❖专论

Comparison on artificial intelligence supplementary training (AIST) and traditional manual training (TMT) for standardized training of prenatal ultrasound standard sections

NIU Lei[1], *HE Xiaoqin*[1], *LIANG Bocheng*[2], *HONG Liwei*[3], *LIAN Min*[1], *HONG Sijie*[1], *LIAO Huifang*[1], *ZHOU Ping*[1], *LI Wenting*[1], *LI Shengli*[2], *ZHONG Xiaohong*[1*]

(1. *Department of Obstetrics and Gynecology, Ultrasound Medicine Division*, 3. *Department of Science and Education, Xiamen University Affiliated Women and Children's Hospital* [*Xiamen Maternal and Child Health Hospital*], *Xiamen* 361000, *China*; 2. *Department of Ultrasound, Shenzhen Maternity and Child Healthcare Hospital, Women and Children's Medical Center, Southern Medical University, Shenzhen* 518028, *China*)

[**Abstract**] **Objective** To compare the effectiveness of artificial intelligence supplementary training (AIST) and traditional manual training (TMT) for standardized training of prenatal ultrasound standard sections. **Methods** Six physicians without prenatal ultrasound experience (trainees) were selected and randomly divided into AIST group (using artificial intelligence [AI] ultrasound workstation) and TMT group (guided by a mentor with over 10 years of experience in a one-on-one manner) after undergoing prenatal ultrasound theory, instrument and operational technique teaching and passing baseline level assessment. Each trainee completed 15 training sessions, with the goal of acquiring 28 standard sections across 22 types within 15 min per session. Then the trainees were evaluated independently of AI or mentor to observe the training effect, and indicators such as standardization rate, completeness rate and repetition rate were calculated and compared between groups. **Results** Before training, no significant difference of baseline levels was found among trainees (all $P>0.05$). During and after training, standardization rate (67.30% *vs*. 45.95%) and completeness rate (74.60% *vs*. 55.08%) in AIST group were both significantly higher than those in TMT group (both $P<0.01$). During training, repetition rate in AIST group was 0, and trainees in AIST group progressed faster in the first 4 training sessions. After training, standardization and completeness rates in both groups were significantly higher than those before training (both $P<0.05$), and the absolute improvement value in AIST group was higher than that in TMT group. **Conclusion** AIST was more effective than TMT for standardized training of prenatal ultrasound standard sections.

[**Keywords**] ultrasonography, prenatal; standard sections; artificial intelligence; training

DOI: 10.13929/j.issn.1003-3289.2025.12.003

对比人工智能辅助培训(AIST)与传统人工培训(TMT)用于产前超声标准切面规范化培训

牛　蕾[1],何晓琴[1],梁博诚[2],洪礼伟[3],练　敏[1],洪思婕[1],廖慧芳[1],周　萍[1],李雯婷[1],李胜利[2],钟晓红[1*]

[1. 厦门大学附属妇女儿童医院(厦门市妇幼保健院)妇产科学系超声医学科,3. 科教部,福建 厦门　361000;2. 南方医科大学妇女儿童医学中心深圳市妇幼保健院超声科,广东 深圳　518028]

[**摘　要**] **目的**　对比人工智能辅助培训(AIST)与传统人工培训(TMT)用于产前超声标准切面规范化培训的效果。**方**

[**基金项目**] 国家重点研发计划(2022YFF0606301)。

[**第一作者**] 牛蕾(1985—),女,吉林长春人,硕士,主治医师。研究方向:超声诊断。E-mail: 644392761@qq.com

[**通信作者**] 钟晓红,厦门大学附属妇女儿童医院(厦门市妇幼保健院)妇产科学系超声医学科,361000。E-mail: 1369946030@qq.com

[**收稿日期**] 2025-08-24　[**修回日期**] 2025-11-02

法　对 6 名无产前超声经验医师（学员）进行产前超声理论、仪器及操作技术教学，待其通过基线水平评估后随机划分 AIST 组[使用人工智能（AI）超声工作站]与 TMT 组（由工作 10 年以上导师进行一对一指导），完成目标为在 15 min 内采集 22 种共 28 个标准切面的 15 次培训后，脱离 AI 或导师验证培训效果，计算标准率、完整率及重复率等并进行组间比较。**结果**　培训前 6 名学员基线水平差异均无统计学意义（P 均＞0.05）。培训期间及培训后 AIST 组标准率（67.30% *vs.* 45.95%）和完整率（74.60% *vs.* 55.08%）均显著高于 TMT 组（P 均＜0.01）。AIST 组在培训期间的重复率为 0，且在前 4 次培训中进步更快。完成培训后 2 组标准率和完整率均显著高于培训前（P 均＜0.05），且 AIST 组提高绝对值均高于 TMT 组。**结论**　AIST 用于产前超声标准切面规范化培训的效果优于 TMT。

[**关键词**]　超声检查，产前；标准切面；人工智能；培训

[**中图分类号**]　R715.3；R445.1　[**文献标识码**]　A　[**文章编号**]　1003-3289(2025)12-1950-04

超声可用于评估胎儿发育及诊断畸形，需依赖操作者经验与技术获取标准切面；规范化培训是提高超声检查质量的基础。传统人工培训（traditional manual training，TMT）由导师面授，主观性强、效率低；且近年出生缺陷防控形势严峻[1]，超声人才短缺、导师水平不一，接受规范化培训比例低[2]。人工智能（artificial intelligence，AI）用于超声诊断进展显著[3]，但用于教学尚处于探索阶段[4-6]。本研究对比观察 AI 辅助培训（AI supplementary training，AIST）与 TMT 用于产前超声标准切面规范化培训效果。

1　资料与方法

1.1　研究对象　以 2024 年 9 月入职厦门市妇幼保健院超声科、无产前超声和 AI 辅助使用经验的 6 名新就职医师为学员，男 1 名、女 5 名，年龄 26～32 岁、中位年龄 27 岁。受检对象为 2025 年 3 月 108 名接受孕中期超声筛查的单胎妊娠孕妇，年龄 20～40 岁、中位年龄 32 岁，孕周 22～26 周、中位孕周 24^{+1}周，体质量指数＜30 kg/m^2，无严重妊娠并发症（如胎儿畸形、胎盘早剥、羊水过多或过少等）。本研究获医院伦理委员会批准（KY-2025-103-K01），所有参与者均自愿并签署知情同意书。

1.2　仪器与方法　采用 Samsung W10 超声诊断仪及配备 AI 工作站。依据《产前超声规范化培训考核标准中国专家共识 2022 版》[7]指南标准进行培训，以《产期筛查和产前诊断质量控制指标》[8]作为参考标准。

1.2.1　切面　①目标切面：包括《产前超声规范化培训考核标准中国专家共识（2022 版）》所要求的 22 种共 28 个孕中期胎儿标准切面，即经丘脑横切面、经侧脑室横切面、经小脑横切面、双眼眶横切面、鼻唇沟冠状切面、颜面正中矢状切面、四腔心切面、左心室流出道切面、右心室流出道切面、三血管气管切面、胸腔矢状切面（双侧）、腹部横切面、脐带腹壁入口切面、膀胱水平横切面及双侧脐动脉血流图、双肾横切面、脊柱矢状切面（颈胸段及腰骶段）、肱骨长轴切面（双侧）、尺桡骨长轴切面（双侧）、股骨长轴切面（双侧）、胫腓骨长轴切面（双侧）、脐带胎盘附着处切面及宫颈管内口切面；目标切面数＝28 个目标切面×病例数；②其他切面：目标切面以外的切面；③重复切面：重复采集的目标切面（视为无效切面）；④有效切面：实际采集的目标切面（不包含重复切面）。目标为对每例均成功采集 28 个标准切面；切面存在重复时，针对其中最高分者进行统计分析。

1.2.2　分组　培训前对 6 名学员产前超声理论、仪器及操作技术教学，各使其采集 3 次目标切面，并上传至超声智慧云平台进行评分。之后随机划分 AIST 组与 TMT 组，每组 3 名。

1.2.3　培训方法及效果验证　AIST 组应用 AI 超声工作站，通过采集卡连接超声仪，同步显示超声仪的动态画面并进行实时动态分析；系统自动识别关键解剖结构并进行标注，即时为学员提供反馈，以强化其学习记忆，并以屏幕下方实时质控评分彩色条提醒其调整探头位置，精准定位关键解剖结构，直至获取高评分标准切面且无重复切面；TMT 组由具有 10 年以上操作经验的导师对学员进行一对一指导。对每名学员进行 15 次培训、每次 15 min，随机选择孕妇作为采集对象，AIST 组图像自动上传 AI 超声工作站，TMT 组图像存入超声诊断仪。培训结束后学员独立完成 3 次切面采集，记录相关信息并收集采集切面。

1.2.4　质控与评分　将切面上传至产前超声智慧云平台进行质控分析，平台自动对切面进行分类，并基于内部结构识别度进行精准评分[9]：≥60 分且＜80 分者为有效基本标准切面，≥80 分为有效标准切面。

1.3　统计学分析　采用 SPSS 27.1.0.1 统计分析软件。以 Pearson χ^2 检验比较分类变量及率，样本量＜40 时采用 Fisher 精确概率法。计算 2 组标准率（有效标准切面数/目标切面数）、完整率（有效切面数/目标切面数）、重复率（重复目标切面数/实采目标切面数）

及其他率(其他切面数/实采切面总数)并进行组间比较。$P<0.05$ 为差异有统计学意义。

2 结果

2.1 基线水平比较　6 名学员基线水平差异均无统计学意义(P 均>0.05),见表 1。

表 1　6 名学员基线水平比较[%(切面数)]

学员编号	标准率	完整率	其他率
1	10.71(9/84)	36.91(31/84)	22.50(9/40)
2	17.67(14/84)	44.05(37/84)	21.28(10/47)
3	10.71(9/84)	45.24(38/84)	22.45(11/49)
4	10.71(9/84)	44.05(37/84)	21.28(10/47)
5	11.9(10/84)	51.19(43/84)	18.87(10/53)
6	15.48(13/84)	51.19(43/84)	20.37(11/54)
P 值	0.743	0.435	0.998

2.2 培训期比较　培训期间每组各对 45 名孕妇进行采集。AIST 组标准率[67.30%(848/1 260) *vs.* 45.95%(579/1 260)]、完整率[74.60%(940/1 260) *vs.* 55.08%(694/1 260)]及其他率[52.31%(1 031/1 971) *vs.* 17.38%(187/1 076)]均明显高于 TMT 组(P 均<0.01),而重复率[0 *vs.* 21.93%(195/889)]低于 TMT 组($P<0.01$),见表 2。培训期间,随扫查次数增加,2 组采集切面标准率、完整率均逐渐提高,AIST 组前 4 次培训中标准率和完整率增加更快(图 1、2)。

表 2　6 名学员培训前、培训期间及验证时采集切面

培训阶段	项目	AIST 组	TMT 组
培训前	目标切面数(9×28)	252	252
	有效切面数	123	106
	有效标准切面数	32	32
培训期间	目标切面数(45×28)	1 260	1 260
	实采切面数	1 971	1 076
	实采目标切面数	940	889
	重复目标切面数	0	195
	其他切面数	1 031	187
	有效切面数	940	694
	有效基本标准切面数	86	102
	有效标准切面数	848	579
验证	目标切面数(9×28)	252	252
	实采切面数	281	232
	实采目标切面数	241	195
	重复目标切面数	56	48
	其他切面数	40	37
	有效切面数	185	147
	有效基本标准切面数	23	8
	有效标准切面数	160	137

2.3 培训后验证　培训结束后 2 组学员均采集 9 名孕妇图像进行效果验证。AIST 组标准率[63.49%(160/252) *vs.* 54.37%(137/252)]及完整率[73.41%

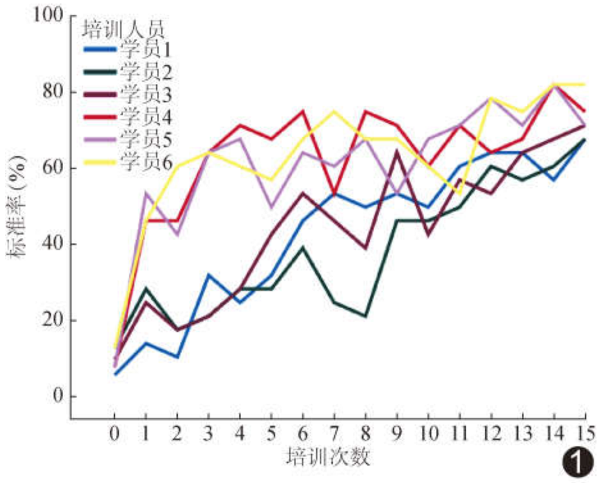

图 1　培训期间学员采集切面标准率图(学员 1~3 为 TMT 组;学员 4~6 为 AIST 组)

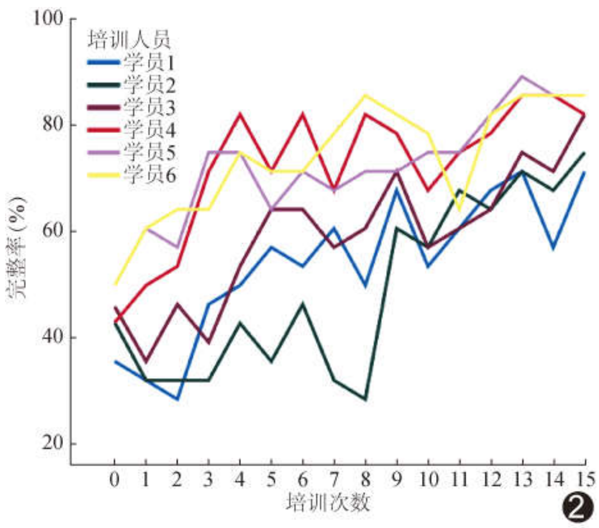

图 2　培训期间学员采集切面完整率图(学员 1~3 为 TMT 组;学员 4~6 为 AIST 组)

(185/252) *vs.* 58.33%(147/252)]均高于 TMT 组(P 均<0.05),而组间重复率[23.24%(56/241) *vs.* 24.62%(48/195)]及其他率[14.23%(40/281) *vs.* 15.95%(37/232)]差异均无统计学意义(P 均>0.05)。见表 2。培训后 2 组标准率、完整率均较培训前显著提高(P 均<0.05),且 AIST 组标准率、完整率提高绝对值均大于 TMT 组(表 3)。

3 讨论

AI 已在超声医学领域取得显著进展[10-12];深度学习和卷积神经网络等技术在图像分类、检测和分割目标任务中表现出色,成为 AI 辅助超声的基础[13-14]。本研究所用超声 AI 质控模型评价切面分类和标准程度的总体准确率达 99.7%及 92.1%,与人工评价总体符合率一致性较高[9,15]。本研究培训期间及验证时,

表 3　比较 6 名学员培训前、后采集标准切面效能

组别	项目	培训前［%（切面数）］	考核［%（切面数）］	P 值	提高绝对值（%）
AIST 组	标准率	12.70(32/252)	63.49(160/252)	<0.001	50.79
	完整率	48.81(123/252)	73.41(185/252)	<0.001	24.60
TMT 组	标准率	12.70(32/252)	54.37(137/252)	<0.001	41.67
	完整率	42.06(106/252)	58.33(147/252)	<0.001	16.27

AIST 组标准率和完整率均显著高于 TMT 组，培训前、后 AIST 组标准率和完整率提升幅度亦均大于 TMT 组，证实 AI 辅助对于提升学员采集超声标准切面能力较传统方法更具优势。

本研究 AIST 组应用 AI 超声工作站，通过实时质量反馈实现自动质控评分，基于自动择优替换机制实现零重复采集超声切面，有效解决了传统培训采集重复率高和效率低等问题；同时利用 AI 智慧云平台自动化体系进行评分，有效规避了人工评估的异质性。

培训期间 AIST 组采集切面标准率显著高于 TMT 组，提示学员在 AI 辅助下较易获得高质量标准切面：AI 通过自动识别解剖结构、实时质控和及时提醒使学员适时调整探头以获取高分切面。TMT 组导师难以实现持续性督导与即时反馈，学员对指导要点理解偏差易诱发紧张情绪而影响操作，使所获切面多局限于基本标准，却难以达到标准水平。AIST 组其他率显著高于 TMT 组，原因在于 AI 超声工作站筛查切面的种类远远超过质控要求的 22 种，还包含筛查需予以关注的 64 个其他重要切面，学员扫描过程中，AI 自动对这些切面进行识别与获取。值得注意的是，AIST 组重复率为零，原因系采集过程中 AI 自动用高分切面替换低分切面以有效避免重复存储。此外，培训期间及验证时 AIST 组完整率显著高于 TMT 组，证实 AI 辅助可显著提高检查效率。培训期间 2 组学员标准率、完整率均逐渐提高，且 AIST 组在前 4 次培训中进步速度更快，有助于缩短培训时间。培训后 AIST 组提升幅度高于 TMT 组，进一步证明 AIST 较 TMT 更具优势。

综上所述，AIST 用于产前超声标准切面规范化培训效果优于 TMT。本研究的主要局限性：为单中心、小样本量观察，难以避免选择偏倚，有待通过设立产前 AI 超声培训中心、增加受训人员数量、构建完整的周期性 AI 产前超声培训方案并予以推广。

利益冲突：全体作者声明无利益冲突。

作者贡献：牛蕾研究设计、图像分析和处理、数据分析、查阅文献、撰写和修改文章；何晓琴和梁博诚审阅文章；洪礼伟统计分析；练敏、洪思婕、廖慧芳、周萍和李雯婷研究实施；李胜利和钟晓红指导。

［参考文献］

[1]　卫生部发布《中国出生缺陷防治报告(2012)》[J].中国药房，2012，23(39)：3693.

[2]　董虹美，冉素真，陈婷，等.PDCA 循环在提升基层医院产前超声筛查短期培训效果的应用[J].医学教育研究与实践，2023，31(1)：104-108.

[3]　余翔，李胜利.人工智能在产前超声应用研究进展[J].实用妇产科杂志，2024，40(11)：866-870.

[4]　WITTEK A，STRIZEK B，RECKER F. Innovations in ultrasound training in obstetrics[J]. Arch Gynecol Obstet，2025，311(3)：871-880.

[5]　雷婷，郑巧，林美芳，等.人工智能辅助培训在产前超声教学中的实证研究[J].现代医院，2023，23(9)：1448-1450.

[6]　LEI T，ZHENG Q，FENG J，et al. Enhancing trainee performance in obstetric ultrasound through an artificial intelligence system: Randomized controlled trial[J]. Ultrasound Obstet Gynecol，2024，64(4)：453-462.

[7]　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.

[8]　国家卫生健康委办公厅关于印发产前筛查和产前诊断质量控制指标的通知[EB/OL].(2023-12-15)[2025-08-01].https://www.nhc.gov.cn/fys/c100078/202312/365527cd4c614965a81a231e46be94ae.shtml.

[9]　谭莹.产科超声图像的智能质量控制研究[D].广州：南方医科大学，2022：1-48.

[10]　郭文佳，罗丹丹，徐荟，等.新型智能产前超声技术 SFA 自动识别与获取胎儿标准切面[J].中国医学影像技术，2023，39(1)：65-69.

[11]　LIN M，HE X，GUO H，et al. Use of real-time artificial intelligence in detection of abnormal image patterns in standard sonographic reference planes in screening for fetal intracranial malformations[J]. Ultrasound Obstet Gynecol，2022，59(3)：304-316.

[12]　YANG C，LIAO S，YANG Z，et al. RDHCformer: Fusing ResDCN and transformers for fetal head circumference automatic measurement in 2D ultrasound images[J]. Front Med (Lausanne)，2022，9：848904.

[13]　CHEN H，WU L，DOU Q，et al. Ultrasound standard plane detection using a composite neural network framework[J]. IEEE Trans Cybern，2017，47(6)：1576-1586.

[14]　YU L，GUO Y，WANG Y，et al. Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks[J]. IEEE Trans Biomed Eng，2017，64(8)：1886-1895.

[15]　谭莹，文华轩，彭桂艳，等.产科超声图像智能质量控制系统的效能[J].中国医学影像技术，2022，38(9)：1361-1366.