的学习。

准确获取早孕期胎儿颜面部标准切面不仅需要 精准识别关键解剖结构,还需要超声医师不断操作练 习,而低年资超声医师的培训以往常需在专家级超声 医师的指导下进行,但这种培训模式存在高水平专家 资源严重匮乏和培养过程费时、费力等缺点。随着 AI 技术的发展,计算机辅助诊断系统则是另一种高性价 比的解决方案[11]。本研究中AI模型对验证集及临床 验证集图像 MSP、RNT 的分类均表现出较好的效能, 与专家级超声医师分类结果的一致性均良好(Kappa 值 分别为 0.888、0.810 和 0.751、0.690), 与以往 AI 识别 中、晚孕期胎儿颜面部超声筛查标准切面效果相 当[12]。因此在临床工作中有望将AI模型嵌入超声检 查设备中,以实现实时评价初级及基层医师获取的早 孕期胎儿颜面部超声筛查切面,督促检查医师不断调 整扫查角度,从而部分替代专家级超声医师,辅助初 级及基层医师标准切面的扫查训练。

标准的超声切面是准确诊断的基础,不同年资、 不同等级医院的超声医师获取的图像质量良莠不齐, 进行广泛的质量评价是提高图像质量的有效方 法[13-14]。人工评价存在费时、费力、主观性强、评审专 家缺乏等缺点,无法对超声图像进行广泛的质量评 价。而应用AI技术评价超声图像质量具有高效、客观 的优势。不同医院的资金实力、仪器性能、人员水平 均有差异,若使用统一的超声图像质量评价标准对基 层医院或许并不合理,同时胎儿检查过程中因肢体、 胎位、羊水等因素影响,采集优秀的标准切面往往十 分困难,而临床对疾病的诊断可能仅需要关键解剖结 构的清晰显示,并非需要所有的超声图像均为优秀的 标准切面,因此在超声质量控制评价中对超声图像进 行等级评定具有十分重要的现实意义。本研究结果 显示,AI模型对标准及非标准切面的分类能力与专家 级超声医师分类具有良好的一致性,表明该模型可较 好地识别早孕期胎儿颜面部超声筛查切面。同时,本 研究的AI模型通过对标准切面中关键解剖结构显示 进行评分,并根据评分结果将图像分为优秀、合格、不 合格,细化了超声图像质量评价准则,且 AI 模型与专 家级超声医师对临床验证集图像质量评价方面表现 出良好的一致性(MSP、RNT对应的Kendall's W分别 为0.760、0.789)。由此可见,具有评分功能的质量评 价系统可以应用于不同等级医院的超声图像质量评 价以达到不同的临床目的。

虽然本研究构建的 AI 模型在识别标准切面上已取得了较好效果,但仍然存在一些问题和不足:①鼻

骨、上颌突识别精度仍较低;②AI模型与专家级超声 医师评价临床验证集图像质量的一致性不够好(优 秀、合格两个等级评定中分歧较大);③部分图像等级 间区分度不够优秀,尚不能很好地应用于超声图像质 量评价。这启发了下一阶段的研究方向:①针对鼻 骨、上颌突识别精度不高、泛化能力较弱等问题,后续 会继续扩充数据集,针对解剖结构的空间位置关系改 进YOLOv4模型,继而提高对关键解剖结构的识别精 度;②增加关键解剖结构数目并对具体解剖结构及形 态进行评价,以增大各等级图像间的评分差异,实现 图像等级评价的完全自动化及改善部分等级间区分 度较差的现状。

综上所述,本研究构建的AI模型对早孕期胎儿颜 面部超声筛查切面的分类具有较高的准确性,可作为 初级医师及基层医师培训和图像质量评价的辅助方法。

参考文献

- [1] Yu Z, Tan EL, Ni D, et al. A deep convolutional neural network-based framework for automatic fetal facial standard plane recognition[J].IEEE J Biomed Health Inform, 2017, 22(3):874-885.
- [2] 李胜利.胎儿畸形产前超声诊断学[M].北京:科学出版社, 2017:89.
- [3] 姜玉新.中国胎儿产前超声检查规范[M].北京:人民卫生出版社, 2016:26-27.
- [4] International Society of Ultrasound in Obstetrics and Gynecology, Bilardo CM, Chaoui R, et al.ISUOG Practice Guidelines (updated): performance of 11-14-week ultrasound scan [J]. Ultrasound Obstet Gynecol, 2023, 61(1):127-143.
- [5 李柯泉,陈燕,刘佳晨,等.基于深度学习的目标检测算法综 述[J].计算机工程,2022,48(7):1-12.
- [6] 中国医师协会超声医师分会.产科超声规范化培训考核标准中国 专家共识(2022版)[J].中华超声影像学杂志,2022,31(5): 369-378.
- [7] 高艳多,杨小红,陈欣林,等.产前超声医生对早孕期胎儿超声筛查相关知识认知情况调查[J].中国产前诊断杂志(电子版), 2019,11(4):9-13.
- [8] Zhang B, Liu H, Luo H, et al. Automatic quality assessment for 2D fetal sonographic standard plane based on multitask learning [J]. Medicine(Baltimore), 2021, 100(4); e24427.
- [9] Plotka S, Klasa A, Lisowska A, et al. Deep learning fetal ultrasound video model match human observers in biometric measurements [J]. Phys Med Biol, 2022, 67(4):045013.
- [10] Walker MC, Willner I, Miguel OX, et al. Using deep-learning in fetal ultrasound analysis for diagnosis of cystic hygroma in the first trimester[J].PLoS One, 2022, 17(6); e0269323.
- [11] 杨晓云, 王丹, 康琦, 等. 人工智能计算机辅助系统联合超声造影 诊断 TI-RADS 5 类结节的价值[J]. 临床超声医学杂志, 2021, 23(3):226-228.

- [12] 刘中华,王小莉,吕国荣,等.人工智能自动识别胎儿颜面部超声标准切面的研究[J].中国医学物理学杂志,2021,38(12):1575-1578.
- [13] Wu L, Cheng JZ, Li S, et al. FUIQA: fetal ultrasound image quality assessment with deep convolutional networks [J]. IEEE Trans Cybern,
- 2017,47(5):1336-1349.
- [14] Hareendranathan AR, Chahal BS, Zonoobi D, et al. Artificial intelligence to automatically assess scan quality in hip ultrasound[J]. Indian J Orthop, 2020, 55(6):1535-1542.

(收稿日期:2023-01-31)

• 病例报道•

Right ventricular diverticulum diagnosed by transthoracic echocardiography combined with contrast-enhanced echocardiography: a case report 经胸超声心动图联合心脏超声造影诊断右室憩室1例

孙丹丹 景香香

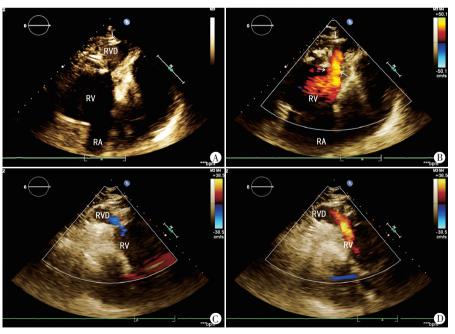
[中图法分类号]R540.45

「文献标识码]B

患者男,60岁,因"反复胸闷、气促1年余,加重1个月"就诊,既往高血压病史3年,规律服药,血压控制可。体格检查:心率84次/min,血压138/86 mm Hg(1 mm Hg=0.133 kPa),心浊音界扩大,三尖瓣区可闻及3级收缩期杂音。心电图提示:心房扑动,完全性右束支传导阻滞。经胸超声心动图检查:右室心尖处"囊袋样"无回声区向心尖凸起,颈部宽约6.5 mm,深约37.0 mm,与右室相通,无回声区周边可见心肌组织与右室心肌

相连续;CDFI于舒张期见右室腔内血流 进入无回声区内,收缩期见无回声区血流 进入右室内(图1)。超声心动图提示:右 室心尖处无回声区,考虑右室心尖部憩室 形成,建议行心脏超声造影进一步检查。 左心声学造影检查:右室心尖处局部呈 "囊袋样" 向外凸出,大小约35 mm×25 mm, 造影剂充盈其内,心内膜显示清晰,其周 围可见心肌组织。左心声学造影提示:右 室心尖部憩室形成(图2)。冠状动脉造 影检查:左主干未见狭窄,左前降支中段 狭窄85%,回旋支管壁欠规则(第二钝缘 支狭窄70%~80%),右冠状动脉近段狭窄 50%。冠状动脉造影提示:冠心病,3支血 管病变。后行冠状动脉搭桥、三尖瓣置换 及临时起搏导线安置术,术中发现右室心 尖部见一大小约20 mm×25 mm憩室,临床 诊断为右室心尖部憩室。

讨论:心脏憩室是一种罕见的先天性 心脏畸形,由于胚胎时期心内膜发育异常,局部心肌组织减少或缺失,导致房室 壁薄弱膨出[1]。心室憩室按组织类型可分为肌型憩室和纤维型憩室,其中肌型憩室与周围心室壁同步舒缩运动,纤维型憩室存在运动障碍^[2]。该病好发于心室心尖部,左室发病多于右室,发病率分别约3.4%和0.6%^[2],多见于儿童,常伴有其他先天性心脏畸;成人较少见,多为孤立性憩室^[3],常无明显临床症状,多于超声心动图检查时偶然发现。本例为老年患者,因胸闷、气促就诊,超声心动图检查发现右室憩室,左心声学造影示其内见心肌



A、B:心尖部无回声区,与右室相通(箭头示);C、D:舒张期见右室腔内血流进入无回声区内,收缩期见无回声区血流进入右室内。RA:右房;RV:右室;RVD:右室憩室

图1 本例患者经胸超声心动图表现

(下转第779页)