

人工智能质控在提高胎儿上腹部水平横切面标准率中的应用价值

彭桂艳 谭莹 曾晴 罗丹丹 黄文兰 江瑶 温昕 李胜利

(南方医科大学附属深圳市妇幼保健院 超声科, 广东 深圳 518000)

【摘要】 目的 探讨人工智能(artificial intelligence, AI)质量控制系统在提高胎儿上腹部水平横切面超声图像标准率中的应用价值。**方法** 应用“产前超声 AI 智慧云平台”对深圳市 60 家医院三个季度在产科超声检查中存储的胎儿上腹部水平横切面图像共 18114 张进行智能质控,受质控医生通过查阅质控结果及图像存在的不足之处针对性地加以改进。应用 χ^2 检验对三个季度图像的标准率、不足原因及申诉情况进行两两比较,评估智能质控对于提高胎儿上腹部水平横切面标准程度的价值。**结果** 各医院第一、二、三季度胎儿上腹部水平横切面图像标准率分别为 80.15%(5649/7048)、86.2%(4391/5096)、90.55%(5406/5970),各季度图像标准率逐步提高,两两比较差异有统计学意义($P < 0.01$);每个季度的不标准图像中,图像不足的主要原因是切面中出现了大片肺脏,分别为 8.9%(631/7048)、7.0%(358/5096)、4.9%(294/5970),其次是脐静脉和门静脉汇合部显示不清,分别为 4.3%(305/7048)、3.5%(181/5096)、2.6%(155/5970),这两类图像均显著减少,两两季度比较差异有统计学意义($P < 0.01$);医生申诉后维持智能评价结果的图像分别为 66.4%(79/119)、49%(25/51)、33.3%(10/30),两两比较差异有统计学意义($P < 0.01$),提示各季度医生对图像认识错误率逐渐降低。**结论** 人工智能质量控制可有效提高胎儿上腹部水平横切面超声图像的标准率,提升医生对标准切面的认识。

【关键词】 人工智能; 上腹部水平横切面; 胎儿; 标准切面

【中图分类号】 R445.1 **【文献标识码】** A

Value of artificial intelligence quality control in improving the standard rate of transverse section of fetal abdomen

Peng Guiyan, Tan Ying, Zeng Qing, Luo Dandan, Huang Wenlan, Jiang Yao, Wen Xin, Li Shengli
Department of Ultrasound, Affiliated Shenzhen Maternity & Child Healthcare Hospital, Southern Medical University, Shenzhen 518028, China

【Abstract】 Objective To explore the application value of the artificial intelligence (AI) quality control system in improving the standard rate of transverse section of fetal abdomen. **Methods** The “Intelligent ultrasonic quality control system” was applied to perform quality control in 18,114 transverse sections of fetal abdomen images stored in 60 hospitals in Shenzhen in three quarters. The doctors subjected to quality control can make targeted improvements by referring to the quality control results and the shortcomings of the images. Image standard rate, reasons for the insufficiency and the instances of the complaint in three quarters were compared by χ^2 test. **Results** The standard rates of transverse section of abdomen in the first, second and third quarters were 80.15% (5649/7048), 86.2% (4391/5096) and 90.55% (5406/5970), respectively. The image standard rate has been gradually increased in each quarter, and the comparative difference between every two quarters was statistically significant ($P < 0.01$). The main

DOI: 10.13470/j.cnki.cjpd.2022.04.002

基金项目:国家重点研发计划资助(2022YFF0606301);深圳市科技计划项目(JCYJ20210324130812035)

reason for the non-standard images in each quarter was the presence of lung in the section (8.9% (631/7048), 7.0% (358/5096), and 4.9% (294/5970), respectively). Another reason for the non-standard images was that umbilical vein and portal vein were not shown in the section (4.3% (305/7048), 3.5% (181/5096), 2.6% (155/5970)). Both reasons for the non-standard images were significantly reduced ($P < 0.01$). The images that maintained the results of AI evaluation after doctors' complaints were 66.4% (79/119), 49% (25/51) and 33.3% (10/30), respectively, and the difference was statistically significant (pairwise comparison between groups, $P < 0.01$), indicating that the error rate of doctors' understanding of images decreased gradually. **Conclusion** The standard rates of transverse section of the fetal abdomen could be improved effectively by AI quality control, and doctors' understanding of the standard section is also improved.

【Key words】 Artificial intelligence; Transverse section of the abdomen; Fetals; Standard section

腹围作为超声评估胎儿体重的重要参数之一,在胎儿体重评估中起着非常重要的作用。腹围测量不准确将直接导致胎儿体重评估出现偏差,临床上常见严重高估或低估胎儿体重的情况,从而增加不必要的医疗处理,加重孕妇身心负担及伤害。获取标准的上腹部水平横切面是准确测量腹围的前提^[13-15]。然而当前各医院部分超声医生对于该切面的认识不足,以致于所留图像不标准,无法满足临床对于准确评估胎儿体重的需求,因此图像质量控制至关重要。人工质控工作量大,效率低,且质控标准不统一,随着人工智能技术在医学图像领域的飞速发展,利用人工智能进行图像质量控制或许能够较好地解决以上问题^[6]。本文主要探讨人工智能(artificial intelligence, AI)对胎儿上腹部水平横切面图像质量控制中的应用价值。

1 资料与方法

1.1 对象 选取2021年深圳市60家医院第一、二、三季度20~40孕周胎儿上腹部水平横切面超声图像共18 114张,参与医生538位。纳入标准:常规超声检查中未发现明显结构异常的单胎。

1.2 方法

1.2.1 使用平台 使用深圳市卫生健康委统一招标采购的“产前超声 AI 智慧云平台”进行产科超声图像质量控制,该智能质控软件搭载在超声智库APP上,可供手机查看质检结果。通过“结果分析”可非常方便地查看图像质量评价,评价内容包括切面名称、切面质量、切面分数、结构评价、不足原因等内容,通过“学习提升”和软件提供的标准图像及模式图进行对比学习,同时还提供视频与文献学习。

1.2.2 智能质控 各医院超声科质控员分别于

2021年第一、二、三季度向“产前超声 AI 智慧云平台”上传本科室所有超声医生本季度中晚孕产科超声图像病例各10例中的所有上腹部水平横切面,由该平台中搭载的智能质控模型对上传图像进行智能质控,模型对于切面中需要显示的关键结构显示外接矩形框,并根据置信度作出评分^[11,12],对于非标准图像提示不足原因,医生通过查看质控结果提高标准切面的认识。

根据国内外的产科超声检查指南^[1-3]中对于上腹部水平横切面的具体要求,标准切面需清楚显示胃泡、脐静脉和门静脉汇合部、脊柱及完整的腹壁,只出现一根肋骨或不出现肋骨,同时不能出现肺脏及肾脏结构(图1)。结合本模型的质控方式,本研究对切面内关键结构的显示情况进行赋分,通过图像的总评分对该切面进行标准程度分级,具体见表1。

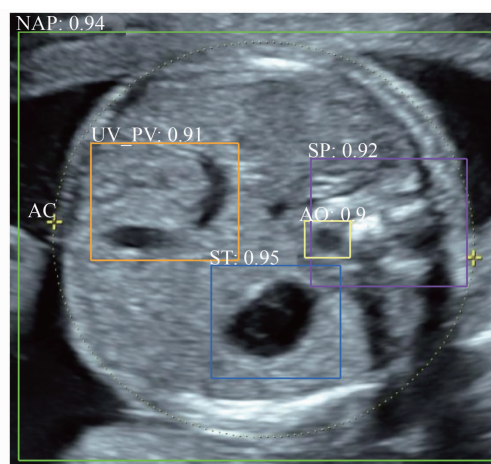


图1 上腹部水平横切面标准图像

注:要求清楚显示胃泡(ST),脐静脉在汇入门静脉处(UV_PV)要清楚显示弧形血管结构,不能显示肺和肾脏。
NAP为腹壁,SP为脊柱,AO为腹主动脉,AC为腹围。

表1 上腹部水平横切面评判规则

结构显示情况	得分(分)
胃泡	
显示清楚	2
显示基本清楚	1
未显示	-2
脐静脉和门静脉汇合部	
显示清楚	2
显示基本清楚	1
未显示	-2
肺脏	
未显示	0
显示大片肺	-4
肾脏	
未显示	0
显示	-4

注:4分:标准;2~3分:基本标准;<2分:非标准。

1.2.3 质控申诉 医生如对质控结果有任何疑问均可以在系统进行申诉。申诉发起后由资深专家团队重新进行人工审核,对于系统判断错误的图像,专家团队予以修正;对于系统判断正确的图像,则维持系统原判,同时专家通过回复疑问以修正医生对于切面认识存在的误区。

1.3 统计学分析 采用 SPSS 23.0 统计软件进行统计分析。计数资料以频数及百分率表示。各季度上腹部水平横切面标准率、非标准图像不足原因及申诉图像的结果两两比较均采用 χ^2 检验。 $P < 0.01$ 为差异有统计学意义。

2 结果

2.1 图像标准程度统计 深圳市各医院第一、二、三季度应用智能质控上腹部水平横切面的总图像数分别为 7048、5096、5970 张,各季度图像标准率逐步

提升,图像非标准率逐渐降低,两两比较差异均有统计学意义($P < 0.01$,详见表2)。

表2 上腹部水平横切面三个季度图像标准率比较[张(%)]

季度	图像总数	标准图像	基本标准图像	非标准图像
1	7048	5649(80.15) ^a	396(5.62) ^a	1003(14.23) ^a
2	5096	4391(86.20) ^b	150(2.95) ^b	553(10.85) ^b
3	5970	5406(90.55) ^c	73(1.23) ^c	491(8.22) ^c

注:a表示第1季度与第2季度相比较存在统计学意义($P < 0.01$);b表示第2季度与第3季度相比较存在统计学意义($P < 0.01$);c表示第1季度与第3季度相比较存在统计学意义($P < 0.01$)。

2.2 不足原因统计 第一、二、三季度非标准图像中出现大片肺脏和肾脏的图像均明显在减少,各季度两两比较差异均有统计学意义($P < 0.01$);未显示胃泡的图像总量不多,第一、二季度及第二、三季度两两比较差异不显著,但第一、三季度比较则差异有统计学意义($P < 0.01$);未显示脐静脉和门静脉汇合部的图像第一、第二季度比较无显著差异,但第二、三季度及第一、三季度两两比较,差异有统计学意义($P < 0.01$), (详见表3)。部分基本标准及非标准图示见图2、图3。

表3 上腹部水平横切面三个季度非标准图像

不足原因比较[张(%)]					
季度	图像总数	出现肺脏	出现肾脏	胃泡显示不清	脐静脉和门静脉汇合部显示不清
1	7048	631(8.9) ^a	136(1.9) ^a	85(1.2)	305(4.3)
2	5096	358(7.0) ^b	51(1.0) ^b	48(0.9)	181(3.5) ^b
3	5970	294(4.9) ^c	24(0.4) ^c	38(0.6) ^c	155(2.6) ^c

注:a表示第1季度与第2季度相比较存在统计学意义($P < 0.01$);b表示第2季度与第3季度相比较存在统计学意义($P < 0.01$);c表示第1季度与第3季度相比较存在统计学意义($P < 0.01$)。

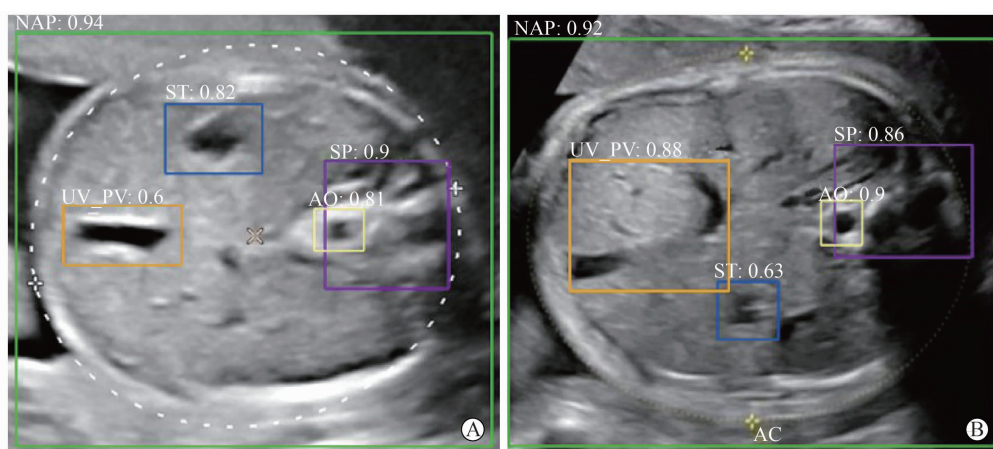


图2 上腹部水平横切面基本标准图像

注:腹壁(NAP)、脊柱(SP)、胃泡(ST)、脐静脉和门静脉(UV_PV)显示,但存在以下问题。A:未显示脐静脉与门静脉汇(UV_PV)合部的弧形血管结构;B:胃泡显示模糊不清。

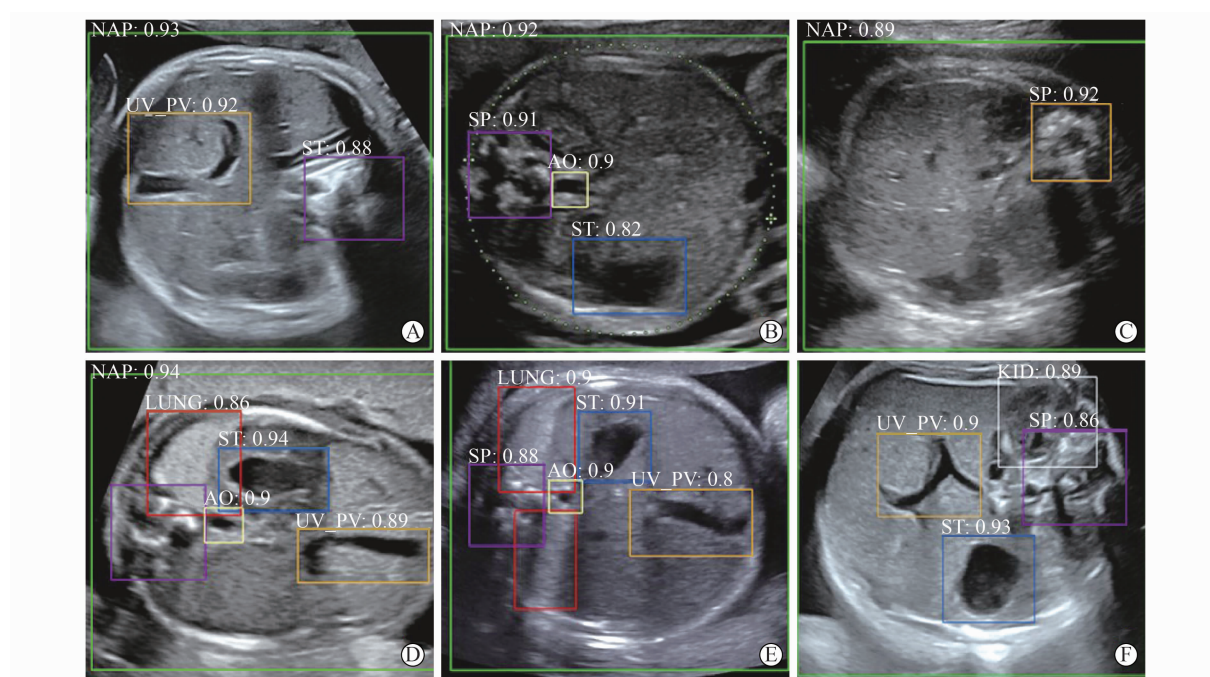


图3 上腹部水平横切面非标准图像

注:腹壁(NAP)、脊柱(SP)清楚显示,但存在以下问题。A:胃泡显示不清;B:脐静脉和门静脉显示不清;C:胃泡、脐静脉和门静脉均显示不清;D:近场出现大片肺脏(LUNG);E:两侧均出现大片肺脏;F:出现了肾脏(KID)。

2.3 申诉图像统计 第一、二、三季度申诉图像中,维持原判图像的申诉率明显降低,其中第一、三季度比较差异有统计学意义($P < 0.01$,详见表4)。

表4 上腹部水平横切面三个季度申诉图像比较[张(%)]

季度	申诉图像数	系统判断错误图像	维持原判图像
1	119	40(33.6)	79(66.4)
2	51	26(51.0)	25(49.0)
3	30	20(66.7)*	10(33.3)*

注: * 表示第1季度与第3季度相比较存在统计学意义($P < 0.01$)。

3 讨论

标准的产科超声切面图像是生物学测量准确和疾病诊断正确的基础^[4]。传统的图像质控主要由各医院的高年资医师作为质控员对科室医师进行抽查,且只能对被抽查医生的少量病例进行人工质控。诚然,该质控方式对于图像质量的提升起到了一定程度的作用,但是由于不同医师各自经验和认知不同,对于标准切面的理解也存在差异,导致不同单位、不同质控员对图像进行质控时标准均不统一。为解决此问题,一些学者^[7-8]提出应用预设评分标准的方法对图像质量进行客观评价,结果证实该方法

是行之有效的,但同时,耗费人力物力和时间的弊端依然存在,最终导致图像质控工作无法大规模全员进行。随着人工智能技术在超声图像分析领域的飞速发展,Wu等^[9-10]与我们团队合作尝试应用人工智能技术对胎儿上腹部水平横切面进行图像质量控制,实验结果非常可观,其准确性可达到与人类专家相当水平,但该研究尚未达到实时判断的效果,且仅停留于实验阶段。本研究使用的质控平台“产科超声智慧云平台”由深圳市妇幼保健院和湖南大学共同自主研发,其应用统一的切面质控标准,明显克服了人工质控的主观性和随意性,同时该系统检测图像的效率相当高,约为0.06s/张,已经达到实时质控要求,充分利用该系统能大大提高产科超声图像的质量质控水平,让质控变得更规范、统一和高效。

本研究结果显示上腹部水平横切面的图像标准率从一季度的80.15%逐步提高至第三季度的90.55%,非标准率同步逐渐下降,表明通过本平台进行质控能显著提升医生留存该切面的标准程度。分析其原因在于,通过质控,医生更清楚地了解自身留存图像的问题所在,并在日后的工作中进行针对性改进;同

时自主质控的意识得到进一步增强,这从实质上改善了各自的图像质量。

分析本研究中的非标准图像发现,肺脏的出现在三个季度中占比最高,尤以第一季度最为明显,此类图像占总图像的8.9%。当上腹部水平横切面出现大片肺脏时,表明切面向胎儿头侧偏斜,若在此切面测量腹围,则会导致腹围测值偏大,进而导致胎儿体重被高估^[5]。部分医生并没有意识到标准的上腹部水平横切面不能出现大片肺脏,所以留存切面时容易忽视;同时,也有部分医生对肺脏的认识不足,常把肺脏误认为肝脏或者脾脏而把此类图像当成了标准图像。AI质控时通过外接矩形框把切面中出现的肺标注出来,医生在查看图像时能清楚地认识到该切面不能显示肺脏,还能对肺的位置及声像特点有更清晰的认识,避免下次留存切面时出现类似的错误。因此出现肺脏的图像逐季度减少,到第三季度降低至4.9%。类似的情况也表现于切面中肾脏的显示上,出现肾脏时表明切面向胎儿足侧偏侧,同样可导致腹围测值的偏大。本研究显示,出现肾脏的图像也逐季度减少,由第一季度的1.9%降低至0.4%。脐静脉和门静脉汇合部是另一个容易被忽视的结构,标准切面上要求这一结构显示为“J”型,若完全不显示或仅显示脐静脉腹内段,则表明切面出现不同程度的偏斜^[5]。通过质控并学习,该类图像在一季度的4.3%下降至第三季度的2.6%,表明医生对于这个细节的认识得到提升。未显示胃泡的图像虽然不多,但是经过智能质控后也明显减少,由1.2%下降至0.9%。以上表明,通过利用该智能质控平台,医生对于标准切面的结构有了更深的认识,切面标准率必然得到提高。

众所周知,智能质控准确性即便能达到人类专家相当水平,仍无法避免存在一些错误。本研究申诉结果显示,医生对于质控结果或存在异议或者疑问,但占比不高,三个季度的申诉率分别为1.7%、1%、0.5%,表明医生对于智能质控结果的认可度是相当高的。在申诉的图像中,维持智能系统原判的图像可认为是医生对于切面的理解不准确所致,第一季度这类图像占比高达66.4%,占申诉图像的大部分。通过后台维护的专家团队的线上书面反馈及

答疑,医生对于切面有了新的认识,这是医生再学习的过程。二、三季度这类图像的申诉显著减少,三季度明显降低至33.3%。受质控医生通过申诉过程可直接与专家团队进行“对话”,无疑是一个与专家交流的好机会,对于偏远地区及基层医院的医生来说更是一个很难得的学习交流平台。

综上所述,智能质控的应用明显克服了人工质控的种种弊端,通过对医生留存图像的标准程度进行智能评价并对不足之处进行反馈,能极大地提升图像质量,推动图像质量控制工作跨越新的里程。本研究主要是通过对医生已经留存好的图像进行智能质控从而提高图像标准率,智能质控的提升空间或许可以尝试在医生留存图像时对图像进行实时质控,这将是进一步持续改进的方向。

参 考 文 献

- [1] 中华医学会超声医学分会妇产超声学组,国家卫生健康委妇幼司全国产前诊断专家组医学影像组. 超声产前筛查指南[J]. 中华超声影像学杂志,2022,31(1):1-12.
- [2] SALOMON L J, ALFIREVIC Z, BERGHELLA V, et al. ISUOG Practice Guidelines (updated): performance of the routine mid-trimester fetal ultrasound scan[J]. Ultrasound in Obstetrics & Gynecol,2022,59(6):840-856.
- [3] 李胜利,文华轩. 产前超声检查规范化应用[J]. 中国实用妇科与产科杂志,2015,31(09):818-822.
- [4] 李胜利,文华轩. 中孕期胎儿系统超声检查切面及临床意义(续)[J]. 中华医学超声杂志(电子版),2010,7(04):617-643.
- [5] 李胜利,罗国阳. 胎儿畸形产前超声诊断学[M]. 北京:科学出版社,2017:11-26.
- [6] 陈彬,严雨霖,杜海雯,等. 人工智能在产科超声中的应用进展[J]. 医学综述,2022,28(14):2903-2907.
- [7] URSEM NICOLETTE T C, PETERS INGRID A, KRAAN-VAN DER EST MIEKE N, et al. An Audit of Second-Trimester Fetal Anomaly Scans Based on a Novel Image-Scoring Method in the Southwest Region of the Netherlands[J]. Journal of Ultrasound in Medicine, 2017,36(6):1171-1179.
- [8] KUMAR AMC, SHRIRAM KS. Automated scoring of fetal abdomen ultrasound scan-planes for biometry[R]. Brooklyn, NY, USA: IEEE 12th International Symposium on Biomedical Imaging (ISBI),2015:862-865.
- [9] 吴凌云. 胎儿腹部超声图像自动质量控制与参数测量[D]. 深圳:深圳大学,2017.

(下转 31 页)

道,进一步丰富了我国南方地区的异常血红蛋白类型数据库,有助于增强和提高医务工作者对异常血红蛋白的认识和重视。

参 考 文 献

- [1] 曾溢滔. 血红蛋白疾病的诊断和治疗[J]. 中华血液学杂志, 1996,17(8):393-394.
- [2] 徐湘民. 血红蛋白病的人群筛查和产前诊断[J]. 海南医学, 2019,30(S1):47-55.
- [3] 秦良谊. 我国异常血红蛋白发生率、分布及遗传多态性[J]. 医学研究通讯,2003,(12):12-14.
- [4] 李春莉,杨梅,李秋红. 重庆地区 34 800 例育龄期夫妇异常血红蛋白病分析[J]. 中国实验血液学杂志,2020,28(4):1316-1320.
- [5] 徐湘民. 地中海贫血预防控制操作指南[M]. 北京:人民军医出版社. 2011: 11- 13.
- [6] 杜丽,秦丹卿,王继成,等. 异常血红蛋白合并地中海贫血的基因诊断及血液学指标分析[J]. 中国产前诊断杂志(电子版), 2019,11(1):34-37.
- [7] 王继成,郭浩,黄华洁,等. 罕见异常血红蛋白的诊断和临床特征分析[J]. 中国实验血液学杂志,2020,28(6):2028-2032.
- [8] 秦良谊. 我国 β 链异常血红蛋白发生率、基因频率和分布[J].

医学研究通讯. 2004,(1):17-18.

- [9] J B CLEGG, M A NAUGHTON, D J WEATHERBALL. Abnormal human haemoglobins. Separation and characterization of the alpha and beta chains by chromatography, and the determination of two new variants, hb Chesapeake and hb J (Bangkok) [J]. J Mol Biol, 1966;19 (1):91-108.
- [10] 李育敏,张水兰,阚丽娟,等. 深圳地区异常血红蛋白病的分子特征与表型分析[J]. 临床检验杂志,2018,36(11):873-875.
- [11] 胡昕听,郭浩,秦丹卿,等. 异常血红蛋白 Hb J-Bangkok 的分子诊断及血液学特征分析[J]. 临床检验杂志,2020,38(3): 171-174.
- [12] R G SCHNEIDER, M E HAGGARD, C W MCNUTT, et al. Hemoglobin G-Coushatta: A New Variant in An American Indian Family [J]. Science, 1964;143(3607):697-698.
- [13] VICHINSKY E, RODGERS GP, RACHMILEWITZ E. Hemoglobinopathies [J]. Curr Mol Med, 2008;8(7):591.

(收稿日期:2022-04-07)

编辑:宋文颖

(上接第 10 页)

- [10] WU L ,CHENG J Z ,LI S , et al. FUIQA: Fetal Ultrasound Image Quality Assessment With Deep Convolutional Networks[J]. IEEE Transactions on Cybernetics, 2017, 47 (5):1336-1349.
- [11] B RAHMATULLAH,I SARRIS, A PAPAGEORGHIOU. Quality control of fetal ultrasound images:Detection of abdomen anatomical landmarks using AdaBoost[R]. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Chicago,IL,USA, 2011:6-9.
- [12] ZHANG BO, LIU HAN, LUO HONG, et al. Automatic quality assessment for 2D fetal sonographic standard plane based on multitask learning [J]. Medicine, 2021, 100 (4): e24427.

- [13] WRIGHT D, WRIGHT A, SMITH E, et al. Impact of biometric measurement error on identification of small- and large-for-gestational-age fetuses[J]. Ultrasound in Obstetrics & Gynecology,2020,55(2):170-176.
- [14] 姚怀齐,刘璇芝,王帆. 超声测量胎儿腹围预测出生体重的探讨[J]. 中国现代医药杂志,2020,22(11):42-44.
- [15] 赵旌,樊慧妮,张天啸. 探讨胎儿腹围在预测巨大儿评估中的临床价值[J]. 延安大学学报(医学科学版),2019,17(4):58-61.

(收稿日期:2022-09-02)

编辑:刘邓浩